# EDP non locales pour modéliser l'adaptation de virus et bactéries en milieu hétérogène : Application à l'évolution de la résistance aux traitements chez les organismes asexués.

Florian Lavigne






# AIX-MARSEILLE UNIVERSITÉ & INRAE
## ECOLE DOCTORALE 184

Institut de Mathématiques de Marseille / UMR 7373
Unité de Biostatistique et Processus Spatiaux / UR 546

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline : Mathématiques

Florian LAVIGNE

# EDP non locales pour modéliser l'adaptation de virus et bactéries en milieu hétérogène : Application à l'évolution de la résistance aux traitements chez les organismes asexués.

Modelling eco-evolutionary dynamics of viruses and bacteria in heterogeneous environments with nonlocal PDEs: Application in the treatment resistance of asexual organisms.

Soutenue le 22/09/2020 devant le jury composé de :

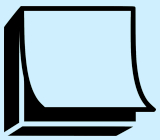
| | | |
|---|---|---|
| Sepideh Mirrahimi | Université Paul Sabatier | Rapporteuse |
| Grégoire Nadin | Sorbonne Université | Rapporteur |
| Jean Clairambault | Sorbonne Université | Examinateur |
| Florence Débarre | Sorbonne Université | Examinatrice |
| Etienne Klein | INRAE Avignon | Examinateur |
| François Hamel | Aix-Marseille Université | Directeur de thèse |
| Guillaume Martin | Université de Montpellier | Directeur de thèse |
| Lionel Roques | INRAE Avignon | Directeur de thèse |
| Guillemette Chapuisat | Aix-Marseille Université | Invitée |

Numéro national de thèse/suffixe local : 2020AIXM0284/014ED184

A ma famille et mes amis,

« Moi, si je devais résumer ma vie aujourd'hui avec vous, je dirais que c'est d'abord des rencontres. Des gens qui m'ont tendu la main, peut-être à un moment où je ne pouvais pas, où j'étais seul chez moi. Et c'est assez curieux de se dire que les hasards, les rencontres forgent une destinée ... »

Chabat, A. (Réalisateur), 2002. Astérix et Obélix : Mission Cléopâtre [film]. *Pathé*.



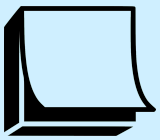
## Résumé

Le sauvetage évolutif est le processus par lequel une population déclinante évite l'extinction en s'adaptant génétiquement. Ce phénomène intervient dans de nombreux contextes biologiques, en présence d'un « stress » : invasion de nouveaux habitats, changement d'hôte chez les pathogènes, émergence de résistance aux pesticides, antibiotiques et fongicides. Il apparaît crucial de développer et de tester des modèles prédictifs de sauvetage évolutif, dans l'objectif de rationaliser et éventuellement d'optimiser les stratégies de gestion de pathogènes dans ces différents contextes. De nombreux problèmes associés au sauvetage évolutif, notamment chez les pathogènes, ont une composante spatiale importante (répartition de l'antibiotique dans l'organisme, des pesticides sur un paysage agricole, *etc.*). Ainsi la rationalisation de l'utilisation de traitements passe par une bonne compréhension et prédiction de l'effet de cette composante spatiale.

Lors de ma thèse, je me suis concentré sur les organismes asexués (pathogènes tels que lignées cancéreuses, virus et bactéries) et ai utilisé des modèles de paysages adaptatifs et des systèmes dynamiques déterministes et stochastiques. L'objectif était d'obtenir un cadre générique de modélisation, permettant de prédire des trajectoires démographiques, sous les effets combinés de la sélection, de la mutation, et de la dispersion des organismes dans un environnement spatialement hétérogène.

Ce cadre de modélisation s'est appuyé sur des équations intégro-différentielles ainsi que sur des EDP de transport non locales vérifiées par les fonctions génératrices des moments de processus stochastiques. *Via* l'analyse mathématique de ces équations, j'ai notamment cherché à calculer les probabilités de sauvetage évolutif ou le temps nécessaire à l'adaptation à un stress.

**Mots clés :** Dynamique Évolutive, EDP Non Locales, Équations Intégro-Différentielles, Processus Stochastiques, Sauvetage Évolutif.



# Abstract

Evolutionary rescue occurs when a population, initially declining because of exposure to a stress, avoids extinction *via* genetic adaptation restoring population growth. This phenomenon underlies a range of biological contexts of fundamental and applied importance: range expansions/contractions, host shifts in pathogens, and the emergence of resistance to herbicides, pesticides, fungicides, antibiotics or chemotherapy in cancer. Better management strategies of resistance emergence is a critical demand for health and agronomy. Such strategies would benefit from an empirically established and general modelling framework to understand and predict the adaptation of asexual organisms (*e.g.*, viruses or bacteria), under the combined effects of mutation and selection, in more or less complex situations, including the case of heterogeneous populations connected by migration.

The aim of this PhD is precisely to develop such a modelling framework and to carry out a rigourous mathematical analysis of the corresponding equations. In that respect, I have embeded a standard phenotype-fitness landscape model into a population genetics model of mutation selection and drift (with migration). This model is described *via* integro-differential equations and nonlocal transport equations satisfied by moment generating functions of some stochastic processes. The mathematical analysis of these equations leads to analytic formulae for evolutionary rescue probabilities and/or characteristic time until resistance emergence occurs.

**Keywords:** Evolutionary Dynamics, Evolutionary Rescue, Integro-Differential Equations, Nonlocal PDEs, Stochastic processes.

# Remerciements

Dans un premier temps, je souhaite remercier mes directeurs de thèse, François HAMEL, Guillaume MARTIN et Lionel ROQUES, de m'avoir guidé et conseillé. Aucun mot assez fort n'arriverait à exprimer la reconnaissance que j'ai envers eux. Grâce à eux, j'ai pu améliorer mes compétences en Mathématiques, que ce soit rigueur ou précision des arguments, mais aussi celles en Biologie. Je les remercie aussi pour la patience dont ils ont pu faire preuve, et enfin d'avoir accepter de m'encadrer pendant ces trois années passées. Je tiens aussi à mettre en avant Vincent CALVEZ, sans qui je n'aurai pas eu la chance de vivre cette expérience.

Je suis reconnaissant envers Sepideh MIRRAHIMI et Grégoire NADIN d'avoir accepté de rapporter ma thèse, mais aussi pour leurs commentaires, remarques et conseils. Je tiens aussi à remercier Jean CLAIRAMBAULT, Florence DÉBARRE et Etienne KLEIN d'avoir accepté de participer à mon jury.

Mes remerciements vous aussi pour mon comité de suivi de thèse : Jérôme COVILLE, Etienne KLEIN et Julien PAPAÏX pour leurs conseils et leurs remarques, leur bienveillance et leurs encouragements.

Lors de ma thèse, j'ai eu le plaisir de côtoyer trois laboratoires, l'unité BioSP (Avignon), l'I2M (Marseille) et l'ISEM (Montpellier). Je remercie ainsi les différents membres de chacun des laboratoires, que j'ai pu rencontrés.

À Avignon, j'ai eu le plaisir de partager des moments conviviaux autour d'une tasse de thé (ou café pour certains). Je remercie Sylvie et Amélie de leur aide pour toute l'aide qu'elles m'ont donnée pour les démarches administratives. Un grand merci aussi à Loïc, Jean-Loup et Jean-François pour leurs connaissances en Informatique (sans lesquelles j'aurai très certainement jeté mon ordinateur par la fenêtre). Je profite de ces quelques lignes pour glisser aussi un message amicale à tous les doctorants/docteurs, que j'ai pu rencontrés : Amel, Candy, Ismael, Julien,

Maria, Marie-Eve, Maryam, Patrizia ; et les stagiaires qui ont partagé mon bureau : Dominica et Julien.

À Aix-Marseille, je remercie Sonia, Nadia et Thierry pour les échanges que j'ai pu avoir en ce qui concerne ma thèse et mon après-thèse. Je remercie aussi toute l'équipe d'enseignement en PeiP que j'ai eu la joie de rencontrer sur le site de Saint-Jérôme, dont Fabienne, Chérifa et Olivier (mais aussi tous les autres enseignants et le personnel), qui m'ont encadré. J'ai eu grâce à eux l'opportunité de tenter la pédagogie inversée, qui s'est révélée forte utile lors du confinement. Bien que cette méthode désintéresse bon nombre d'enseignants, ils m'ont fait confiance pour la mettre en place, et cela a porté ses fruits : meilleur compréhension de la part des étudiants et un plus grand intérêt de leur part. Je remercie aussi mes étudiants de PeiP 2.4, avec qui j'ai passé d'excellents moments, et sans qui cette expérience n'aurait pas vu le jour. On refait un goûter quand vous le voulez (avec champagne ?).

Je remercie aussi toute l'équipe de *MT180* (Olivier, Delphine et tous les doctorants), que ce soit pour les exercices de respirations, ou pour les tortues qui trottent sur trois toits très étroits. Je tiens à vous présenter d'excises excuses pour tout le temps passé à m'aider à m'améliorer sur la diction et la vulgarisation. Cela m'a beaucoup aidé pour mes présentations, mais aussi pour mes cours. J'en profite aussi pour remercier aussi tous les jeunes doctorants/docteurs qui ont organisé la session de 2019 des *13 Minutes Jeunes Chercheurs* de Marseille et les participants. Je me souviendrai de tout le monde, mais surtout de notre mascotte Rémy.

À Montpellier, je remercie toute l'équipe de leur accueil chaleureux.

Je tiens à remercier aussi Amélie, Alvina, Emilie, Olivier et Romain, qui m'ont permis de jouer à de nombreux jeux pendant ma thèse. Je me souviendrai des fous rires que nous avons eu suite à nos nombreuses discussions. Par la suite j'espère pouvoir revenir continuer nos parties endiablées. "Si je peux, je peux, mais si je peux pas ... je peux pas." Je remercie aussi Céline et Maé pour les bons moments que nous avons passé : un jour nous réussirons à connaître le menu par cœur, et nous pourrons remplacer Emilie ! Un grand merci aussi à la boutique GameGoodies de Morières les Avignon pour leur sympathie et leur accueil. Je remercie aussi Cédric, Clément, Esteban, François, Joris, Kevin, Maxime, les Nicolas, Luis et Tony, pour nos différentes parties de cartes. Nous nous reverrons tous j'en suis persuadé.

Je tiens à remercier toute la collocation de Montpellier qui m'a hébergé pendant mes différents séjours sur Montpellier, et qui m'a fait découvrir le jeux *Evolution*.

Je remercie tout le personnel du CIRM qui a permis à moi mais aussi aux autres personnes du CEMRACS de passer un moment inoubliable. Grâce à leurs conseils, j'ai pu découvrir de magnifiques endroits à visiter dans les calanques. Je n'oublierai pas tous les origamis que j'ai pu faire pendant ce séjour. En espérant qu'ils soient exposés et que personne ne vole l'éléphant cette fois !

Je remercie aussi tous mes amis éparpillés partout, que je n'ai pas forcément eu l'occasion de revoir très souvent pendant ces trois ans, mais qui restaient tout de même à mes côtés. Merci donc à Alexandre, Camille, Margaux, Mathieu, Pauline, Perrine, Pierre-Emmanuel, Thomas, Valentin et William. Merci aussi à mon cher MacGy. Nos fous rires me manquent beaucoup. Je remercie aussi tout ceux que j'ai rencontré pendant mon doctorat : Alain, Alvaro, Antoine, Bastien, Benjamin, Esteban, François, Gabriel, Gérard, Igor, Jason, Johanny, Jordan, Julien, Kelyan, Luca, tous ceux que j'ai déjà cités précédemment, et tous ceux que j'ai oublié de citer et auprès desquels je m'excuse par avance.

Je termine ici en remerciant toute ma famille pour le soutien que chaque membre a pu m'apporter. Sans mes parents, mon frère, mes grands-parents et ma tante (et sa famille), je n'aurai pas pu arriver où j'en suis aujourd'hui. Bien que le premier, je suis le dernier qui finit mes études. Je sais que j'ai pris mon temps, mais c'est enfin terminé.

# Table des matières

# Table des figures

# Glossaire

**asexué**
: Individu donnant naissance, sans partenaire, à des individus enfants, génétiquement identiques (à mutations près) au parent (*e.g.* fission binaire des bactéries ou des champignons, reproduction végétative des fraisiers par le biais des stolons, parthénogenèse des phasmes). 24

**capacité de Charge**
: Effectif maximal qu'une population peut atteindre dans le contexte d'un modèle de croissance logistique dans un habitat donné. 36

**épistasie**
: Ecart à l'additivité (ou la multiplicativité) des effets de plusieurs mutations (non homologues) sur un trait donné. Dans notre cas, nous ne traiterons que des cas d'épistasie pour la fitness. Pour un ensemble d'allèles, l'épistasie impliquera que la distribution de leurs effets dépend du fond génétique dans lequel ils apparaissent (dans notre cas dépendent de la fitness du parent exclusivement). 55

**équation de Malthus**
: Modèle proposé par T. Malthus en 1798, décrivant la croissance d'une population biologique, à l'aide de l'équation différentielle linéaire (1.1). 34

**équation logistique**
: Modèle proposé par P. Verhulst en 1838, décrivant la croissance d'une population biologique, en prenant en compte l'existence d'une taille maximale de population (due à une limitation des ressources ou un effet négatif de la compétition à forte densité), à l'aide de l'équation différentielle linéaire (1.4). 36

**fitness Malthusienne absolue**
: Espérance du taux de croissance exponentiel (instantané au temps $t$) d'un génotype en temps continu, notée $r(t)$. 35

**Modèle Géométrique de Fisher (FGM)**
: Paysage adaptatif (*cf.* ci-dessous) où un phenotype continu détermine la fitness, et où cette dernière décroit avec la distance du phénotype à une valeur unique, appelée optimum phenotypique, noté $\mathbf{x}^*$. 55

**mutagénèse létale**
: Processus d'extinction de la population généré par un trop fort taux de mutation avec un effet moyen délétère. 77

**mutations**
: Erreurs produites lors de la réplication de l'ADN, comme la disparition d'un gène sur un chromosome ou l'échange d'une partie de chromosome. Ici nous nous interesserons implicitement aux mutations héritables, *i.e.*, affectant la lignée germinale, et transmise au descendant : pour les individus unicellulaires, toute mutation est héritable. 47

**paramètres abiotiques**
: Paramètres de l'environnement liés au non-vivant, comme la luminosité, le sol, la température, le vent, *etc*. 21

**paramètres biotiques**
: Paramètres liés au vivant, comme la compétition entre différentes espèces, la prédation, les parasites, *etc*. 21

**paysage adaptatif**
: Relation mathématique explicite liant le phénotype d'un individu avec sa fitness malthusienne absolue. 55

**sauvetage évolutif**
: Phénomène, apparaissant par le biais de l'évolution (mutations par exemple), et permettant à une population d'éviter l'extinction, dans un nouvel environnement. 26

**stress environnemental**
: Modification des conditions abiotiques (en comparaison avec un habitat de référence, généralement une source) qui réduit la fitness Malthusienne d'un génotype focal en-dessous de zéro (transformant l'habitat en puits). 25

**système Source–Puits**
: Ensemble de populations connectées par migration, dont certaines ont des conditions propices à la croissance de la population, par elle-même (source), et d'autres non (puits) qui se maintiennent uniquement grâce à l'influx de migrants). 25

# 1 Introduction

Cette thèse est dédiée à l'étude de l'adaptation d'une population asexuée, par le biais de modèles mathématiques, pour comprendre au mieux les phénomènes transitoires d'adaptation. En effet, l'adaptation est un phénomène biologique complexe, dont on connaît quelques rouages, et qui a été dépeint par R. A. Fisher en 1934 :

" *It is almost impossible with any brevity to exemplify the notion of adaptation. Just because adaptation consists, even in the simplest cases, in a multiplicity of correspondences between one sufficiently complicated system, the organism itself, and another equally complicated, the environment in which it finds itself. It is, indeed, just this multiplicity that makes the thing recognizably adaptive.*"

R. A. Fisher (1934)

Ces recherches ont mené à la publication de 3 articles. Un premier (HAMEL ; LAVIGNE ; MARTIN ; ROQUES 2020) est publié dans le journal *Nonlinear Analysis : Real World Applications* et correspond au Chapitre 2. Un second (LAVIGNE ; MARTIN ; ANCIAUX ; PAPAIX et al. 2020) est publié dans *Evolution*, et fait l'objet du Chapitre 4. Un dernier (LAVIGNE ; ROQUES 2020), publié dans le journal *Expositiones Mathematicae*, est mis en annexe (voir Annexe B). Les Chapitres 7 et 8 sont basés sur un proceeding, fait à la suite de la session 2018 du CEMRACS. Le Chapitre 5 a été récemment soumis. Enfin les Chapitres 3 et 6 proviennent de recherches en cours, mais en voie de finalisation.

Ce chapitre permet de faire un éventail historique sur les modèles mathématiques qui ont inspiré ceux étudiés dans les chapitres suivants. Toutes les simulations numériques ont été réalisées sur Matlab®.

# Sommaire

## 1.1. Contexte biologique

Dans un certain environnement donné, chaque groupe d'êtres vivants, appelé *population*, évolue de façon différente, selon l'habitat dans lequel il vit. Chaque individu sera alors plus ou moins adapté selon différents paramètres regroupés en deux classes différentes :

- ◇ Paramètres biotiques : paramètres liés au vivant, comme la compétition entre espèces, les proies, les prédateurs, les parasites, *etc.* ;
- ◇ Paramètres abiotiques : paramètres liés au non-vivant, comme la luminosité, le sol, la température, le vent, *etc.*

Cependant l'une des grandes questions encore en suspens en biologie est de comprendre et de décrire quantitativement la réponse (démographique et évolutive) d'une population aux changements de ces paramètres. Ainsi nous allons commencer par présenter le cheminement historique, que la recherche a emprunté, pour étudier les phénomènes étudiés dans cette thèse.

### 1.1.1. Historique

L'étude des facteurs, qui déterminent la diversité des êtres vivants et sa dynamique, a elle-même énormément évolué, depuis Aristote (IVe siècle avant J.C.). Il fut le premier (connu de nos jours) à avoir développé une théorie évolutive : celle du processus de développement de l'embryon (*épigenèse*), en opposition à la théorie de la préformation, qui persistera jusqu'au XIXe siècle dans l'idée populaire (voir Chapitre 2 – *Evolution and the Crisis of Neoclassical Biology de* DEPEW*;* WEBER *1996 pour plus de détails)*. En opposition à cette dernière, P de MAUPERTUIS 1745, éclairé par les théories faites entre autre par Aristote, défend la théorie de l'hérédité : les deux parents auraient la même influence, lors de la reproduction. Elle est étaillée par J.-B. de Lamarck (XIXe siècle) dans ces différents ouvrages, dans lesquels il énonce différentes lois de transmission des caractères acquis :

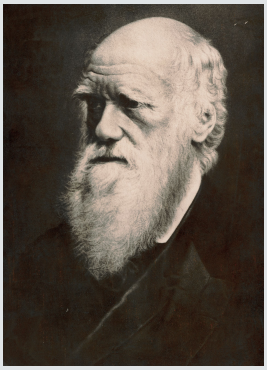
> « *Tout ce que la nature a fait acquérir ou perdre aux individus par l'influence des circonstances où leur race se trouve depuis longtemps exposée, et, par conséquent, par l'influence de l'emploi prédominant de tel organe, ou par celle d'un défaut constant d'usage de telle partie ; elle le conserve par la génération aux nouveaux individus qui en proviennent, pourvu que les changements acquis soient communs aux deux sexes, ou à ceux qui ont produit ces nouveaux individus.* »
>
> LAMARCK 1873

**Alfred Russel Wallace** (1823 – 1913) est un naturaliste, anthropologue et biologiste britannique, et l'un des premiers chercheurs à émettre l'hypothèse de sélection naturelle, en collaboration avec C. Darwin.

Dans son œuvre *Darwinism* de 1889, il expose une nouvelle théorie, appelée plus tard *effet Wallace*. Selon celle-ci, la sélection naturelle provoquerait des barrières contre l'hybridation, et séparerait ainsi des variétés différentes et leur développement : la descendance, entre deux populations souches différentes d'une même espèce, serait moins bien adaptée que ces parents et ainsi éliminée par la sélection naturelle.

Il est aussi connu pour ses théories controversées en spiritisme (par exemple, la sélection naturelle n'aurait pas lieu dans des domaines spirituels comme le génie mathématique ou artistique ou encore l'esprit et l'humour).

A. Wallace discuta dans de nombreux articles, suite à ses expériences de terrains, du fait que des espèces de même morphologie devaient vivre dans des régions du monde proches, même si celles-ci sont coupées par des barrières naturelles, comme l'Amazone. Inspiré par les travaux de T. Malthus, il correspondit avec C. Darwin, sur différents concepts évolutifs, telle que la sélection naturelle. Suite à l'essai *On the Tendency of Species to form Varieties* de A. Wallace en 1858, C. Darwin publia son ouvrage *On the Origin of Species*, décrivant en particulier le phénomène de *sélection naturelle* : comme le mentionne DEPEW ; WEBER 1996 (Chapitre 1 – *Darwin's Darwinism*), cette force extérieure, similaire à la gravité newtonienne, tendrait à démultiplier les formes les mieux adaptées, puisqu'elles se reproduisent plus que les autres et que leurs caracteristiques se transmettent à leur descendants. Ceci crée donc une adaptation des formes vivantes à leur environnement, progressivement, si celui–ci reste constant.



Le naturaliste et paléontologue anglais **Charles Darwin** (1809 – 1882) est connu pour ses travaux sur la théorie de l'évolution, et son ouvrage *On the Origin of Species* (1858).

En 1831, il part faire le tour du monde en cinq ans au bord du *HMS Beagle*, lui permettant de faire des relevés géologiques dans différents endroits, mais aussi de récupérer de nombreuses espèces et fossiles. Pendant son voyage, il écrit un journal au départ à destination de sa famille, mais qui finit publié en 1839, dans lequel sont décrits les peuples indigènes et les populations coloniales.

Bien qu'étant un scientifique reconnnu, de nombreux collègues critiquent ses pensées sur la sélection naturelle et l'évolution : il n'arrivait ni à expliquer la diversité, à partir de la sélection naturelle, ni à donner de preuves tangibles à sa théorie.

Il manquait à C. Darwin un mécanisme pour l'hérédité, que J. Mendel lui fournit à la fin du XIXe siècle : il décrivit la manière dont les gènes se transmettent de génération en génération. Contrairement à la vision de C. Darwin de changement par "petits pas" continus, il suppose que l'évolution est discontinue. Ces travaux peuvent se résumer en trois lois, appelées plus tard *lois de Mendel* par C. Correns :

- ▷ **Loi d'uniformité des hybrides de première génération** : Si les parents proviennent de souches pures, alors leurs descendants possèdent tous le même génotype.
- ▷ **Loi de disjonction des allèles** : Lors de la deuxième génération, provenant de deux parents de souches pures différentes, peuvent être observés deux groupes d'individus, séparés par des traits distincts. Cette loi implique la présence d'*allèle dominant* et d'*allèle récessif*.
- ▷ **Ségrégation indépendante des caractères héréditaires** : Si deux parents ont différents traits biologiques, ces caractères sont transmis indépendamment les uns des autres.

Alors que toutes ces études ne sont principalement basées que sur une description du monde vivant, Sir F. Galton utilisa des méthodes statistiques pour décrire mathématiquement la théorie de la sélection naturelle. Cette relation entre biologie évolutive et mathématiques est la mère de bon nombre d'autres modèles biologiques. Elle donna naissance entre autre à la dynamique des populations en écologie (menée par A. Lotka et V. Volterra) et à la génétique des populations, développée par Sir R. Fisher, J. Haldane et S. G. Wright. Ces derniers étudièrent en particulier l'effet conjoint de la sélection et des mutations sur les fréquences des gènes (*cf.* Section 1.3.3). Le chapitre 5 de PROVINE 2001 fait une référence historique de leurs recherches, qui divergent dans les hypothèses faites sur les effets de la dérive génétique.

Nous avons donc un corpus théorique avec de nombreux modèles, créés depuis plus d'un siècle, dont les conclusions dépendent fortement de modalités particulières, notamment des hypothèses faites sur l'écologie (un ou plusieurs habitats, connectés ou pas, autres espèces ou pas, *etc.*) et sur la base génétique de l'adaptation (reproduction sexuée ou asexuée, le grand ou le faible nombre de loci polymorphes, effets des mutations sur la fitness, *etc.*). Nous devons donc choisir un contexte particulier en partie, puisqu'il n'existe pas de théorie/prédiction générale et analytique à ce jour.

Dans cette thèse, seuls les organismes haploïdes asexués, c'est-à-dire dont les descendants sont identiques à leur (unique) parent, aux mutations près (comme les virus et les bactéries), seront principalement traités (sauf dans le Chapitre 7). Comprendre la réponse démographique et evolutive à des changements environnementaux, dans ces espèces, a de profondes conséquences en santé humaine, vétérinaire et en agronomie :

- ◇ **Traitement antibiotique** : Une bactérie, comme *Escherichia Coli* (Elena ; Lenski 2003 ; Harmand ; Gallet ; Martin ; Lenormand 2018 ; Lenski 2017 ; Lenski ; Rose ; Simpson ; Tadler 1991 ; Wiser ; Ribeck ; Lenski 2013) ou *Citrobacter Freundii* (Harmand ; Gallet ; Martin ; Lenormand 2018), peut être initialement sensible à un traitement antibiotique, et donc son effectif dans un environnement (corps malade, intestins, *etc.*) diminue. Cependant, elle peut finir par s'adapter génétiquement et devenir résistante.
- ◇ **Changement d'hôte pour un virus** : Un parasite, comme le Virus de la Mozaïque de l'Endive, peut infecter certaines plantes comme le salsifi sauvage. Cependant, par le biais de vecteurs naturels (pucerons par exemple), il peut être en contact avec un nouvel hôte, telle que la laitue, et la contaminer (expériences menées actuellement dans l'Unité de Pathologie Végétale de l'INRAE Avignon). D'autres exemples expérimentaux peuvent se trouver dans Dennehy ; Friedenberg ; McBride ; Holt et al. 2010.
- ◇ **Pollution des sols** : Suite à l'utilisation de certains produits chimiques, des éléments chimiques comme le Chlorure de Sodium (NaCl) ou l'Oxyde de Cuivre (CuO-NP) se retrouvent de plus en plus dans la composition des sols. Suite à cela, des chercheurs veulent utiliser des plantes, telles que la jacynthe d'eau *Eichhornia Crassipes* (Guezo ; Fiogbe ; Tobias 2017) ou les lentilles d'eau *Landoltia Punctata* (Shi ; Abid ; Kennedy ; Hristova et al. 2011) (espèces se reproduisant en partie de façon asexuée), pour qu'elles absorbent ces produits, et ainsi dépolluer la terre.
- ◇ **"Relapse"** : Les tumeurs cancéreuses, initialement bien endiguées par un traitement, peuvent reprendre une dynamique de croissance après un temps soumis à ce traitement (Pattison ; Mitchell ; Lade ; Leong et al. 2017).

Dans les deux cas, si l'individu arrive à s'adapter (soit quand la bactérie finit par être résistante, soit quand le virus arrive à contaminer un nouvel hôte), on parle de contournement de résistance.

L'évolution de telle résitance résulte de l'effet cumulé de différents processus évolutifs – sélection (Section 1.3.1) et mutations (Section 1.3.2)– et démographiques – croissance et extinction (Section 1.2.3.2). Ces phénomènes peuvent avoir lieu dans un habitat isolé. Cependant, lorsque des habitats sont connectés, la migration peut en plus affecter cette évolution. En effet, chaque environnement provoque des contraintes évolutives différentes pour une même population : changer d'environnement provoque ce qu'on appelle un stress environnemental.

### 1.1.2. Effets de la migration : Environnements sources vs environnements puits

Dans certaines situations, on ne peut se borner à ne considérer qu'un seul environnement, même pour une dynamique évolutive à court ou moyen terme. Une population souffrira de différents changements de paramètres environnementaux (ressources, prédation, *etc.*). Ainsi nous devons de nous intéresser à un environnement changeant (spatiallement et/ou temporellement). Le premier pas est de considérer deux environnements connectés par un flux migratoire, causé par l'une des raisons suivantes (LOREAU ; DAUFRESNE ; GRAVEL ; GUICHARD et al. 2013) :

- ◇ l'environnement n'est pas assez viable pour la population locale. Suite à un manque de ressources, sans migration avec un autre habitat, la population finirait par s'éteindre ;
- ◇ les environnements voisins proposent de meilleures conditions de vie à la population, l'incitant à migrer.

Ces deux habitats étant différents, ils ne causeront pas les mêmes conditions évolutives et démographiques pour une même population. Par exemple, les individus pourraient être mieux adaptés dans le premier environnement, que dans le second. Une autre possibilité est que l'habitat 1 ait plus de ressources et du coup soit occupé par un plus grand nombre d'individus, ce qui peut influer sur l'asymétrie de la migration, *etc*. Cela nous amène à la notion de système Source–Puits.

Un environnement est une source, si la population peut être maintenue sans migration : le taux de natalité surpasse le taux de mortalité – la population est bien adaptée. À l'inverse, un puits est un habitat dans lequel sans migration la population est vouée à l'extinction : le taux de mortalité est plus important que le taux de natalité, et donc la population décroît. On suppose typiquement que le flux migratoire partant d'une population donnée est proportionnel à son effectif. Une source étant typiquement plus peuplée qu'un puits, le flux migratoire entre source et puits sera *a priori* asymétrique, majoritairement de la source vers le puits.

D'après LOREAU ; DAUFRESNE ; GRAVEL ; GUICHARD et al. 2013, la notion de source-puits existe en dehors de la biologie (et lui est antécédente). Dans les an-

nées 1970, la géologie souhaitait décrire les différentes composantes de la Terre : celles d'où une substance chimique provenait (sources) et celles qui l'absorbaient ou la détruisaient (puits). Une décennie plus tard, ce concept est appliqué à l'étude d'évolution de population (HOLT 1985 ; PULLIAM 1988 ; SHMIDA ; ELLNER 1984). FURRER ; PASINELLI 2016 recense de nombreux systèmes source-puits dans de nombreux clades (*e.g.* amphibiens et mammifères) du règne animal.

La migration entre ces deux types d'environnements (source et puits) peut avoir différentes formes, ce qui mène à différentes notions de système source / puits (référencées dans LOREAU ; DAUFRESNE ; GRAVEL ; GUICHARD et al. 2013 et SOKURENKO ; GOMULKIEWICZ ; DYKHUIZEN 2006). Nous allons nous attarder ici sur certaines d'entre elles. Mais avant cela, nous devons faire la remarque qu'« être une source » et « être un puit » n'est pas un état figé dans le temps, bien que PULLIAM 1988 montre qu'une source et un puits peuvent être continûment liés, et maintenir un état d'équilibre dans lequel la source reste viable, tandis que le puits reste non viable.

**Puits fermé / Population isolée.** À un moment précis, des individus d'une source migrent vers le puits, et y restent isolés (voir Fig. 1.1) : aucune migration n'a lieu entre les deux environnements, après cet événement initial. On appelle ce système « puits fermé » ("Closed Sink", *cf.* SOKURENKO ; GOMULKIEWICZ ; DYKHUIZEN 2006). On peut donc résumer leur dynamique à celle qui a lieu dans un unique environnement, puisque leur absence de migration retour garantit l'invariance de la source (si elle était initialement dans un état d'équilibre). De nombreuses recherches ont déjà été menées sur ce sujet du point de vue expérimental (ELENA ; LENSKI 2003 ; LENSKI 2017 ; LENSKI ; ROSE ; SIMPSON ; TADLER 1991 ; TSIMRING ; LEVINE ; KESSLER 1996 ; WISER ; RIBECK ; LENSKI 2013) comme théorique (ANCIAUX ; CHEVIN ; RONCE ; MARTIN 2018 ; BELL ; GONZALEZ 2009 ; GOMULKIEWICZ ; HOLT ; BARFIELD ; NUISMER 2010).

Lorsque le changement d'environnement est trop grand, le taux de croissance de la population, initialement positif (cas d'une croissance de population), peut devenir négatif, et donc mènerait à l'extinction. Cependant, avant que cela ne se produise, des individus mieux adaptés (soit déjà présents initialement, soit étant apparus par des mutations) peuvent se multiplier et ainsi éviter l'extinction : c'est ce qu'on appelle le sauvetage évolutif (ANCIAUX ; CHEVIN ; RONCE ; MARTIN 2018 ; BELL ; GONZALEZ 2009). Ici la question clé de ce sujet est de connaître la probabilité que le puits soit envahi (probabilité de rescue), qui dépend de différents paramètres évolutifs, comme la taille de la population, les taux de mutation et de migration, *etc*.

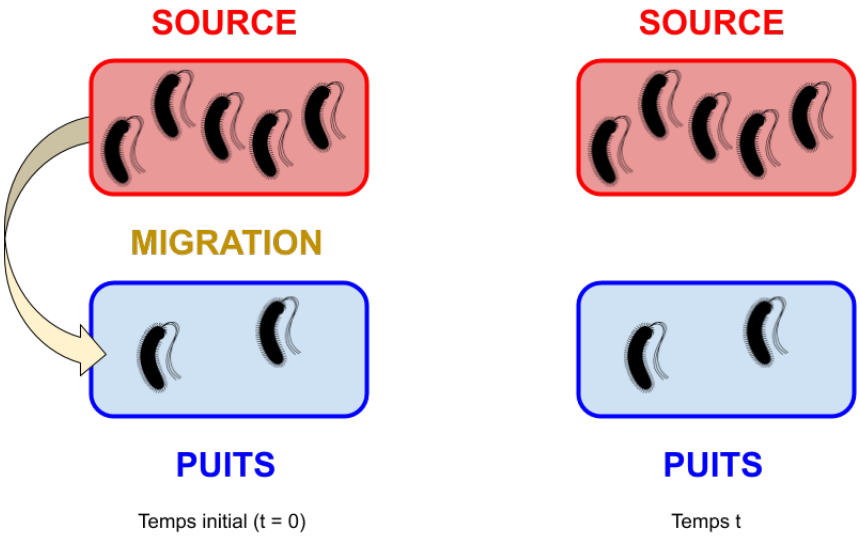



FIGURE 1.1. – **Schéma du système source / puits fermé.** Dans l'environnement rouge, la population possède un taux de croissance positif. Des individus migrent de la source au puits au temps $t = 0$. Les individus, après avoir migré, subissent la sélection du nouvel environnement, et ont un taux de croissance négatif dans l'environnement bleu.

**Puits à effet trou noir.** La source envoie de façon récurrente des individus dans le puits (voir Fig. 1.2). Cependant, celui-ci n'en renvoie pas en retour et les individus s'y éteignent (intialement du moins) : d'où l'idée de puits « trou noir ». Lorsque, par adaptation, le puits cesse d'être un « trou noir », *i.e.*, lorsque la population devient stable dans le puits, même en l'absence de migration, alors le puits a été « envahi ».

Ce genre de système inclut l'invasion de nouveaux habitats par des individus étrangers (COLAUTTI ; ALEXANDER ; DLUGOSCH ; KELLER et al. 2017), le changement d'hôte pour un virus (DENNEHY ; FRIEDENBERG ; MCBRIDE ; HOLT et al. 2010) ou encore la résistance antibiotique (JANSEN ; COORS ; STOKS ; DE MEESTER 2011 ; SOKURENKO ; GOMULKIEWICZ ; DYKHUIZEN 2006). L'asymétrie dans la migration est le premier pas de l'étude d'environnements connectés continuellement, permettant de comprendre les phénomènes démographiques et/ou évolutifs jouant un rôle pour le succès d'une invasion. Bon nombre de modèles se focalisent sur l'effet démographique en négligeant les processus d'évolution (sélection et mutations) comme DRURY ; DRAKE ; LODGE ; DWYER 2007.

Deux modèles de colonisation d'un « puits à effet trou noir » seront présentés dans ce manuscrit, incluant certains effets éco-evolutifs (migration, croissance, mutation et selection). Nous nous concentrerons sur la dynamique transitoire du système, puisque dans chaque cas l'invasion ultime du puits est certaine dans ces modèles. Le premier (*cf.* Annexe D) est un aperçu d'un modèle non épistatique, et met en lumière les conséquences de négliger l'épistasie. Le second (*cf.* Chapitre 6) sera basé sur le FGM (*cf.* Section 1.4) pour modéliser l'évolution dans chaque environnement. Dans ces deux modèles, des résultats explicites peuvent être obtenus, car il n'existe pas de boucle éco-évolutive entre le puits et la source. Celle-ci a sa dynamique propre (typiquement stable) et ne subit pas l'effet de l'évolution du puits. Cela simplifie le traitement mathématique à un problème à un environnement.

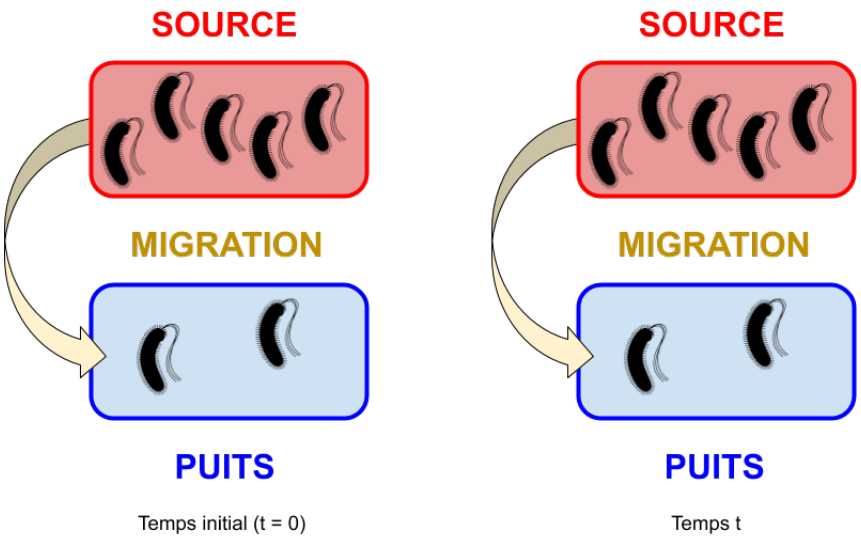



FIGURE 1.2. – **Schéma du système source / puits à effet trou noir.** Dans l'environnement rouge, la population possède un taux de croissance positif. Des individus migrent de la source au puits dès le temps $t = 0$. Les individus, après avoir migré, subissent la sélection du nouvel environnement, et ont un taux de croissance négatif dans l'environnement bleu.

**Puits réciproques.** Le puits et la source ont un échange continuel d'individus allant de l'un vers l'autre (voir Fig. 1.3). SOKURENKO ; GOMULKIEWICZ ; DYKHUIZEN

2006 parlent alors de système “Reciprocal Sink”. Différents scenarii peuvent se produire, selon les paramètres évolutifs. Considérons l'exemple particulier où deux habitats différents sont connectés par une migration symétrique. L'un des habitats peut être viable pour la population initiale et l'autre non viable. Contrairement au cas précédent du puits à effet trou noir, où la migration était purement unidirectionnelle (de la source vers le puits), autoriser la migration retour du puits vers la source peut parfois modifier grandement les dynamiques. En effet, DÉBARRE ; RONCE ; GANDON 2013 et RONCE ; KIRKPATRICK 2001 font resortir quatre comportements asymptotiques différents :

- ▷ La population s'éteint dans les deux environnements ;
- ▷ Les distributions deviennent asymétriques bimodales : on ne trouve dans chacun des habitats que des individus bien adaptés à l'un des deux environnements ;
- ▷ Les distributions sont symétriques bimodales. Dans chaque environnement, des individus sont spécialisés aux conditions abiotiques, et donc très bien adaptés à leur environnement, mais pas à l'autre ;
- ▷ Les distributions sont symétriques unimodales. La même distribution se trouve dans les deux environnements. Elle favorise des individus qui sont aussi bien adaptés dans un environnement comme dans l'autre : ce phénomène s'appelle le généralisme, tandis que les deux précédents cas correspondent à la spécialisation.

Ces différents états stationnaires ont été étudiés par differentes methodes. Ont d'abord été considérés des distributions normales des phénotypes dans chaque habitat (RONCE ; KIRKPATRICK 2001), puis des sommes pondérées de distributions normales des phénotypes, menant à une distribution bimodale (DÉBARRE ; RONCE ; GANDON 2013). Toutefois, l'hypothèse de normalité est souvent peu précise en présence de migration, encore plus pour des asexués : la migration entre les deux habitats entraîne un excès de phénotypes d'un côté de la médiane et non de l'autre, créant une asymétrie de la distribution, appelée *skewness*. Des approches récentes pour contourner cette difficulté ont consisté à étudier de faibles déviations à cette distribution normale, soit *via* des approches de dynamique adaptative, soit *via* des équations aux dérivées partielles de type Hamilton-Jacobi (GANDON ; MIRRAHIMI 2017a ; MIRRAHIMI ; GANDON 2019). Cela a mené à l'étude quantitative de la spécialisation et du généralisme : si le changement d'environnement est trop important, comparé au taux de migration (ce qui se traduit en l'inégalité $2s\theta^2 < m$ par exemple dans MIRRAHIMI ; GANDON 2019), la population sera bimodale et la spécialisation a lieu. Dans le cas contraire, la population évolue vers le généralisme.

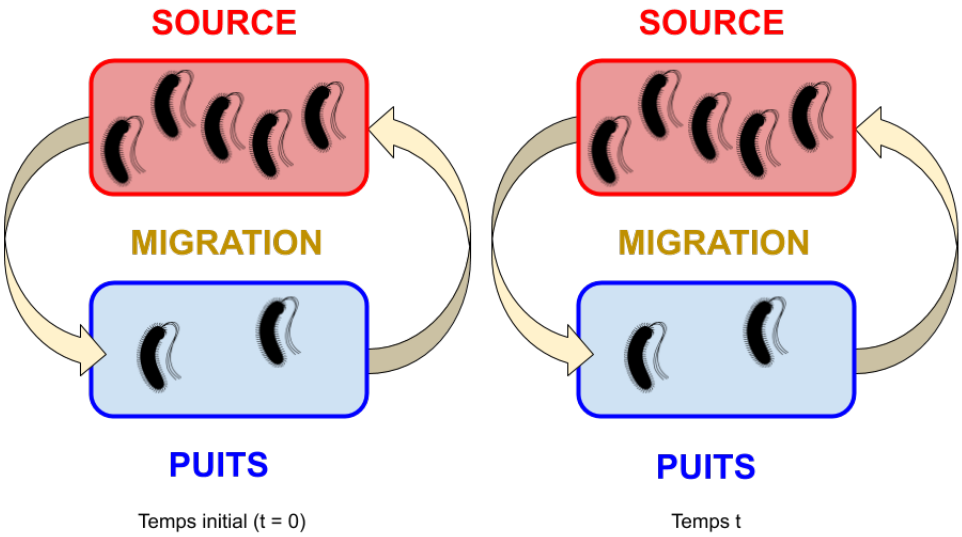



FIGURE 1.3. – **Schéma du système source / puits réciproque.** Des individus migrent entre les deux habitats, dans les deux sens, à tout temps, et subissent la sélection.

Si les taux de migrations sont différents, alors d'autres phénomènes peuvent avoir lieu, comme le souligne KAWECKI ; HOLT 2002. En effet, si l'asymétrie est trop forte (trop forte migration vers le puits, et faible migration vers la source), le système source-puits peut être renversé : la source devient un puits, et vice et versa.

D'autres modèles souhaitent plutôt comprendre l'effet de la migration sur la prédation, dans la continuité des travaux de A. Lotka et V. Volterra. Le modèle de SEGEL ; LEVIN 1976 implique qu'un état d'équilibre est instable, si les prédateurs se dispersent beaucoup plus que les proies.

Notons enfin que presque tous les modèles utilisés ici et dans les références citées sont deterministes et n'ont donc aucune composante de stochasticité, distribution des temps d'invasion, *etc*. Pour pouvoir détailler le processus transitoire dans un contexte simple, nous nous focaliserons sur le cas simple à deux habitats discrets ou moins, sauf mention contraire.

## 1.2. Modèle démographique – processus de reproduction

Pour pouvoir prédire l'évolution et donc l'adaptation d'une population, il nous faut faire un lien avec des modèles mathématiques. Ceux-ci se basent sur une observation, qui mène vers une question, à laquelle une réponse va être cherchée par le biais des mathématiques (probabilités, géométrie, algèbre, équations différentielles, *etc.*). Ainsi selon le type d'outils et d'hypothèses mathématiques faites, plusieurs modèles peuvent être créés pour répondre à une même question, chacun ayant ses qualités mais aussi ses défauts. Plus le nombre de phénomènes considérés par le modèle est important, plus celui-ci sera complet, mais plus celui-ci sera compliqué à résoudre et à étudier. Ainsi la modélisation doit simplifier le problème, tout en préservant l'essence même de celui-ci, selon LEVINS 1966. R. D. Levins précise qu'il est impensable de travailler avec des modèles pouvant maximiser la généralisation (application à plusieurs systèmes réels dans le monde), le réalisme (prise en compte d'un nombre conséquent de variables ayant un effet sur le système étudié) et la précision (précision des prédictions), tout en voulant comprendre et prédire des phénomènes précis. Il faut donc trouver un compromis entre ces trois points comme il le mentionne :

"*The world is stranger than we can imagine and surprises are inevitable in science. Thus we found, for example, that pesticides increase pests, antibiotics can create pathogens, agricultural development creates hunger, and flood control leads to flooding. But some of these surprises could have been avoided if the problems had been posed big enough to accommodate solutions in the context of the whole.*"

AWERBUCH ; KISZEWSKI ; LEVINS 2002

Il existe donc différentes stratégies énoncées par LEVINS 1966 (résumées sur la Fig. 1.4) :

- ◇ Sacrifier la généralisation en faveur du réalisme et de la précision, comme en gestion halieutique (HOLLING 1959 ; WATT 1956). Le nombre de paramètres peut y être réduit pour ne garder que ceux qui agissent significativement sur le problème étudié. Ces modèles, pouvant être résolu par des calculs numériques, permettent des mesures assez précises.
- ◇ Sacrifier le réalisme en faveur de la précision et de la généralité. Souvent critiqués pour des hypothèses trop simplistes, les travaux s'incluant dans ce mode de pensée souhaitent développer des théories générales. Cependant, LEVINS 1966 souligne le défaut de croire à tort qu'en négligeant des petits facteurs réels la conclusion ne sera que légèrement modifiée.

◇ Sacrifier la précision en faveur de la généralité et du réalisme. Les chercheurs de ce mouvement, comme R. H. MacArthur et R. D. Levins lui-même, étudient des modèles en faisant le moins d'hypothèses possibles. Par exemple, au lieu de donner une fonction particulière, ils n'utilisent qu'une hypothèse de monotonie, de convexité, *etc.*

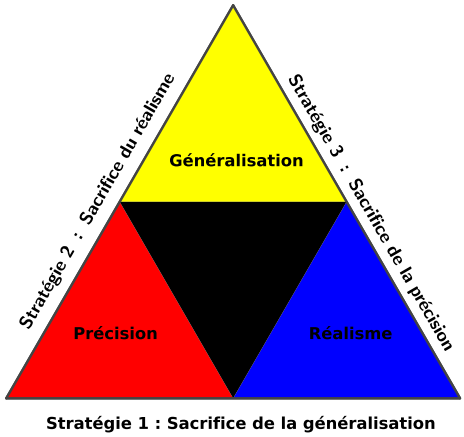
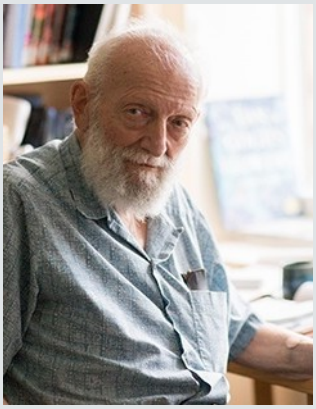



FIGURE 1.4. – **Triangle de R. Levins sur les approches de modélisation.**

Ainsi, pour résoudre un problème général, un premier modèle va être créé de façon simpliste, puis va être complété pour rajouter de la complexité mathématique, mais aussi de la précision et du réalisme. Les modèles en dynamique des populations en sont un très bon exemple, comme nous allons le voir par la suite.



**Richard "Dick" Levins** (1930 – 2016) est un généticien, biomathématicien, écologue et philosphe des sciences, d'origine américaine. Il est connu pour ses travaux sur les *métapopulations* (*id est* population de populations). Il est l'un des initiateurs des études sur l'effet d'un changement environnemental, avec par exemple la découverte surprenante que des espèces peuvent se sélectionner elles-mêmes et être menées à l'extinction, suite au changement environnemental.

Il travaille aussi sur la modélisation et publie son ouvrage *The Strategy of Model Building in Population Biology* en 1966, décrivant les différentes méthodes de modélisation.

### 1.2.1. Modèle de Malthus (1798)

Suite à son ouvrage *Essay on the principle of population* (publié pour la première fois en 1798), le pasteur T. Malthus fit trembler la population anglosaxone, en prédisant que la croissance économique de ce pays n'aurait pu alimenter la population, au bout d'un certain temps. Pour cela, il fit les hypothèses suivantes :

- ▷ La population augmente géométriquement, et la taille de la population double tous les 25 ans ;
- ▷ La production agricole, quant à elle, ne pouvait se développer que linéairement. En effet, sur 25 ans, si la population double la quantité de production, T. Malthus remarqua l'impossibilité (à son époque du moins) de quadrupler cette quantité sur les 25 années suivantes.

Ces hypothèses ont été démenties très rapidement. En effet, BRITANNICA et al. 1993 annonce qu'« Il a été démontré de manière définitive que la célèbre proposition, "la population augmente de manière géométrique, et la nourriture de manière arithmétique" est fausse ». Même de nos jours, certains chercheurs étudient ce modèle. BEAUJEU-GARNIER 1989 montre en effet que le monde ne semble pas se diriger vers un scenario de pénurie alimentaire, mais par contre dénonce l'écart grandissant du développement entre les différentes classes sociales, chose que T. Malthus n'avait pas pris en compte.

**Thomas Robert Malthus** (1766 – 1834) est un économiste britannique. Suite à la publication de *Enquiry concerning Political Justice, and its Influence on General Virtue and Happiness* (1793, W. Godwin), qui prône la prospérité et la justice d'une population en constante croissance, il publie un pamphlet (*Essay on the principle of population*), dans lequel il prédit une famine au Royaume-Uni. Il base son raisonnement sur des hypothèses de croissance exponentielle de l'effectif de la population et linéaire des ressources alimentaires. Suite à ses conclusions, il conseille de limiter les naissances pour ainsi éviter une catastrophe qui lui semble inévitable.

Suite aux scandales produits par son œuvre, il passe sa vie à développer des raisonnements scientifiques pour défendre sa thèse. Il en vient à réfuter la loi des débouchés, et à annoncer que l'offre ne créée pas la demande.

#### Application biologique

Bien que ce modèle ne soit pas suffisant pour décrire l'évolution économique d'une nation, il reste fort utile pour prédire la dynamique démographique d'une population (LENSKI ; ROSE ; SIMPSON ; TADLER 1991 ; MESZÉNA ; GYLLENBERG ;

JACOBS ; METZ 2005 ou GOMULKIEWICZ ; HOLT 1995 en version discrète). Considérons pour cela une population isolée, dont les générations ne se chevauchent pas. Notons $N(t)$ l'effectif des individus, $b(t)$ le taux de reproduction (inverse du temps d'attente pour qu'un individu ait son premier enfant) et $d(t)$ le taux de mortalité au temps $t$. Ainsi notre population peut être représentée par un processus stochastique comme nous le verrons plus tard. En prenant une limite déterministe de ce processus, l'espérance de l'effectif au temps $(t + \mathrm{d}t)$ vaut :

$$N(t+\mathrm{d}t) = N(t) + \underbrace{b(t)\,\mathrm{d}t\,N(t)}_{\text{Nb. d'enfants}} - \underbrace{d(t)\,\mathrm{d}t\,N(t)}_{\text{Nb. de décès}},$$

pour tout temps $t$ et pour toute variation temporelle $\mathrm{d}t \geq 0$. Cette dernière équation peut se réécrire sous la forme :

$$\frac{N(t+\mathrm{d}t) - N(t)}{\mathrm{d}t} = [b(t) - d(t)]\ N(t),$$

tant que $\mathrm{d}t > 0$. Finalement, en faisant tendre $\mathrm{d}t$ vers 0, nous obtenons une équation différentielle ordinaire (EDO) linéaire du premier ordre, appelée équation de Malthus :

$$N'(t) = r(t)\,N(t), \qquad \text{avec } r(t) = b(t) - d(t), \tag{1.1}$$

pour tout temps $t \geq 0$. Si nous connaissons la taille $N_0$ de la population au temps $t = 0$, cette équation admet une unique solution donnée par :

$$\forall t \geq 0,\ N(t) = N_0\ \exp\left[\int_0^t r(s)\,\mathrm{d}s\right]. \tag{1.2}$$

Bien que cette formule ne donne aucune information sur l'évolution de la population, pour une fonction $r$ quelconque, nous pouvons prédire l'avenir d'une telle population, si nous connaissons explicitement $r(t)$ en tout temps $t \geq 0$.

Revenons aux hypothèses faite par T. Malthus, se traduisant ici par des taux $b(t)$ et $d(t)$ constants. Dans ce cas, $r(t)$ est égale à une constante $r$, pour tout $t \geq 0$, et l'équation (1.2) se simplifie en :

$$\forall t \geq 0,\ N(t) = N_0\,e^{r\,t}. \tag{1.3}$$

Ainsi trois cas de figure (illustrés sur la Fig. 1.5) se présentent devant nous :

- ▷ Si $r < 0$, dans ce cas, $N(t)$ tend vers 0, quand $t$ tend vers $+\infty$. La population finira par s'éteindre ;
- ▷ Si $r = 0$, il y a autant de morts que de naissances, et donc l'effectif de la population ne varie pas au cours du temps ;
- ▷ Si $r > 0$, l'effectif de la population explose en temps grand, *id est*, $N(t)$ tend vers $+\infty$, quand $t$ tend vers $+\infty$.

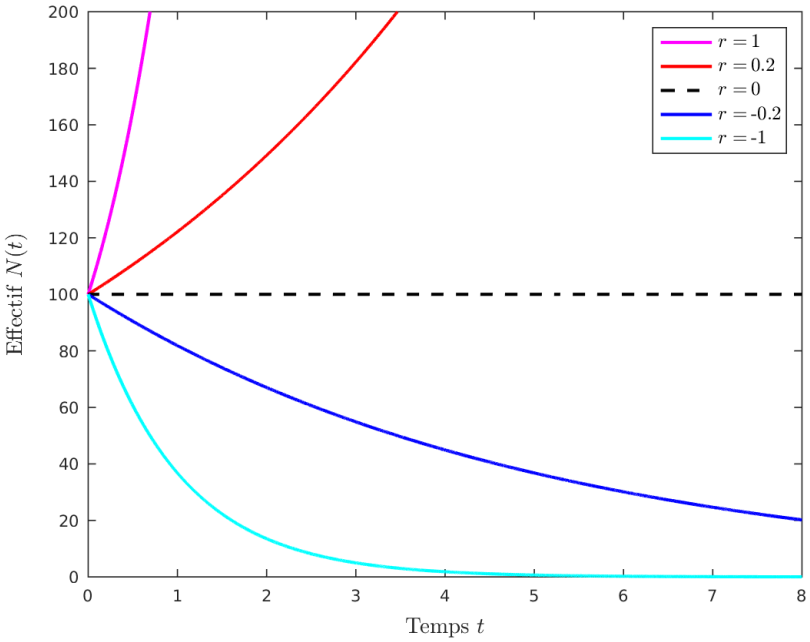



FIGURE 1.5. – **Évolution malthusienne de l'effectif $N(t)$ de population pour différentes valeurs de $r$.** Initialement, chaque population possède 100 individus. Les différentes trajectoires représentent l'évolution de population, donnée par (1.3) ayant une fitness malthusienne absolue $r$ valant 1 (courbe magenta), 0.2 (courbe rouge), 0 (courbe noire en pointillé), - 0.2 (courbe bleue) et - 1 (courbe cyan).

Cette quantité $r$ (et même plus généralement $r(t)$) décrit ainsi la croissance de la population au cours du temps : on appelle la fonction $r(t)$ *taux de croissance de la population* ou encore fitness Malthusienne absolue.

### Limites du modèle

Bien que ce modèle explique les phénomènes de colonisation et de repeuplement, qui donnent suite à des croissances exponentielles, il ne prend pas en considération les limites auxquelles les individus font face : ressources alimentaires, limitation de l'espace pour les gîtes de chaque individu, compétition entre les individus pour l'accès à des ressources, prédation, reproduction sexuée, *etc.*

T. Malthus, certes conscient que ces périodes de croissance exponentielle ne pouvaient avoir lieu que sur une durée limitée, n'a donné aucun modèle mathématique pour prendre en compte des limitations naturelles. Ce ne sera que bien plus tard que d'autres scientifiques proposeront d'autres modèles démographiques. En 1838, P. Verhulst développe un modèle évitant cette l'explosion de la taille de

population en temps infini, par le biais de l'équation logistique :

$$N'(t) = r(t)\, N(t) \left[1 - \frac{N(t)}{K}\right], \tag{1.4}$$

avec $r(t)$ la fitness malthusienne absolue au temps $t$, et $K$ l'effectif maximal que peut avoir la population dans la niche écologique où elle vit. Dans ce modèle, nous faisons l'hypothèse que le ratio entre le taux de naissance et le taux de mort dépend linéairement de la taille de population, *i.e.*, $d(t) = b(t)N(t)/K$. On appelle $K$ la capacité de Charge de la population. Dans le cas où $r$ est constant, les solutions de cette EDO sont :

$$\forall t \geq 0,\ N(t) = \frac{N_0\, K}{N_0 + (K - N_0)\, e^{-rt}}. \tag{1.5}$$

Là encore, si $N_0 \neq K$, le signe de $r$ implique que la population va soit s'éteindre ($r < 0$), soit croître exponentiellement ($r > 0$). Cependant, dans le cas de la croissance exponentielle, cette fois-ci l'effectif $N(t)$ converge vers $K$, comme nous pouvons le voir sur la Fig. (1.5).

**Pierre-François Verhulst** (1804 – 1849) est un mathématicien belge. Il soutient sa thèse sur la résolution des équations binomiales en 1825. Soutenu par L. Quételet, il se lance dans des travaux de statistiques, appliqués à la démographie.

En 1838, les travaux de T. Malthus lui inspire un nouveau modèle, qui décrit la croissance d'une population, de façon « non exponentielle ». Dans sa publication de 1845, il lui donne le nom de modèle logistique.

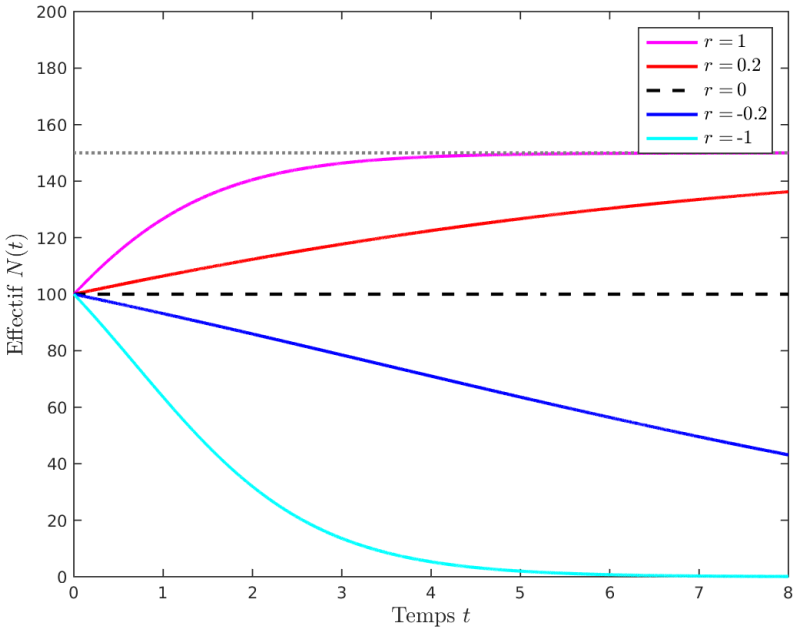



FIGURE 1.6. – **Évolution logistique de l'effectif $N(t)$ de population pour différentes valeurs de $r$.** Initialement, chaque population possède 100 individus. Les différentes trajectoires représentent l'évolution de population, donnée par (1.5) ayant une fitness malthusienne absolue $r$ valant 1 (courbe magenta), 0.2 (courbe rouge), 0 (courbe noire en pointillé), - 0.2 (courbe bleue) et - 1 (courbe cyan). La courbe grise représente l'asymptote de la taille de population $N(+\infty) = K$, pour $K = 150$.

### 1.2.2. Phénomènes migratoires

Dans cette section, nous considèrons une population d'individus qui occupent deux environnements. Dans chaque environnement, les tailles de population $N_1(t)$ et $N_2(t)$ évoluent différemment. On note $r_1$ et $r_2$ respectivement les fitness malthusiennes absolues dans l'environnement 1 et 2. Comme vu dans la Section 1.2.1, si les deux environnements ne sont pas connectés par un phénomène migratoire, les effectifs dans chacun d'eux vérifient les EDOs suivantes :

$$\forall t \geq 0, \begin{cases} N_1'(t) = r_1(t)\, N_1(t), \\ N_2'(t) = r_2(t)\, N_2(t). \end{cases} \tag{1.6}$$

Supposons maintenant que les individus migrent entre les deux environnements. Entre deux instants $t$ et $(t+dt)$, il y a $\delta_{1\to 2}\, N_1(t)\, \mathrm{d}t$ (resp. $\delta_{2\to 1}\, N_2(t)\, \mathrm{d}t$) qui partent de l'environnement 1 (resp. 2) vers l'environnement 2 (resp. 1), avec $\delta_{1\to 2} \geq 0$,

$\delta_{2\to1} \geq 0$ et $\delta_{1\to2} + \delta_{2\to1} > 0$.

Suite à la migration, une approximation déterministe peut montrer que l'environnement 1 aura un effectif au temps $(t + \mathrm{d}t)$ valant :

$$N_1(t + \mathrm{d}t) = N_1(t) + r_1(t)\ \mathrm{d}t\ N_1(t) - \underbrace{\delta_{1\to2}\ \mathrm{d}t\ N_1(t)}_{\text{Nb. d'émigrants}} + \underbrace{\delta_{2\to1}\ \mathrm{d}t\ N_2(t)}_{Nb.d'immigrants}.$$

Par les mêmes arguments que ceux avancés dans la Section 1.2.1, la fonction $N_1(t)$ vérifie l'EDO :

$$N_1'(t) = (r_1(t) - \delta_{1\to2})\ N_1(t) + \delta_{2\to1}\ N_2(t).$$

De même, on obtient que $N_2'(t) = (r_2(t) - \delta_{2\to1})\ N_2(t) + \delta_{1\to2}\ N_1(t)$. Finalement, nous avons un système différentiel ordinaire d'ordre 1, qui peut s'écrire sous forme matricielle $\mathbf{N}'(t) = A(t)\ \mathbf{N}(t)$ (type de modèle introduit par Leslie 1945 pour étudier des sous-populations classés par rapport à l'âge), avec le vecteur $\mathbf{N}(t) = (N_1(t), N_2(t))^T$ et :

$$A(t) = \begin{pmatrix} r_1(t) - \delta_{1\to2} & \delta_{2\to1} \\ \delta_{1\to2} & r_2(t) - \delta_{2\to1} \end{pmatrix}. \tag{1.7}$$

Pour la suite de cette section, nous considérons que $r_1$ et $r_2$ sont deux fonctions constantes et noterons $A$ la matrice $A(t)$ (qui devient constante). Ce modèle est la version continue du modèle discret étudié par Kawecki 1995 puis Holt 1996b, avec $r_2 < 0$ (l'environnement 2 est un puits). Selon eux, si les mutations restent faibles, les individus de la source (environnement 1) n'arrivent jamais à s'adapter au puits (environnement 2) dans différents cas :

- ▷ si $\delta_{1\to2}$ est faible, le puits reste faiblement peuplé. Ainsi les mutations n'ont pas beaucoup d'opportunité pour faire apparaître un individu suffisamment adapté, parce que même avec un $\delta_{1\to2} > 0$, on a une invasion certaine en temps long.
  - ◇ si $\delta_{2\to1}$ est faible – cas limite à effet trou-noir (*cf.* Section 1.1.2).
  - ◇ si la sélection est trop forte dans l'environnement 2 ($r_2 \ll 0$), et si les deux taux de migration ($\delta_{1\to2}$ et $\delta_{2\to1}$) sont plus petits que $1/2$.

Dans les autres cas, l'évolution permet à la population de s'adapter au nouvel environnement et ainsi l'invasion réussit.

Passons maintenant à l'étude du modèle continu $\mathbf{N}'(t) = A\mathbf{N}(t)$.

**Lemme 1.** *Supposons, quitte à permuter les indices, que $\delta_{2\to1} > 0$. La matrice $A$ donnée par (1.7) n'est pas diagonalisable dans $\mathbb{R}$ si et seulement si $\delta_{1\to2} = 0$ et $r_1 = r_2 - \delta_{2\to1}$.*

*Démonstration.* Le ploynôme caractéristique de $A$ est :

$$\chi_A = X^2 - (r_1 + r_2 - \delta_{1\to2} - \delta_{2\to1})X + (r_1 - \delta_{1\to2})(r_2 - \delta_{2\to1}) - \delta_{1\to2}\delta_{2\to1},$$

de discriminant :

$$\Delta = (r_1 + r_2 - \delta_{1\to2} - \delta_{2\to1})^2 - 4[(r_1 - \delta_{1\to2})(r_2 - \delta_{2\to1}) - \delta_{1\to2}\delta_{2\to1}].$$

On remarque que :

$$\Delta = (r_1 - r_2 - \delta_{1\to2} + \delta_{2\to1})^2 + 4\delta_{1\to2}\delta_{2\to1}.$$

Deux cas se présentent :

- ▷ Si $\Delta > 0$, $\chi_A$ admet deux racines distinctes, ce qui implique la diagonalisabilité de $A$ ;
- ▷ Supposons que $\Delta = 0$, ce qui correspond à $\delta_{1\to2}\delta_{2\to1} = 0$ et $r_1 = r_2 - \delta_{2\to1}$. Comme $\delta_{2\to1} \neq 0$, nous avons $\delta_{1\to2} = 0$. On peut ainsi réécrire $A$ comme étant :

  $$A = \begin{pmatrix} r_1 & \delta_{2\to1} \\ 0 & r_1 \end{pmatrix}.$$

  Sa seule valeur propre est $r_1$ avec $E_{r_1}(A) = \ker((1,0)^T)$. Donc $A$ n'est pas diagonalisable.

□

Par le biais du Lemme 1, nous allons pouvoir résoudre le système différentiel $\mathbf{N}'(t) = A\mathbf{N}(t)$, selon si $A$ est diagonalisable ou non, puis en déduire des critères d'explosion démographique.

**Proposition 1.** *Supposons, quitte à permuter les indices, que $\delta_{2\to1} > 0$. La population globale connaît une croissance exponentielle si et seulement si :*

$$(r_1 - \delta_{1\to2})\,(r_2 - \delta_{2\to1}) \geq \delta_{1\to2}\,\delta_{2\to1}.$$

*Démonstration.* On remarque que la condition est symétrique, tout comme le système. Il suffit donc de considérer seulement la population de l'habitat 1.

**Cas $A$ diagonalisable dans $\mathbb{R}$.** Il existe $\lambda_1$ et $\lambda_2$ (valeurs propres de $A$) telles que $A = PDP^{-1}$, avec :

$$P = \begin{pmatrix} \delta_{2\to1} & \delta_{2\to1} \\ \lambda_1 - r_1 + \delta_{1\to2} & \lambda_2 - r_1 + \delta_{1\to2} \end{pmatrix} \text{ et } D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Plus précisément, nous avons :

$$\lambda_1 = \frac{(r_1 + r_2 - \delta_{1\to 2} - \delta_{2\to 1}) - \sqrt{\Delta}}{2}, \text{ et } \lambda_2 = \frac{(r_1 + r_2 - \delta_{1\to 2} - \delta_{2\to 1}) + \sqrt{\Delta}}{2}.$$

En posant $\mathbf{V}(t) = P^{-1}\mathbf{N}(t)$, on a $\mathbf{V}'(t) = D\mathbf{V}(t)$. On sait alors que :

$$\mathbf{V}(t) = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \mathbf{V}(0) = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} P^{-1} N(0).$$

Finalement, on a :

$$\mathbf{N}(t) = P \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} P^{-1}\mathbf{N}(0),$$

ce qui implique que :

$$N_1(t) = \frac{N_1(0)}{\lambda_1 - \lambda_2}\left[(\lambda_1 - r_1 + E_1)e^{\lambda_2 t} - (\lambda_2 - r_1 + E_1)e^{\lambda_1 t}\right] + \frac{E_2 N_2(0)}{\lambda_1 - \lambda_2}\left[e^{\lambda_1 t} - e^{\lambda_2 t}\right].$$

Finalement soit l'habitat 1 n'aura plus d'habitant, soit il connaîtra une explosion démographique, selon les paramètres.

**Cas $A$ non diagonalisable.** Dans ce cas, on remarque que $N_2'(t) = r_2 N_2(t)$ admet pour solution $N_2(t) = N_2(0)e^{r_2 t}$. De plus, $N_1$ vérifie :

$$N_1'(t) = r_1 N_1(t) + N_2(0)e^{r_2 t},$$

ce qui implique que :

$$N_1(t) = N_1(0)e^{r_1 t} + N_2(0)\,\frac{e^{r_2 t} - e^{r_1 t}}{r_2 - r_1}.$$

Comme $r_2 = r_1 + \delta_{2\to 1} > r_1$, $N_1(t)$ est équivalent quand $t$ tend vers $+\infty$ à :

$$N_1(t) \underset{t\to 0}{\sim} \frac{N_2(0)}{\delta_{2\to 1}}\, e^{r_2 t}.$$

□

Dans le cas où les fitness malthusiennes absolues $r_1$ et $r_2$ dépendent de $N_1$ et $N_2$, le modèle est dit *à contexte-dépendence* (par exemple $r_1(N_1) = r_1(1 - N_1/K_1)$), l'étude devient plus complexe. L'étude ne semble être possible qu'en un état d'équilibre (voir HASTINGS 1983 ; HOLT 1985 ; HOLT 1996a pour plus de détails sur les équilibres). De plus, le coefficient de migration peut aussi dépendre des

effectifs sous la forme $\delta_{1\to 2} = f_1(N_1)$, mais là encore les études ne s'arrêtent qu'à celle de l'équilibre (GOMULKIEWICZ ; HOLT ; BARFIELD 1999 ; KAWECKI ; HOLT 2002).

### 1.2.3. Modèle de Wright-Fisher sans sélection ni mutation et équation de Feller

Nous venons de voir un modèle permettant de prédire les modifications au cours du temps de l'effectif d'une population de façon purement déterministe : si deux populations ont les mêmes taux de reproduction et de mortalité, le modèle de Malthus prédit que l'évolution de ces deux populations serait identique. Cependant, ce modèle purement déterministe n'est en fait qu'une description de l'espérance de la dynamique, sur des réplicats aléatoires. La dynamique individuelle de populations particulières présente une variation autour de cette espérance, qui peut être cruciale. L'ensemble des effets démographiques et génétiques aléatoires est regroupé dans le terme de *dérive génétique* qui décrit l'écart à l'espérance de la dynamique des fréquences génotypes.

#### 1.2.3.1. Modèle de Wright-Fisher sans sélection ni mutation

On considère ici une population d'effectif constant $N$, ne possédant que deux allèles d'un même gène, qu'on notera $a$ et $A$. Notons que pour une population diploïde, le modèle qui suit pourra s'appliquer. En effet, chaque individu possède l'un des génotypes suivants : $(a, a)$, $(a, A)$ ou $(A, A)$. Lorsque les deux allèles $A$ et $a$ à un locus donné sont codominants, et lorsque les individus sont sexués et se reproduisent en panmixie, on pourra réduire le problème à la dynamique des fréquences alléliques, en ignorant les fréquences génotypiques (qui restent à l'équilibre de Hardy-Weinberg) : nous sommes alors ramenés à étudier une population à $2N$ individus caractérisés par un allèle $a$ ou $A$.

Passons à la description du processus de reproduction. Lorsque le nombre de gamètes est bien supérieur au nombre d'adultes, on peut simplifier le processus en considérant que les allèles $A$ et $a$ dans les descendants sont obtenus par tirage avec remise dans l'ensemble des allèles parentaux. Il en résulte une distribution binomiale des fréquences alléliques, après reproduction, conditionnée par la fréquence dans la génération précédente. Ce processus correspond alors à la chaîne de markov suivante : l'état de la population de la génération $(n + 1)$ ne dépend que de la génération $n$. On note $Z^{(n)}$ l'effectif du gène $a$ dans la population à la génération $n$ et les proportions $Y^{(n)} = Z^{(n)}/N$. Ainsi la *probabilité de transition* de passer de l'événement $\{Z^{(n)} = i\}$ à l'événement $\{Z^{(n+1)} = j\}$ vaut :

$$p_{ij} = \binom{N}{j} p_i^j q_i^{N-j},$$

avec $p_i = i/N$ (resp. $q_i = 1 - p_i$) la probabilité qu'un parent donne naissance à un

individu de gène $a$ (resp. $A$) et $i$ le nombre de gène $a$ dans la population. Comme $(Y^{(n)})_{n\in\mathbb{N}}$ est une chaîne de Markov, on sait que $\mathbb{E}[Y^{(n)} - Y^{(n-1)}] = 0$. D'après les résultats classiques sur les lois binomiales, on sait que :

$$\text{Var}\left[Y^{(n)} - Y^{(n-1)} \,|\, Z^{(n-1)} = i\right] = \frac{1}{N^2}\, N\, p_i\, q_i = \frac{p_i\, q_i}{N}.$$

#### 1.2.3.2. Équation de Feller

KIMURA 1954 proposa une approximation de ce processus par des équations de diffusion, permettant ainsi d'obtenir des résultats sous forme plus explicite, en étendant ces précédents résultats à des processus markoviens à valeurs réelles ($Y^{(n)} \in \mathbb{R}$ au lieu de $\mathbb{N}$) et indexé sur le temps $t \in \mathbb{R}_+$, et nous avons que :

$$\text{Var}\left[\Delta Y(t) \,|\, Z(t_0) = x\right] = \frac{x(1-x)}{N},$$

pour tout $t_0 \geq 0$. En utilisant le théorème de Kolmogorov, il approxime le modèle de Wright-Fisher, par un modèle différentiel, lorsque les changements sont faibles et que l'effectif $N$ n'est pas trop petit. La densité de probabilité $u(t,x)$ de $(Y(t))_{t\geq 0}$ est alors solution de l'équation de Fokker-Planck :

$$\partial_t u(t,x) = \frac{1}{2N}\,\partial_{xx}\left[x(1-x)u(t,x)\right]. \tag{1.8}$$

Par le changement de variable $(t,x) \to (Nt,x)$, nous pouvons nous ramener à l'EDP $\partial_t u(t,x) = \frac{1}{2}\,\partial_{xx}\left[x(1-x)u(t,x)\right]$. Cette équation, avec la condition initiale $u(0,x) = \delta_p(x)$ (pour $p \in [0,1]$), admet une unique solution dans l'ensemble $H$ défini ci-dessous (voir TRAN ; HOFRICHTER ; JOST 2013) :

$$H = \left\{ f : [0,1] \to [0,+\infty] \text{ mesurable avec } \int_{[]0,1]} f(x)\, g(x)\, \mathrm{d}x < \infty,\ \forall g \in C^\infty([0,1]) \right\}.$$

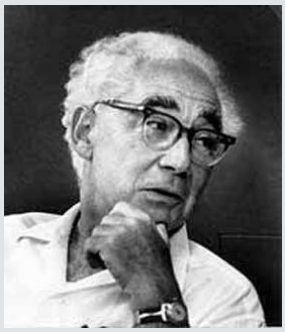
Cette solution admet même une formulation explicite en série, faisant intervenir les polynômes de Gegenbauer. Des propriétés plus classiques, comme le principe du maximum, l'unicité ou la solution fondamentale, ont été démontrées par EPSTEIN ; MAZZEO 2010, dans le cas où le support de $x$ n'était pas borné et plus précisément dans le cas où $x > 0$.

On suppose maintenant que l'équation (1.8) reste valide si l'effectif $N$ dépend du temps. Cette dernière relation peut se traduire en une Équation Différentielle Stochastique (EDS), appelée *Équation de Feller*, développée par FELLER 1951 :

$$\forall t \geq 0,\ \mathrm{d}N(t) = r(t)\, N(t)\, \mathrm{d}t + \sqrt{\sigma^2\, N(t)}\, \mathrm{d}W(t), \tag{1.9}$$

avec $r(t)$ la fitness malthusienne absolue, $\sigma^2$ la variance de la reproduction stochastique et $(\mathrm{d}W(t))_{t\geq 0}$ un *bruit blanc* (voir Annexe A pour plus de détails).

Dans le cas où $r$ est constant, cette EDS provient de la limite diffusive d'un processus de naissances et morts (BAAKE ; WAKOLBINGER 2015 ; FELLER 1951 ; LAMBERT 2008). Elle décrit l'évolution de la taille de population $N(t)$ avec un taux de naissance $b = k + \frac{b_1}{N}$ et un taux de mortalité $d = k + \frac{d_1}{N}$, pour $b_1,\ d_1 \geq 0$ et $k > 0$ : dans ce cas, $r = b - d$ et $\sigma^2 = b + d$.



**William Feller** (1906 – 1970) obtient son doctorat en 1926, sous la direction de R. Courant. En 1933, fuyant le régime Nazi, il quitte l'Allemagne pour le Danemark, puis la Suède, et finit par s'installer aux États-Unis, où il se fait naturalisé en 1944.

Ses nombreux articles portent sur la théorie de la mesure, l'analyse fonctionnelle, la géométrie et les équations différentielles. Il joue un rôle majeur dans l'avancée en probabilités (marches aléatoires, processus de diffusion, *etc.*) et développe des relations entre chaînes de Markov et EDP.

Comme nous pouvons le remarquer sur la Fig. 1.7, cette EDS ajoute un bruit stochastique autour de la trajectoire déterministe donnée par (1.2). Cependant, en moyenne, les simulations correspondent à la formule donnée par le modèle de Malthus.

L'un des intérêts de cette équation est de pouvoir exprimer le temps d'extinction de la population, plus facilement. Pour cela, nous définissons le premier temps de sortie de $\mathbb{R}_+^*$ par :

$$\tau = \inf\{t \geq 0,\ \exists s \leq t, N(s) = 0\}.$$

Quand $r$ est constant, la distribution des temps d'extinction $\tau$ est connue (voir par exemple proposition 4.4 de la section 4.3 de DAWSON 2017). Sa fonction de répartition est donnée par :

$$\forall\, t > 0,\ \forall\, N_0 \geq 0,\ \mathbb{P}(\tau \geq t | N_0) = 1 - \exp\left[-\frac{2rN_0\sigma^{-2}}{1 - e^{-rt}}\right].$$

Des travaux récents (BANSAYE ; SIMATOS 2015), basés sur des résultats de probabilités sur la convergence de processus (de branchement) de Galton-Watson, ont étudié le cas inhomogène ($r$ dépendant du temps $t$) :

$$\forall\, t > 0,\ \forall\, N_0 \geq 0,\ \mathbb{P}(\tau \geq t | N_0) = 1 - \exp\left[-\frac{2\, N_0\sigma^{-2}}{\int_0^t \exp\left(-\int_0^\xi r(\gamma)d\gamma\right) d\xi}\right]. \quad (1.10)$$

Cependant, une autre preuve, basée sur des méthodes propres aux équations aux dérivées partielles (EDP), a été proposée par LAVIGNE ; ROQUES 2020, pour calculer $\mathbb{P}(\tau \geq t | N_0)$, pour tout $\tau \geq 0$ (voir Annexe B pour plus de détails).

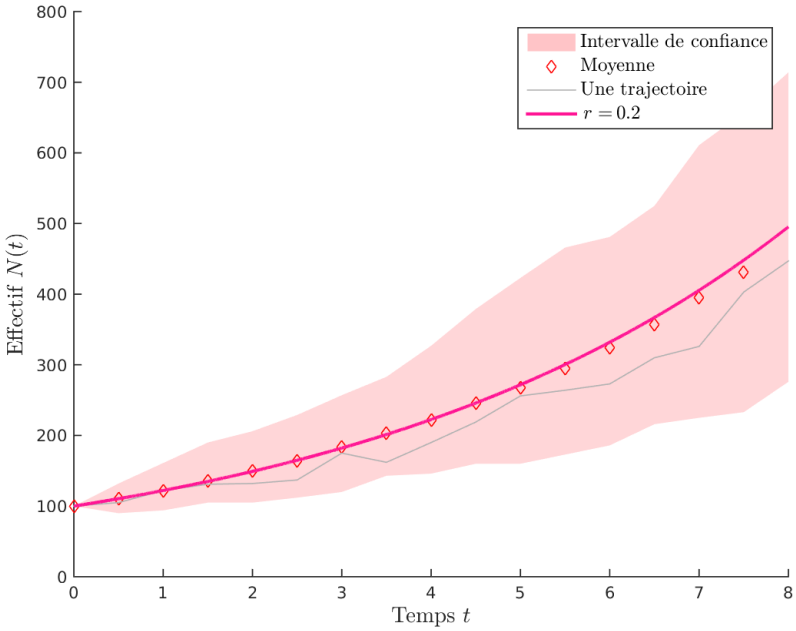



FIGURE 1.7. – **Évolution malthusienne de l'effectif $N(t)$ de population avec stochasticité démographique.** Initialement, la population possède 100 individus et a une fitness malthusienne absolue $r = 0.2$. La courbe magenta représente l'évolution déterministe donnée par (1.3). La zone colorée représente l'intervalle entre les quantiles 0.025 et 0.975 sur des simulations de la solutions de l'EDS (1.9) et les points rouges en forme de diamant correspondent à la moyenne sur ces 100 simulations. La courbe grise correspond à l'évolution d'une des simulations stochastiques. Ici la variance stochastique vaut $\sigma^2 = 1$.

## 1.3. Modèle phénotypique

Le modèle démographique présenté dans la Section 1.2 fait l'hypothèse que chaque individu possède les mêmes taux de reproduction et de mortalité. Cependant, ces derniers dépendent bien évidemment de chaque individu et plus précisément de leur phénotype hérité. Cela implique donc de considérer non pas la population entière de l'habitat, mais des sous-populations de même part héritable (*i.e., breeding value*) définie par $y = \mathbb{E}[\text{phénotype} \,|\, \text{génotype}]$.

Avant de passer aux différents modèles qui nous intéressent, nous devons expliquer comment représenter un phénotype, c'est-à-dire un ensemble de traits biologiques (par exemple, la taille, le poids, *etc.*). Chacun de ces traits biologiques peut être mesuré, et donc représenté par un réel $x \in \mathbb{R}$. Étant donné qu'un phénotype est l'ensemble de $n$ traits biologiques, chaque phénotype et donc

chaque individu sera représenté par un vecteur $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

### 1.3.1. Modélisation de la sélection

Considérons le groupe d'individus d'une population possèdant un phénotype $\mathbf{x} \in \mathbb{R}^n$ donné. On note alors le taux de croissance associé $\varrho(\mathbf{x})$. Définissons maintenant les différentes fitness malthusiennes absolues $r$. Bien que deux individus n'aient pas forcément le même phénotype, ils peuvent partager la même fitness malthusienne absolue. En effet, celle-ci ne dépend pas directement du phénotype mais de la part héritable $\mathbf{y}$ définie comme étant le phénotype moyen, sachant le génotype. Dans ce cas, la fitness malthusienne absolue $r(\mathbf{x})$ sera définie comme la valeur moyenne du taux de croissance $\varrho$ sachant le génotype. Dans la suite de ce manuscrit, nous confondrons, pour simplifier les raisonnements, les notions de *breeding value* avec les phénotypes, et le taux de croissance avec la fitness malthusienne absolue.

Nous avons alors vu dans la Section 1.2 que dans la limite deterministe :

$$\partial_t N(t, \mathbf{x}) = r(\mathbf{x})\, N(t, \mathbf{x}), \tag{1.11}$$

où $N(t, \mathbf{x})$ est la densité des individus au temps $t$ ayant le phénotype $\mathbf{x}$ et $\partial_t$ est la dérivée partielle par rapport à la variable $t$. Posons $N(t)$ la taille de la population globale au temps $t$ :

$$N(t) = \int_{\mathbb{R}^n} N(t, \mathbf{x})\, \mathrm{d}\mathbf{x}.$$

En intégrant formellement (1.11) par rapport à $\mathbf{x} \in \mathbb{R}^n$, nous obtenons :

$$N'(t) = \overline{r}(t)\, N(t), \qquad \text{avec} \qquad \overline{r}(t) = \frac{1}{N(t)} \int_{\mathbb{R}^n} r(\mathbf{x})\, N(t, \mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{1.12}$$

Ainsi l'évolution de la population dépend non plus d'une fitness commune à chaque individu, mais de la *fitness malthusienne absolue moyenne* $\overline{r}(t)$.

Cette quantité $\overline{r}(t)$ traduit l'effet de la sélection naturelle sur les individus. Pour voir cela, intéressons-nous plutôt à la distribution $q(t, \mathbf{x}) = N(t, \mathbf{x})/N(t)$ d'individus ayant le phénotype $\mathbf{x}$ au temps $t$, qui vérifie :

$$\partial_t q(t, \mathbf{x}) = \frac{\partial_t N(t, \mathbf{x})}{N(t)} - \partial_t N(t)\, \frac{N(t, \mathbf{x})}{N(t)^2} = [r(\mathbf{x}) - \overline{r}(t)]\; q(t, \mathbf{x}). \tag{1.13}$$

L'équation intégro–différentielle (1.13) implique deux comportements d'évolution du groupe des individus de phénotype $\mathbf{x} \in \mathbb{R}^n$ :

- ▷ Si la fitness $r(\mathbf{x})$ est inférieure à $\overline{r}$, dans ce cas, $q(t, \mathbf{x})$ décroît : ces phénotypes moins adaptés que la moyenne de la population tendent à disparaitre ;
- ▷ Si la fitness $r(\mathbf{x})$ est supérieure à $\overline{r}$, la proportion $q(t, \mathbf{x})$ va croître au cours du temps : les phénotypes sont bien adaptés. De plus si $r$ est majorée et si

$r(\mathbf{x}) = \max(r)$, alors ces phénotypes optimaux deviendront majoritaires.

Ainsi la sélection tend à homogénéïser la population en ne gardant que certains phénotypes, comme nous pouvons le voir sur la Fig. 1.8, représentant la discrétisation de la solution de l'EDP (1.13), donnée par un schéma numérique explicite. Pour cet exemple, la fitness malthusienne absolue est donnée par $r(x) = 1 - x^2/2$ (avec $n = 1$). Cette forme quadratique se révèlera être importante dès la Section 1.4.

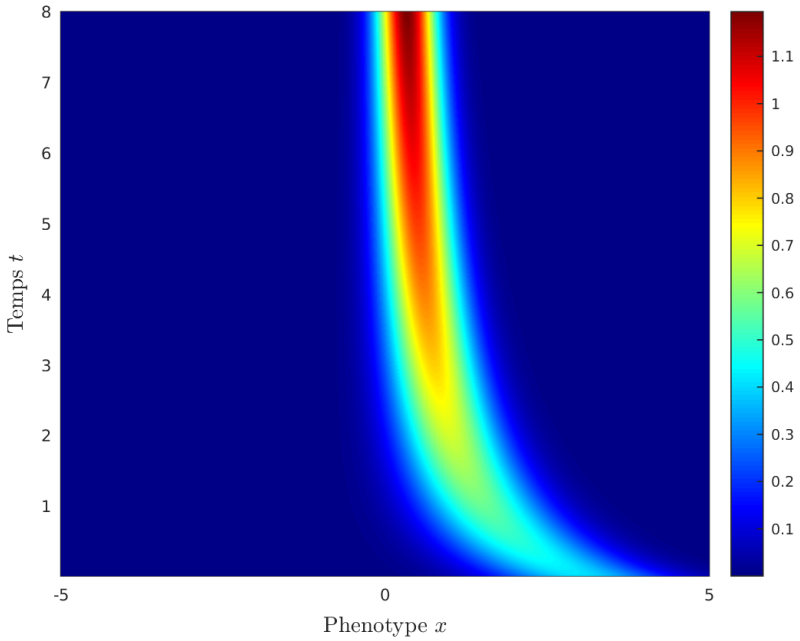



FIGURE 1.8. – **Évolution de la distribution des phénotypes, sous l'effet de la sélection,** donnée par un schéma numérique explicite. Le phénotype d'un individu n'est composé que d'un seul trait ($n = 1$). Initialement les phénotypes sont distribués selon une gaussienne réduite centrée en $x = 3$, tandis que la fitness malthusienne absolue est donnée par $r(x) = 1 - x^2/2$.

### Modèle individu-centré

Faisons une rapide description stochastique du processus évolutif et démographique. Pour cela, nous allons étudier un modèle discret (en temps et en phénotype), et donc non continu, contrairement au modèle de diffusion de Feller.

Une population, dont la fitness malthusienne absolue vaut $r$ est contante, est composée de $N_0$ individus initialement. On suppose que les générations ne se chevauchent pas, et que chaque individu donne naissance à un nombre $k$ d'individus, pour $k$ une variable aléatoire suivant une loi de Poisson de paramètre

$e^r$. Avec ces informations, nous appliquons le diagramme suivant pour faire évoluer notre population.

Pour pouvoir approcher ce modèle par le modèle de Feller, nous devons faire une approximation en effets faibles : les taux doivent être petits devant 1, tandis que $\mathrm{d}t = 1$.

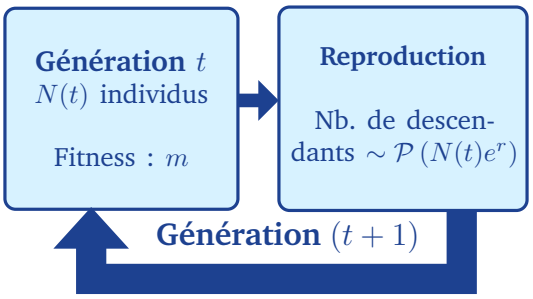



FIGURE 1.9. – Modèle individu-centré de reproduction.

En simulant 100 populations indépendantes, avec $N_0 = 100$ et $r = 0.2$, à partir du modèle décrit par la Fig. 1.9, nous obtenons la Fig. 1.10. La courbe grise est l'évolution d'une des simulations (appelées aussi *réplicats*). La moyenne sur tous les réplicats, représentée par des diamants magenta, correspond parfaitement à la théorie de Malthus, présentée en Section 1.2.1.

### 1.3.2. Effets des mutations

Bien que le modèle précédent sélectionne les individus les mieux adaptés, nous pouvons observer dans la nature une certaine diversité, que celui-ci ne peut expliquer en l'état. Modifions-le en considérant la possibilité de mutations. En effet, lors de la transcription de l'A.D.N., des erreurs (insertion ou délétion de portions d'A.D.N. sur un chromosome, remplacement d'un nucléotide par un autre, échange d'une partie de chromosome, *etc.*) peuvent se produire, appelées mutations. L'EDP (1.11) peut être modifiée pour prendre en compte ce phénomène. Nous le faisons en modélisant l'effet des mutations sur les phénotypes, qui à leur tour affectent les taux de croissance. Pour décrire la grande diversité possible d'effet des mutations, KIMURA 1965 a proposé une approche intégro–différentielle.

Un individu parent de phénotype $\mathbf{y}$ donne naissance à un individu enfant, qui par les mutations a un phénotype $\mathbf{x}$, selon une loi donnée. Supposons que cette loi est à densité, notée $K(\mathbf{x}-\mathbf{y})$. Ainsi la densité d'individus, nés au temps $t$ et ayant un phénotype $\mathbf{x}$ après mutation vaut $\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\, N(t,\mathbf{y})\, \mathrm{d}\mathbf{y}$. Cependant, des individus de parents $\mathbf{x}$ peuvent aussi muter aussi vers d'autres phénotypes : il y en a $\int_{\mathbb{R}^n} K(\mathbf{y}-\mathbf{x})\, N(t,\mathbf{x})\, \mathrm{d}\mathbf{y} = N(t,\mathbf{x})$. Finalement, l'effectif $N(t,\mathbf{x})$ est solution

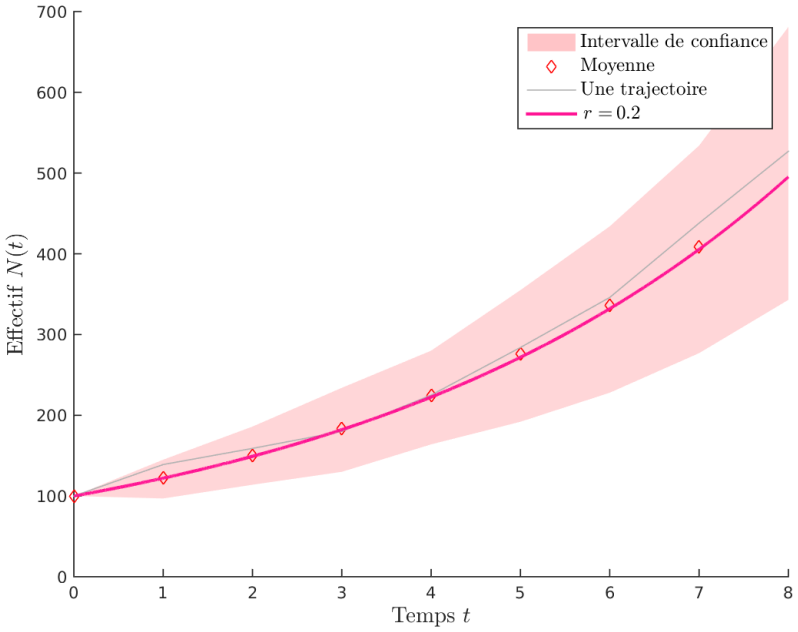



FIGURE 1.10. – **Évolution malthusienne de l'effectif $N(t)$ de population**, donnée par modèle individu-centré. Initialement, la population possède 100 individus et a une fitness malthusienne absolue $r = 0.2$. La courbe magenta représente l'évolution déterministe donnée par (1.3). La zone colorée représente l'intervalle entre les quantiles 0.025 et 0.975 et les points rouges en forme de diamant correspondent à la moyenne sur ces 100 simulations. La courbe grise correspond à l'évolution d'une des simulations stochastiques.

de :

$$\partial_t N(t,\mathbf{x}) = r(\mathbf{x})\, N(t,\mathbf{x}) + U\left[\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\, N(t,\mathbf{y})\,\mathrm{d}\mathbf{y} - N(t,\mathbf{x})\right], \qquad (1.14)$$

avec $U > 0$ le *taux de mutation*.

Du point de vue discret, le nombre de mutants produits en une génération est tiré dans une loi de Poisson de paramètre $UN(t)$. Le taux de mutations correspond donc au paramètre de cette loi, pour le modèle individu-centré. Pour un modèle continu, ce taux est le taux de naissance $b$ multiplié par la probabilité $p_U$ qu'une naissance donne un mutant, *i.e.*, $U = bp_U$. Dans ce cas, le taux de reproduction est considéré comme étant fixe (entre les phénotypes et au cours du temps) et seul le taux de mortalité est affecté par les phénotypes et l'environnement. Ainsi $U$ reste fixe.

En intégrant (1.14) par rapport à $\mathbf{x} \in \mathbb{R}^n$, $N(t)$ vérifie toujours (1.12), *via* la

propriété suivante, conséquence immédiate du théorème de Fubini :

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\,N(t,\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{x}\,N(t,\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_{\mathbb{R}^n} N(t,\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

Les deux équations (1.12) et (1.14) permettent d'avoir l'EDP satisfaite par la distribution $q(t,\mathbf{x})$ :

$$\partial_t q(t,\mathbf{x}) = [r(\mathbf{x}) - \overline{r}(t)]\,q(t,\mathbf{x}) + U\left[\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\,q(t,\mathbf{y})\,\mathrm{d}\mathbf{y} - q(t,\mathbf{x})\right], \quad (1.15)$$

pour $\overline{r}(t) = \int_{\mathbb{R}^n} r(\mathbf{x})\,q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ la fitness malthusienne absolue moyenne, $U > 0$ le taux de mutation et $K$ le noyau des effets des mutations.

Dans le cas où le noyau mutationnel ne possède pas une moyenne nulle, *i.e.*, $c := \int_{\mathbb{R}^n} \sum_{i=1}^n x_i K(\mathbf{x})\mathrm{d}x \neq 0$, alors les mutations provoquent une translation de la distribution des phénotypes, en plus d'un étalement, alors que si $c$ est nulle, nous n'observerons qu'un étalement, comme sur la Fig. 1.11). Pour le voir mathématiquement, nous utilisons dans ce paragraphe une hypothèse supplémentaire sur $K$ : le support de $K$ est supposé compact de rayon plus petit que $\varepsilon > 0$, et $K$ est de la forme $K(\mathbf{x}-\mathbf{y}) = \frac{1}{\varepsilon}J((\mathbf{x}-\mathbf{y})/\varepsilon)$, avec $J$ une fonction indépendante de $\varepsilon$. Dans le cadre de faible variance mutationelle, *i.e.*, $\varepsilon \to 0$, un développement formel du terme intégral implique que :

$$\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\,q(t,\mathbf{y})\,\mathrm{d}\mathbf{y} \simeq q(t,\mathbf{x}) + \varepsilon^n \int_{\mathbb{R}^n} \mathbf{y}'\,J(\mathbf{y}')\,\mathrm{d}\mathbf{y}' \cdot \nabla q(t,\mathbf{x}),$$

tant que $\mathbf{c} = \int_{\mathbb{R}^n} \mathbf{y}'\,J(\mathbf{y}')\,\mathrm{d}\mathbf{y}'$ est fini non nul (pour plus de détails voir PERTHAME 2009). On peut donc approximer l'EDP (1.15) par :

$$\partial_t q(t,\mathbf{x}) = [r(\mathbf{x}) - \overline{r}(t)]\,q(t,\mathbf{x}) + U\varepsilon^n\,\mathbf{c}\cdot\nabla q(t,\mathbf{x}).$$

Ainsi les mutations translatent approximativement la distribution $q$ à vitesse $\mathbf{c}$ dans l'espace phénotypique, au cours du temps. Cette équation s'applique à l'étude de virulence dans des systèmes immunitaires (voir BELLOMO ; DE ANGELIS ; PREZIOSI 2003 pour un modèle plus général de virulence).

Cependant, nous supposerons comme dans les modèles précédents que les mutations ont une moyenne nulle, ce qui implique une autre approximation, dite *diffusive*.

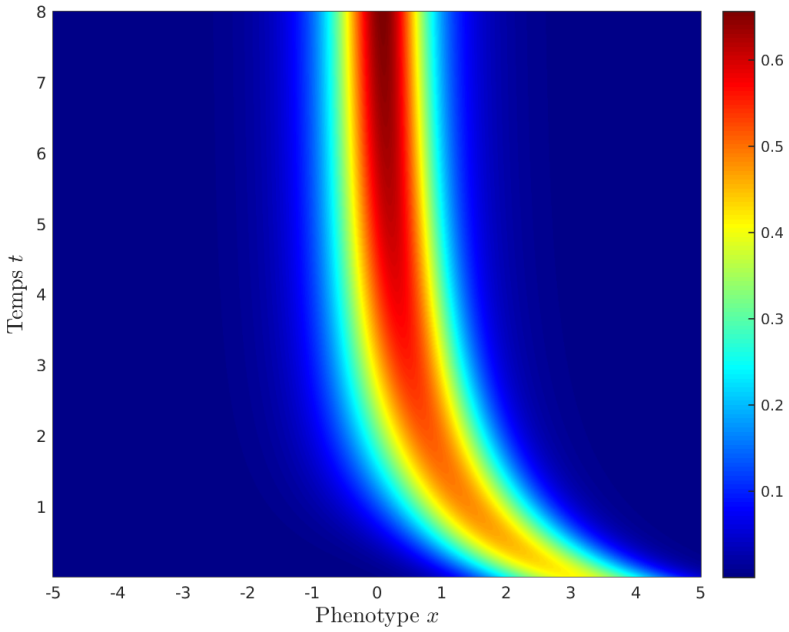



FIGURE 1.11. – **Évolution de la distribution des phénotypes, sous l'effet de la sélection et des mutations,** donnée par un schéma numérique explicite. Le phénotype d'un individu n'est composé que d'un seul trait ($n = 1$). Initialement les phénotypes sont distribués selon une gaussienne réduite centrée en $x = 3$, tandis que la fitness malthusienne absolue est donnée par $r(x) = 1 - x^2/2$. Les mutations sont distribuées selon une Gaussienne centrée réduite, et ont un taux $U = 0.5$.

**Approximation diffusive de (1.15).**

Supposons que $K$ soit pair et admette des moments d'ordre 2. Écrivons le développement de Taylor de $q(t, \mathbf{x} - \mathbf{y})$ en $\mathbf{x}$ :

$$q(t, \mathbf{x} - \mathbf{y}) = \sum_{k_1,\dots,k_n=0}^{\infty} (-1)^{k_1+\cdots+k_n} \frac{y_1^{k_1} \cdots y_n^{k_n}}{k_1! \cdots k_n!} \frac{\partial^{k_1+\cdots+k_n} q}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(t, \mathbf{x}),$$

et définissons les moments centrés de $K$ :

$$\omega_{k_1,\dots,k_n} = \int_{\mathbb{R}^n} y_1^{k_1} \cdots y_n^{k_n} \, K(\mathbf{y}) \, \mathrm{d}\mathbf{y}.$$

Cela implique formellement que :

$$\begin{aligned}\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\, q(t,\mathbf{y})\,\mathrm{d}\mathbf{y} &= \int_{\mathbb{R}^n} K(\mathbf{y})\, q(t,\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y},\\ &= \int_{\mathbb{R}^n} \sum_{k_1,\dots,k_n=0}^{\infty} \frac{(-1)^{k_1+\cdots+k_n}}{k_1!\cdots k_n!}\, y_1^{k_1}\cdots y_n^{k_n}\, K(\mathbf{y})\, \frac{\partial^{k_1+\cdots+k_n} p}{\partial x_1^{k_1}\cdots\partial x_n^{k_n}}(t,\mathbf{x})\,\mathrm{d}\mathbf{y},\\ &= \sum_{k_1,\dots,k_n=0}^{\infty} \frac{(-1)^{k_1+\cdots+k_n}}{k_1!\cdots k_n!} \int_{\mathbb{R}^n} y_1^{k_1}\cdots y_n^{k_n}\, K(\mathbf{y})\, \frac{\partial^{k_1+\cdots+k_n} p}{\partial x_1^{k_1}\cdots\partial x_n^{k_n}}(t,\mathbf{x})\,\mathrm{d}\mathbf{y},\\ &= \sum_{k_1,\dots,k_n=0}^{\infty} \omega_{k_1,\dots,k_n} \frac{(-1)^{k_1+\cdots+k_n}}{k_1!\cdots k_n!}\, \frac{\partial^{k_1+\cdots+k_n} p}{\partial x_1^{k_1}\cdots\partial x_n^{k_n}}(t,\mathbf{x}).\end{aligned}$$

Par parité de $K$, $\omega_{k_1,\dots,k_n}$ est nul si l'un des $k_i$ est impair. De plus, la matrice $M$ des moments d'ordre 2 – matrice de variance-covariance de $K$ – est symétrique réelle : il existe une matrice orthogonale $P \in O_n(\mathbb{R})$ et une matrice diagonale $\Lambda = \mathrm{diag}_n(\lambda_1,\dots,\lambda_n)$ avec $M = P^T \Lambda P$. En supposant que $\max_{1\leq i\leq n} \lambda_i \ll 1$ (voir Section 2.5 pour plus de détails dans le cas gaussien), les moments d'ordres supérieurs peuvent être formellement négligés, ce qui implique que :

$$\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\, q(t,\mathbf{y})\,\mathrm{d}\mathbf{y} \approx q(t,\mathbf{x}) + \sum_{i=1}^{n} \frac{\lambda_i}{2} \partial_{x_i x_i} q(t,\mathbf{x}),$$

puis :

$$\partial_t q(t,\mathbf{x}) \approx \sum_{i,j=1}^{n} m_{ij} \partial_{x_i x_j} q(t,\mathbf{x}) + [r(\mathbf{x}) - \overline{r}(t)] q(t,\mathbf{x}),$$

qu'on peut réécrire matriciellement en :

$$\partial_t q(t,\mathbf{x}) \approx \nabla\left[M \cdot \nabla q(t,\mathbf{x})\right] + [r(\mathbf{x}) - \overline{r}(t)] q(t,\mathbf{x}).$$

Étudions maintenant la fonction $\widetilde{q}(t,\mathbf{x}) = q(t,P\mathbf{x})$, solution de :

$$\partial_t \widetilde{q}(t,\mathbf{x}) \approx \nabla\left[\Lambda \cdot \nabla \widetilde{q}(t,\mathbf{x})\right] + [\widetilde{r}(\mathbf{x}) - \overline{r}(t)] \widetilde{q}(t,\mathbf{x}).$$

pour $\widetilde{r}(\mathbf{x}) = r(P\mathbf{x})$. Finalement, nous avons pu nous ramener à l'équation :

$$\partial_t q(t,\mathbf{x}) \approx \sum_{i=1}^{n} \frac{\mu_i^2}{2} \partial_{x_i x_i} q(t,\mathbf{x}) + [r(\mathbf{x}) - \overline{r}(t)] q(t,\mathbf{x}). \tag{1.16}$$

avec $\mu_i^2 = U\,\lambda_i$.

L'équation (1.16), connue sous le nom d'*équation replicator-mutator*, se retrouve appliquée dans bon nombre de domaines différents, comme le mentionne KOMAROVA 2004, dans le cadre d'un ensemble discret de valeurs de $\mathbf{x}$ : modèle de génétique des populations (BIKTASHEV 2014 ; HADELER 1981a ; HADELER 1981b ; PAGE ; NOWAK 2002 ; TAYLOR ; JONKER 1978), réseaux de réaction autocatalytique

(STADLER ; SCHUSTER 1992), modèle d'évolution du langage (KOMAROVA 2004 ; NOWAK ; KOMAROVA ; NIYOGI 2002).

Le cas $n = 1$ est étudié par ALFARO ; CARLES 2014b pour différentes fonctions $m(x)$. Si la fitness est linéaire – $m(x) = x$ – la solution est définie en tout temps, mais se déplace en accélérant vers la droite, quand $t$ tend vers $+\infty$. Le deuxième cas est $m(x) = x^2$. Celui-ci produit l'extinction de la population en un temps fini, dans le sens où il existe un temps $t_{\text{ext}} > 0$ tel que pour tout temps $t > t_{\text{ext}}$ et pour tout $x \in \mathbb{R}$, $q(t, x) = 0$. Le dernier cas est le plus intéressant (voir 1.4) : il s'agit de $m(x) = -x^2$. Ici, la population va finir par être distribué selon une loi gaussienne. Notons que dans les deux cas quadratiques précédents, une formule explicite peut être donnée pour $q(t, x)$ (voir aussi ALFARO ; CARLES 2017). ALFARO ; VERUETE 2018 développent, quant à eux, le comportement en temps long de $q(t, x)$ pour une gamme plus large de fonctions $m(x)$ dites *confinantes*, dans le sens où $m(x) \to -\infty$ quand $|x| \to +\infty$.

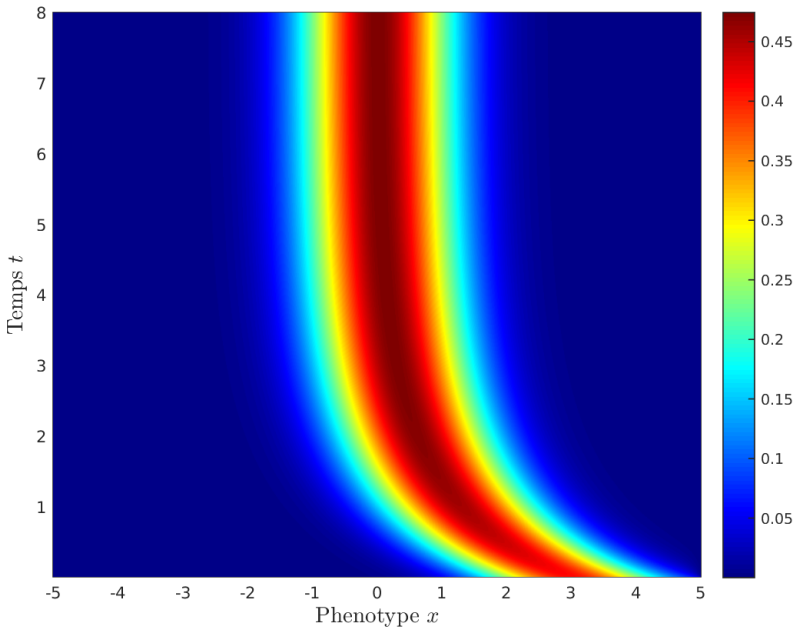



FIGURE 1.12. – **Évolution de la distribution des phénotypes, sous l'effet de la sélection et des mutations, avec approximation diffusive,** donnée par un schéma numérique explicite. Le phénotype d'un individu n'est composé que d'un seul trait ($n = 1$). Initialement les phénotypes sont distribués selon une gaussienne réduite centrée en $x = 3$, tandis que la fitness malthusienne absolue est donnée par $r(x) = 1 - x^2/2$. Le coefficient de diffusion $\mu$ est pris tel que $\mu^2 = 0.5$.

### 1.3.3. Modèle de Wright-Fisher avec sélection et mutations

Cette section est dédiée à la description d'un processus stochastique, décrivant l'évolution d'une population sous l'effet conjoint de la sélection et des mutations.

**Sewall Green Wright** (1889 – 1988), généticien américain, connu pour ses travaux sur la théorie de l'évolution et sur l'analyse des relations structurelles en statistiques, est l'un des pionniers de la théorie de l'évolution et de la génétique des populations de nos jours, aux côtés de R. A. Fisher et de J. Haldane. Convaincu que les forces évolutives – sélection, mutations, migration – sont des processus importants pour l'adaptation d'une population, tout comme les phénomènes aléatoires (brassage génétique, *etc.*), regroupés sous le nom de *dérive génétique*, il développe un modèle stochastique, permettant de décrire la dérive génétique.

Ce modèle suppose que la population est asexuée et de taille constante $N$. Supposons qu'à la génération $t$, la population soit composée de $N$ individus de phénotypes $(\mathbf{x}_t^1, \ldots, \mathbf{x}_t^N)$. On note $m_t^i$ la fitness associée au phénotype $\mathbf{x}_t^i$. Entre les générations $t$ et $(t+1)$, la reproduction et la sélection vont opérer par un tirage avec remise de $N$ individus parmis les individus de la génération précédente : cela correspond à l'utilisation d'une loi multinomiale avec pour poids de sélection $\left(\exp[m_t^1], \ldots, \exp[m_t^N]\right)$, à normalisation près. La population est alors composée de phénotypes $(\widetilde{\mathbf{x}}_t^1, \ldots, \widetilde{\mathbf{x}}_t^N)$. Ensuite, les mutations agissent sur chaque individu au nombre de $k$, pour $k$ tiré selon une loi de Poisson de paramètre le taux de mutation $U$. Chacune des mutations transforme le phénotype d'un individu $\widetilde{\mathbf{x}}_t^i$ en $\widetilde{\mathbf{x}}_t^i + \mathrm{d}\mathbf{x}_t^i$, pour $\mathrm{d}\mathbf{x}_t^i$ une variable aléatoire. Ici nous considérerons que la loi de $\mathrm{d}\mathbf{x}_t^i$ est une loi gaussienne centrée de matrice de variance-covariance $\lambda I_n$. Ce processus est reporté sur le schéma de la Fig.1.13.

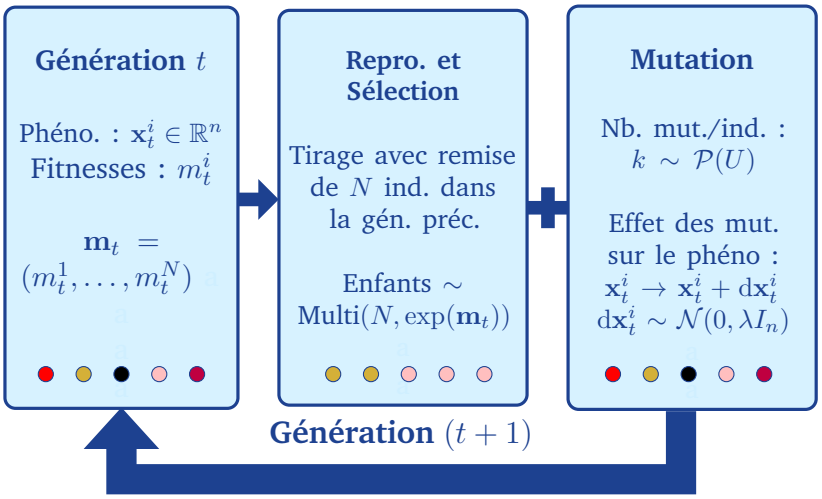



FIGURE 1.13. – **Modèle de Wright-Fisher avec effet mutationnel gaussien sur les phénotypes.**

### Généralisation avec démographie

Ce dernier modèle peut être généralisé pour prendre en compte une taille de population variable. Cette généralisation est nécessaire dans le cadre de l'étude de l'impact de la migration sur l'évolution de la population, étant donné que la composition de la population changera à chaque génération, et donc aussi le nombre d'individus.

**Sir Ronald Aylmer Fisher** (1890 – 1962), biologiste statisticien britannique, considéré comme le « plus grand des successeurs de Darwin » (DAWKINS 1995), a « presque à lui tout seul fondé les statistiques modernes » (ANDERS 1998), en ayant introduit des notions importantes telle que le maximum de vraisemblance, et des méthodes d'estimation et d'inférence incontournables aujourd'hui (inférence fréquentiste ou objective), décrites dans son article *On the mathematical foundations of theoretical statistics*.

Grâce à ces études de statistiques, au départ appliquées en physique et en astronomie (théorie des erreurs, théorie des gaz), il construit les fondements de la génétique des populations, en collaboration avec les généticiens S. G. Wright et J. Haldane. Il relie la sélection naturelle à la génétique mendélienne, et explique l'existence de traits neutres, par sa *théorie de l'emballement fisherien*.

Il développe aussi des principes cruciaux pour la planification expérimentales, dans *The Design of Experiments* (1935).

## 1.4. Modèle de Fisher – Paysage adaptatif

Maintenant que nous avons décrit les modèles théoriques représentant les effets respectifs de la sélection et des mutations, il nous faut expliciter le choix de la fonction fitness $r(\mathbf{x})$. Pour cela nous allons utiliser le Modèle Géométrique de Fisher (FGM), qui donne une relation explicite entre le phénotype d'un individu et sa fitness : ce modèle est un exemple de paysage adaptatif. Il en existe de nombreuses autres formes.

Sir R. Fisher fait l'hypothèse de l'existence d'un phénotype optimal unique, au sens où il existe un phénotype $\mathbf{x}^* \in \mathbb{R}^n$ (qu'on peut supposer sauf mention contraire être à l'origine du repère phénotypique, *id est*, $\mathbf{x}^* = (0, \dots, 0)$) tel que :

$$\forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{x}^*\},\ r(\mathbf{x}) < r(\mathbf{x}^*).$$

De plus, la fitness malthusienne absolue décroît, quand le phénotype s'éloigne de l'optimum $\mathbf{x}^*$. Dans nos modèles ainsi que dans LANDE 1980 et KIMURA 1965, cette décroissance est choisie quadratique, ce qui implique que :

$$r(\mathbf{x}) = r_{\max} - \frac{\|\mathbf{x} - \mathbf{x}^*\|^2}{2}.$$

Ce choix implique la définition auxiliaire de *fitness malthusienne relative* :

$$m(\mathbf{x}) = -\frac{\|\mathbf{x} - \mathbf{x}^*\|^2}{2},$$

et donc de *fitness malthusienne relative moyenne* $\overline{m}(t) = \overline{r}(t) - r_{\max}$.

Comme TENAILLON 2014 le mentionne, ce modèle prend en compte bon nombre de propriétés importantes du point de vue biologique de l'effet des mutations sur la fitness. En effet, l'une des plus importantes est l'épistasie – la dépendance de la distribution des effets mutationnels sur la fitness, par rapport à la fitness du parent. Comme nous pouvons le voir sur le schéma de la Fig. 1.14, le phénotype parent $\mathbf{x}_1$ peut avoir un enfant qui mute, et qui se trouve dans une des trois zones colorées autour de $\mathbf{x}_1$ :

- ▷ Si le phénotype de l'enfant est dans le disque ouvert de centre $\mathbf{x}^*$ et passant par le phénotype parent $\mathbf{x}_1$ (zone jaune dans la Fig. 1.14), alors il est plus proche de l'optimum $\mathbf{x}^*$ que $\mathbf{x}_1$. Sa fitness est donc plus grande que la fitness de son parent : il sera mieux adapté, grâce à des mutations, dites *bénéfiques* ;
- ▷ Si le phénotype de l'enfant est sur le cercle de centre $\mathbf{x}^*$ et passant par $\mathbf{x}_1$ (ligne verte dans la Fig. 1.14), il aura la même fitness que le parent : il sera aussi bien adapté. La mutation est alors dite *neutre* ;
- ▷ Si le phénotype de l'enfant est en dehors du disque fermé de centre $\mathbf{x}^*$ et passant par $\mathbf{x}_1$ (zone bleue dans la Fig. 1.14), il sera plus loin de l'optimum

$\mathbf{x}^*$ que son parent, et donc aura une fitness plus petite : il sera moins bien adapté, à cause de mutations, dites *délétères*.

Sur la Fig. 1.14, nous remarquons que ces zones diffèrent selon le phénotype parental, et que les phénotypes proches (comme $\mathbf{x}_1$) de l'optimum $\mathbf{x}^*$ ont moins de mutations bénéfiques que les phénotypes loin de celui-ci (comme $\mathbf{x}_3$). Ainsi le noyau des mutations agissant sur la fitness dépend bien de la fitness du parent.

Au-delà d'allier une certaine simplicité et une part de réalisme, le modèle est particulièrement utile pour aborder les habitats hétérogènes. En effet, dans ce cas, la notion d'optimum permet de caractériser ces habitats, et la notion d'adaptation locale. Un phénotype est plus ou moins bien adapté à un habitat donné selon qu'il se trouve proche ou non de l'optimum specifique à cet habitat. Ainsi des individus bien adaptés dans un environnement 1, peuvent être très mal adaptés dans un environnement 2, et vice et versa. L'effet des mutations peut être décrit comme étant dépendant des fitness de l'individu dans chaque environnement. MARTIN ; LENORMAND 2015 détaillent les différentes distributions des mutations sur un couple de fitness $(m_1, m_2)$ avec $m_i$ la fitness de l'individu s'il était dans l'environnement $i$.

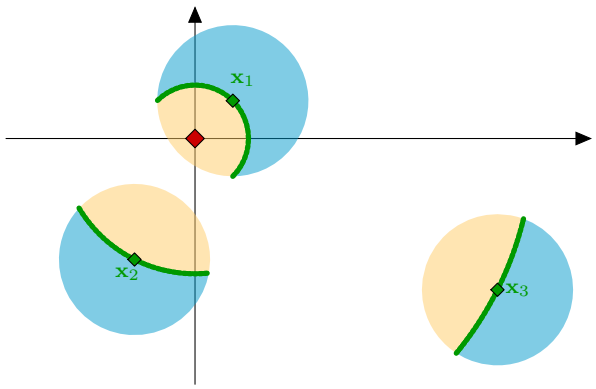



FIGURE 1.14. – **Schéma du paysage adaptatif à un seul optimum phénotypique en 2D, décrit par le Modèle Géométrique de Fisher.** Les points verts représentent différents phénotypes parents $\mathbf{x}_1$, $\mathbf{x}_2$ et $\mathbf{x}_3$, tandis que les lignes vertes représentent des niveaux d'isofitness. Les zones jaunes (resp. bleues) représentent les mutations bénéfiques (resp. délétères) du parent associé. On remarque que les phénotypes proches de l'optimum (point rouge) ont moins de mutations délétères : le noyau mutationnel dépend donc de la fitness du parent.

## 1.5. Méthodologie générale et outils mathématiques

Au cours de cette thèse, nous avons cherché à :

(i) Modéliser l'évolution d'une population asexuée par le biais de différents IBM ;

(ii) Prendre en considération l'épistasie, en utilisant le modèle géométrique de Fisher ;

(iii) Intégrer une structure spatiale, *via* des flux migratoires entre deux environnements ;

(iv) Modéliser l'évolution d'une population asexuée par le biais d'EDP linéaires, permettant ainsi d'obtenir des prédictions analytiques de ces évolutions.

Sous l'hypothèse générale de mutations anisotropes, ces recherches se sont basées sur l'équation d'évolution suivante, vérifiée par la distribution phénotypique $q$ :

$$\forall t > 0,\ \forall \mathbf{x} \in \mathbb{R}^n, \quad \partial_t q(t,\mathbf{x}) = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\, \partial_{x_i x_i} q(t,\mathbf{x}) + [r(\mathbf{x}) - \overline{r}(t)] q(t,\mathbf{x}),$$
$$\text{avec} \quad \int_{\mathbb{R}^n} q(t,\mathbf{x}) \mathrm{d}\mathbf{x} = 1, \quad (1.17)$$

décrivant l'adaptation d'une population asexuée dans un habitat fixe, se faisant à l'aide de la sélection et de mutations. Nous avons vu dans la Section 1.3.2 que cette équation est une approximation d'une équation intégro-différentielle plus complexe (cf Annexe C pour plus de détails). En effet, cette dernière ne permet pas d'avoir facilement des résulats analytiques en tout temps, mais des résultats en temps longs.

Par ailleurs, une population asexuée subit aussi la dérive génétique (*cf.* Section 1.2.3.2 ou Annexes A et B). Cependant, MARTIN ; ROQUES 2016 ont montré qu'une hypothèse de taille de population suffisamment grande (ou taille infinie) permet de négliger son effet sur la distribution phénotypique. Nous ferons ainsi cette hypothèse tout au long du manuscrit, sauf mention contraire.

Ces équations paraboliques, que nous considérons, ne présentent pas de solutions explicites générales, sauf dans certains cas particuliers (*e.g.*, hypothèse gaussienne dans les Chapitres 2 et 3). Cependant, il est possible de décrire quantitativement différentes propriétés sur la solution de ces équations, comme nous allons le voir dans la Section 1.5.1.

Étant donné que nous n'arrivons pas à obtenir des informations en tout temps, *via* l'équation phénotypique, nous n'allons pas nous focaliser sur la distribution des phénotypes, mais plutôt sur la fitness malthusienne absolue moyenne $\overline{r}(t)$, ce qui permettra alors de savoir, si la population s'adaptera en moyenne à l'environnement (cas où $\lim_{t\to+\infty} \overline{r}(t) > 0$) ou pas (cas où $\lim_{t\to+\infty} \overline{r}(t) \leq 0$). Pour ce faire,

nous étudierons la distribution de fitness qui vérifient une équation parabolique dégénérée (voir Section 2.2 pour plus de détails), et sa *transformée de Laplace*, appelée *Fonction Génératrice des Moments* (MGF). Cette dernière vérifie une équation de transport, dont l'étude est généralement menée par la méthode des caractéristiques, présentée en Section 1.5.2.

### 1.5.1. Outils d'analyses d'équations paraboliques

Cette équation différentielle parabolique (1.17) n'est pas standard :

▷ Le terme de sélection $f(q,t,\mathbf{x}) = [r(\mathbf{x}) - \overline{r}(t)]q(t,x)$ est non local, mais aussi non linéaire, suite au terme :

$$\overline{r}(t) = \int_{\mathbb{R}^n} r(\mathbf{x})q(t,\mathbf{x})\mathrm{d}\mathbf{x}.$$

Des équations de réaction-diffusion avec un terme de réaction $f$ non local spécifique ont été étudiées (*e.g.* $f(q,t,x) = q(t,x)(1-\Phi * q(t,x)$, pour $\Phi$ une fonction positive d'intégrale 1, voir BERESTYCKI ; NADIN ; PERTHAME ; RYZHIK 2009 ; FANG ; ZHAO 2011 ; FAYE ; HOLZER 2015 ; GENIEYS ; VOLPERT ; AUGER 2006b ; GOURLEY 2000 ; HAMEL ; RYZHIK 2014a, ou $f(q,t,x) = q(t,x)(q(t,x)-\theta)(1-\Phi * q(t,q))$, pour $0 < \theta < 1$, voir ALFARO ; COVILLE ; RAOUL 2014), bien que l'une des principales barrières à leurs études soit l'absence de principe de comparaison.

▷ Le terme de sélection $f(q,t,\mathbf{x}) = [r(\mathbf{x}) - \overline{r}(t)]q(t,\mathbf{x})$ possède un coefficient qui est non borné ($\mathbf{x} \mapsto r(\mathbf{x})$ sur $\mathbb{R}^n$). Cela pourrait, si le coefficient était positif, entrainer une explosion de la solution en temps fini (ALFARO ; CARLES 2014a ; ALFARO ; CARLES 2017 ; ALFARO ; VERUETE 2018).

Nous verrons que cette équation (munie d'une condition initiale admet une unique solution régulière (CHABROWSKI 1970 ; LUNARDI 1998), bien que les coefficients de notre opérateur ne soit pas bornés. Bien qu'elle ne soit pas explicite, une telle solution pourra être comparée à des fonctions, par le biais de principes du maximum et de comparaison (ARONSON ; BESALA 1967 ; PROTTER ; WEINBERGER 1967).

#### Cas des systèmes paraboliques

Lorsque la migration est prise en compte, le modèle se complexifie en un système différentiel parabolique de la forme :

$$\partial_t \mathbf{q}(t,\mathbf{x}) + L\mathbf{q}(t,\mathbf{x}) = H\mathbf{q}(t,\mathbf{x}),$$

avec $L$ une matrice diagonale :

$$L = \begin{pmatrix} L_1 & & (0) \\ & \ddots & \\ (0) & & L_k \end{pmatrix},$$

où $L_\mu$, opérateur modélisant la sélection, les mutations et l'émigration dans l'environnement $\mu$, est de la forme :

$$L_\mu q(t,\mathbf{x}) = -\sum_{i=1}^{n} \frac{\mu_{\mu,i}^2}{2}\, \partial_{x_i x_i} q(t,\mathbf{x}) \\ - \left[ -\frac{\|\mathbf{x}-\mathbf{O}_\mu\|^2}{2} + \int_{\mathbb{R}^n} \frac{\|\mathbf{y}-\mathbf{O}_\mu\|^2}{2} q(t,\mathbf{y})\mathrm{d}\mathbf{y} \right] q(t,\mathbf{x}) + \delta q(t,\mathbf{x}),$$

tandis que $H$ représentant immigration est la matrice :

$$H = \frac{\delta}{k-1} \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}.$$

Ici le coefficient $\delta > 0$ représente le taux d'émigration qui est supposé indépendant de l'environnement $\mu$. De plus, nous faisons l'hypothèse que chaque individu migrant peut aller avec la même probabilité dans chaque environnement.

Étant donné la forme de la matrice $H$, ce système évolutif rentre dans la catégorie des systèmes *coopératifs entièrement couplés* (Besala 1979 ; Sweers 1992). Ainsi l'existence de solutions (Besala 1979 ; Sweers 1992), des principes du maximum (Weinberger 1975) et des inégalités de Harnack (Sirakov 2009) pourront être étendus au cas de système.

Étant donnée la complexité de ce modèle nous nous contenterons de traiter seulement le cas de deux environnements connectés, *i.e.*, $k = 2$.

### 1.5.2. Fonctions génératrices et équations de transport

Supposons qu'au temps $t$ la population soit composée de phénotypes $\{\mathbf{x}_i\}_{i=1,\dots,N}$ (avec $N \in \mathbb{N}$). On note les fitness malthusiennes relatives associées $\{m_i(t)\}_{i=1,\dots,N}$. En l'absence de démographie et de migration, l'évolution de la distribution des fitness est généralement étudiée grâce à une approximation par clôture des moments : les variations du moment d'ordre $k$ dépend linéairement du moment d'ordre $(k+1)$, *via* une EDO linéaire, et ce système se résout alors en faisant l'hypothèse que les moments d'ordre supérieur à une valeur fixée s'annulent. L'exemple le plus connu en biologie évolutive est le « théorème fondamental de Fisher » (Fisher 1930) :

**Théorème 1. (THEOREME FONDAMENTAL DE FISHER)**
*Sous l'effet de la sélection seule, le taux de croissance de la fitness malthusienne relative moyenne d'une population est égale à la variance génétique de la fitness (*FISHER *1930).*
*Si de plus les mutations agissent (avec un taux $U > 0$ et de variance $\lambda$ sur les $n$ traits biologiques), alors la fitness malthusienne relative moyenne $\overline{m}(t)$ vérifie :*

$$\overline{m}'(t) = V - \frac{\sqrt{U\lambda}n}{2}.$$

Cette équation ne peut être résolue sans hypothèse supplémentaire, telle qu'une hypothèse de moments fermés. Récemment, l'étude des *fonctions génératrices des cumulants (CGF)* (BURGER 1991 ; MARTIN ; ROQUES 2016 ; TRAN ; HOFRICHTER ; JOST 2014) a permis de fermer cette équation (voir la méthode en Section 1.5.2.1). L'équation satisfaite par la CGF contient en fait toutes celles vérifiées par les moments (comme dans le Théorème 1), et peut être utilisée pour décrire la dynamique de la fitness malthusienne relative moyenne dans des modèles avec ou sans épistasie (MARTIN ; ROQUES 2016). Par exemple, la fonction génératrice des moments (MGF) et celle des cumulants (CGF), étant définies respectivement par :

$$M(t,z) = \frac{1}{N}\sum_{i=1}^{N} e^{m_i(t)\,z}, \qquad \text{et} \qquad C(t,z) = \log M(t,z),$$

pour tout $z \geq 0$, sont reliées à la fitness malthusienne relative moyenne et à la variance en fitness par $\partial_z C(t,0) = \overline{m}(t)$ et $\partial_{zz} C(t,0) = V(t)$.

Toutefois, les fonctions $M(t,z)$ et $C(t,z)$ sont en soit des fonctionnelles d'un processus aleatoire multivarié avec un nombre infini d'états (puisque la mutation ici peut produire une infinité de nouveaux phénotypes aux fitness différentes). On va donc étudier l'espérance de ces fonctionnelles sur l'ensemble des processus stochastiques démographiques et mutationnels. De là, nous tenterons de trouver une dynamique déterministe approximative à cette espérance. Nous continuerons de noter $C(t,z)$, *etc.* toutes les approximations déterministes des quantités vues précédemment. Si une confusion peut être faite, nous utiliserons la notation $\langle X \rangle$ pour représenter la moyenne d'une variable aléatoire $X$ sur les réplicats.

**Remarque 1.** *Pour certains cas (cf. Chapitres 2 et 3), nous n'utiliserons pas directement la fitness $m$ mais un vecteur fitness $(m_1, \ldots, m_n)$. Cela conduira à une transformée de Laplace évaluée en $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{R}_+^n$.*

Cette dernière fonction vérifiera un certain type d'équations aux dérivées

partielles bien connues, appelées *équation de transport non locale* de la forme :

$$\partial_t M(t,\mathbf{z}) = a(t,\mathbf{z})\cdot\nabla M(t,\mathbf{z}) - a(t,0)\cdot\nabla M(t,0) + b(\mathbf{z})\, M(t,\mathbf{z}), \tag{1.18}$$

pour des fonctions $a:\mathbb{R}^n\to\mathbb{R}^n$ et $b:\mathbb{R}^n\to\mathbb{R}$ continues particulières et $\nabla M(t,\mathbf{z})$ le vecteur composé par les différentes dérivées partielles de $M$ par rapport aux composantes de $\mathbf{z}$. L'intérêt de cette équation est d'avoir une solution explicite (pour certaines fonctions $a$ et $b$) et ainsi avoir la fitness malthusienne relative moyenne $\overline{m}(t)$ qui n'est autre que $a(t,0)\cdot\nabla M(t,0)$ (voir les Chapitres C, 2 et 3 pour plus de détails). Cette section a pour but d'expliquer la méthode générale, dite *méthode des caractéristiques*, pour trouver une solution explicite de l'EDP (1.18).

#### 1.5.2.1. Méthode des caractéristiques

Dans cette section, nous développons la méthode des caractéristiques pour résoudre (1.18), dans le cas où $n=1$ et les fonctions $a$ et $b$ ne dépendent pas du temps.

Supposons qu'il existe $y\in C^1(\mathbb{R}_+,\mathbb{R}_+)$ inversible solution de :

$$\begin{cases} y'(z) = a(y(z)), \quad \forall z>0,\\ y(0)=0. \end{cases}$$

Posons la fonction auxiliaire $F(t,z)=\log[M(t,y(z))]$, satisfaisant :

$$\begin{aligned}\partial_t F(t,z) &= \frac{\partial_t M(t,y(z))}{M(t,y(z))},\\ &= a(y(z))\frac{\partial_z M(t,y(z))}{M(t,y(z))} - a(0)\frac{\partial_z M(t,y(0))}{M(t,y(z))} + b(y(z)),\\ &= \partial_z F(t,z) - \partial_z F(t,0) + b(y(z)).\end{aligned}$$

Maintenant, si nous arrivons à déterminer une formule pour $F$, dans ce cas, nous aurons une formule pour $M$, étant donné que $M(t,z)=\exp\left[F(t,y^{-1}(z))\right]$.

Fixons $z\in\mathbb{R}_+$ et définissons $G(t)=F(t,z-t)$ pour $t\in[0,z]$. Cette fonction est solution de :

$$G'(t) = -\partial_z F(t,0) + b(y(z-t)).$$

Ainsi nous avons pour tout $0\le t\le z$ :

$$F(t,z-t) = G(t) = F(0,z) - \int_0^t \partial_z F(s,0)\,\mathrm{d}s + \int_0^t (b\circ y)(z-s)\,\mathrm{d}s,$$

ce qui implique que pour tout $t \geq 0$ et pour tout $z \geq 0$ :

$$\begin{aligned} F(t,z) &= F(0,z+t) - \int_0^t \partial_z F(s,0)\,\mathrm{d}s + \int_0^t (b\circ y)(z+t-s)\,\mathrm{d}s, \\ &= F(0,z+t) - \int_0^t \partial_z F(s,0)\,\mathrm{d}s + \int_0^t (b\circ y)(z+s)\,\mathrm{d}s. \end{aligned}$$

Cette relation n'est pas fermée : $F(t,z)$ est exprimée en fonction de $\partial_z F(t-s,0)$, que nous ne connaissons pas. Dérivons cette relation par rapport à $z$ et évaluons ensuite en $z=0$ :

$$\partial_z F(t,0) = \partial_z F(0,t) + \int_0^t (b\circ y)'(s)\,\mathrm{d}s = \partial_z F(0,t) + b(y(t)) - b(0).$$

Ainsi nous obtenons une formule explicite pour $\partial_z F(t,0)$, donc aussi pour $F(t,z)$, puis pour $M(t,z)$.

**Remarque 2.** *Cette méthode de calcul peut se généraliser à $n \geq 0$ et à des fonctions $a$, $b$ dépendant du temps $t$. Pour cela, il suffit de trouver une fonction $\mathbf{y} : \mathbb{R}_+ \times \mathbb{R}_+^n \to \mathbb{R}_+^n$ de classe $C^1$ et inversible vérifiant :*

$$\begin{cases} \partial_t y_1(t,\mathbf{z}) - \partial_{z_1} y_1(t,\mathbf{z}) &= -a_1(t,\mathbf{y}(t,\mathbf{z})), \\ \qquad\vdots & \qquad\vdots \\ \partial_t y_n(t,\mathbf{z}) - \partial_{z_n} y_n(t,\mathbf{z}) &= -a_n(t,\mathbf{y}(t,\mathbf{z})), \\ \mathbf{y}(t,0) = 0, \end{cases}$$

*pour $a(t,\mathbf{z}) = (a_1(t,\mathbf{z}),\dots,a_n(t,\mathbf{z}))$, et de définir $F(t,\mathbf{z}) = \log[M(t,\mathbf{y}(t,\mathbf{z}))]$.*

### 1.5.3. Fonctions hypergéométriques

Ces équations nous amènerons à utiliser des fonctions particulières appelées *fonctions hypergéométriques*. Il s'agit de fonctions dépendant de deux paramètres $(a_1,\dots,a_p) \in \mathbb{R}^p$ et $(b_1,\dots,b_q) \in \mathbb{R}^q$.

La première fonction hypergéométrique à avoir été définie fut la fonction hypergéométrique de Gauss. Cette fonction est définie par la série entière :

$${}_2F_1(a_1,a_2;b;z) = 1 + \sum_{k=1}^{\infty} \frac{a_1(a_1+1)\cdots(a_1+k-1)\,a_2(a_2+1)\cdots(a_2+k-1)}{b(b+1)\cdots(b+k-1)}\,\frac{z^k}{k!},$$

pour $|z|<1$.

**Remarque 3.** *La fonction ${}_2F_1(a_1, a_2; b; z)$ peut être aussi définie comme étant la solution développable en série entière de l'EDO :*

$$z\,(1-z)\,y''(z) + [b - (a_1 + a_2 - 1)]\,y'(z) - a_1\,a_2\,y(z) = 0,$$

*avec* $y(0) = 1$.

Pour simplifier cette définition, nous pouvons utiliser les *symboles de Pocchamer (croissants)* :

$$(a)_k = \begin{cases} 1, & \text{si } k = 0, \\ a(a+1)\cdots(a+k-1), & \text{sinon.} \end{cases}$$

Cela implique que ${}_2F_1(a_1, a_2; b; z) = \displaystyle\sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k}{(b)_k}\,\frac{z^k}{k!}$.

**Remarque 4.** *Les symboles de Pochhammer peuvent être définis par le biais de la fonction Gamma* $\Gamma$. *Comme pour tout* $z$ *avec* $Re(z) > 0$, $\Gamma(z+1) = z\Gamma(z)$, *on sait que pour tout* $a$ *et pour tout* $k \in \mathbb{N}$, $\Gamma(a+k) = (a+k-1)\cdots a\Gamma(a)$, *ce qui implique que :*

$$(a)_k = \frac{\Gamma(a+k)}{\Gamma(a)}.$$

Nous pouvons donc maintenant définir plus généralement la *fonction hypergéométrique* ${}_pF_q$, définie pour $|z| < 1$ par la série entière :

$${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k}\,\frac{z^k}{k!},$$

pour des paramètres $(a_1, \ldots, a_p) \in \mathbb{R}^p$ et $(b_1, \ldots, b_q) \in \mathbb{R}^q$.

**Exemple.** *Certaines fonctions hypergéométriques particulières se ramènent à des fonctions usuelles :*

- ◇ *Pour* $p = q = 0$, *on a* ${}_0F_0(;;z) = \exp(z)$ ;
- ◇ *Pour* $(p, q) = (1, 0)$ *et* $a_1 = \alpha$, *on a* ${}_1F_0(\alpha;;z) = (1-z)^{-\alpha}$ ;
- ◇ *Pour* $(p, q) = (0, 1)$ *et* $b_1 = 1/2$, *on a* ${}_0F_1(;1/2;z^2/4) = \cosh(z)$ ;
- ◇ *Pour* $(p, q) = (0, 1)$ *et* $b_1 = 3/2$, *on a* ${}_0F_1(;3/2;z^2/4) = \sinh(z)/z$.

■

### 1.5.4. Fonction de Lambert

Il nous reste une dernière fonction à étudier : la fonction dite *de Lambert*. Il s'agit d'un ensemble de fonctions, et plus précisément des différentes branches inverses de la fonction $f(x) = x\, e^x$. Il existe uniquement deux branches à valeurs réelles, notées :

$$W_0 : [-\,e^{-1}, +\infty) \to [-1/e, +\infty) \text{ and } W_{-1} : [-\,e^{-1}, 0) \to (-\infty, -1/e].$$

**Proposition 2.** *Considérons quatre réels $a$, $b$, $c$ et $d$ tels que $a\, c \neq 0$. Posons le réel $\Delta$ défini par :*

$$\Delta = \frac{a\, b}{c} \exp\left(-\,\frac{d\, b}{c}\right).$$

*Alors :*

- *si $\Delta \geq 0$ ou $\Delta = -e^{-1}$, il existe une unique solution $x \in \mathbb{R}$ à l'équation :*

$$a\, e^{bx} + c\, x + d = 0.$$

- *si $\Delta \in (-e^{-1}, 0)$, il existe exactement deux solutions $x \in \mathbb{R}$ à l'équation :*

$$a\, e^{bx} + c\, x + d = 0.$$

- *si $\Delta < -\,e^{-1}$, l'équation $a\, e^{bx} + c\, x + d = 0$ n'admet pas de solution réelle.*

*De plus, une solution de cette équation (si elle existe) peut s'écrire sous la forme :*

$$x = -\,\frac{d}{c} - \,\frac{1}{b}\, W(\Delta).$$

*Démonstration.* L'idée de la preuve est de réécrire l'équation $a\, e^{bx} + c\, x + d = 0$ en $X\, e^X = Y$ :

$$\begin{aligned} a\, e^{bx} + c\, x + d = 0 &\Leftrightarrow \frac{a\, b}{c} e^{bx} = -b\, x - \,\frac{d\, b}{c}, \\ &\Leftrightarrow e^{-bx}\left(-b\, x - \,\frac{d\, b}{c}\right) = \frac{a\, b}{c}, \\ &\Leftrightarrow X\, e^X = \Delta, \end{aligned}$$

avec $X = -b\, x - \,\frac{d\, b}{c}$. On en déduit le théorème et la formule, grâce à la définition de $W$. □

**Proposition 3.** *Soit* $(a, b, c) \in \mathbb{R}^\star \times \mathbb{R}_\star^+ \times \mathbb{R}^\star$. *Les solutions de l'équation* $e^{ax} = b\,(1+cx)^2$ *sont :*

$$x = -\frac{1}{c} - \frac{2}{a} W\left[\frac{\pm a}{2c\sqrt{b}} \exp\left(-\frac{a}{2c}\right)\right].$$

*Démonstration.* Nous avons :

$$\begin{aligned} e^{ax} = b\,(1+cx)^2 &\Leftrightarrow \left[\sqrt{b}(1+cx)e^{-ax/2}\right]^2 = 1, \\ &\Leftrightarrow \sqrt{b}(1+cx)e^{-ax/2} = \pm 1, \\ &\Leftrightarrow \pm e^{ax/2} + \sqrt{b}cx + \sqrt{b} = 0, \end{aligned}$$

ce qui donne le résultat souhaité, en utilisant le Théorème 2. □

## 1.6. Plan, contributions et perspectives

La Partie I se concentre sur l'équation (1.16), permettant de comprendre l'effet de l'anisotropie mutationnelle sur l'évolution. Elle permet ainsi de modéliser et prédire l'évolution d'une population asexuée (*e.g.*, virus, bactéries) dans un milieu unique, dont les paramètres abiotiques peuvent changer au cours du temps, à cause d'une modification de la concentration d'antibiotique par exemple. La Partie II se concentre sur l'effet de la migration sur l'adaptation, et donc sur l'invasion d'un nouveau milieu par des pathogènes. Comme dit dans les précédentes sections, nous nous concentrerons sur le cas de deux environnements connectés. La dernière partie (Partie III) permet de faire un premier pas vers des modèles de migration dans un environnement continu. Elle se concentrera sur l'invasion d'individus dont le trait évolue, et l'effet de la reproduction (asexuée et sexuée) sur celle-ci.

### Partie I

#### Chapitre 2

Le Chapitre 2 (HAMEL ; LAVIGNE ; MARTIN ; ROQUES 2020) étudie la dynamique adaptative d'une population asexuée, sous l'effet conjoint de la sélection et des mutations, supposées anisotrope. De récentes recherches se sont penchées sur des modèles d'évolution pour suivre la trajectoire complète au cours du temps de la distribution de fitness. La clé de ces études est l'utilisation d'équations aux dérivées partielles ou d'équations intégro-différentielles, représentant la dynamique de la

distribution d'un unique trait phénotypique $x \in \mathbb{R}$ :

$$\partial_t q(t,x) = \mathcal{M}_1[t,x,q(t,x)] + (m(x) - \overline{m}(t))\, q(t,x),$$

où l'opérateur (différentiel ou intégral, voir Section 1.3.2) $\mathcal{M}_1$ représente l'effet des mutations sur la distribution phénotypique $q$ au cours du temps. Certains modèles (ALFARO ; CARLES 2014a ; GIL ; HAMEL ; MARTIN ; ROQUES 2017 ; TSIMRING ; LEVINE ; KESSLER 1996) relient directement ce trait à la fitness elle-même, entraînant alors l'équation suivante, satisfaite par la distribution de fitness :

$$\partial_t p(t,m) = \mathcal{M}_2[t,m,p(t,m)] + (m - \overline{m}(t))\, p(t,m),$$

avec $m \in \mathbb{R}$ la fitness, tandis que $\mathcal{M}_2$ est un opérateur intégral ou différentiel. Nous utiliserons le modèle phénotypique général, en utilisant une relation explicite entre le trait phénotypique et la fitness $m$, appelée modèle de Fisher (*cf.* Section 1.4 pour plus de détails). Ce modèle implique, dans le cas général de $n$ traits phénotypiques, regroupés dans un seul et même vecteur $\mathbf{x} \in \mathbb{R}^n$, la relation :

$$m(\mathbf{x}) = -\frac{\|\mathbf{x} - \mathbf{O}\|^2}{2},$$

pour $\mathbf{O} \in \mathbb{R}^n$ un optimum phénotypique. Par une translation de l'espace phénotypique, nous pouvons supposer que $\mathbf{O} = (0,\dots,0) \in \mathbb{R}^n$. Finalement, le modèle considéré dans le Chapitre 2 peut se réécrire sous la forme complète :

$$\partial_t q(t,\mathbf{x}) = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\, \partial_{x_i x_i} q(t,\mathbf{x}) + (m(x) - \overline{m}(t))\, q(t,\mathbf{x}),$$

pour tout $t > 0$ et $\mathbf{x} \in \mathbb{R}^n$, avec $\overline{m}(t)$ la fitness moyenne au temps $t$ de la population, définie par :

$$\overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x}) q(t,\mathbf{x}) \mathrm{d}\mathbf{x}.$$

Cette équation a été récemment traitée dans le cas $n = 1$ (ALFARO ; CARLES 2017 ; ALFARO ; VERUETE 2018, pour une forme plus générale de fitness). Le but de ce chapitre est de traiter le cas général $n \geq 1$, et ainsi de comprendre le rôle de l'anisotropie.

Le problème de Cauchy associé admet certes une unique solution, comme le mentionne le théorème suivant, mais il semble difficile d'obtenir une formule explicite pour la distribution phénotypique $q(t,\mathbf{x})$, avec des mutations anisotropes, contrairement au cas isotrope.

**Théorème 2.**
*Soit $q_0 \in C^{2+\alpha}(\mathbb{R}^n)$, pour $\alpha \in (0,1)$ une mesure de probabilité, i.e., telle que :*

$$\int_{\mathbb{R}^n} q_0(t,\mathbf{x})\mathrm{d}\mathbf{x} = 1.$$

*Supposons qu'il existe une fonction $g : \mathbb{R}_+ \to \mathbb{R}_+$ décroissante telle que :*

- $0 \leq q_0(\mathbf{x}) \leq g(\|\mathbf{x}\|)$*, pour tout $\mathbf{x} \in \mathbb{R}^n$ (avec $\|\cdot\|$ la norme euclidienne de $\mathbb{R}^n$) ;*
- *la fonction $\mathbf{x} \mapsto m(\mathbf{x})g(\|\mathbf{x}\|)$ soit bornée et intégrable sur $\mathbb{R}^n$.*

*Il existe une unique solution positive $q \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ du problème :*

$$\begin{cases} \partial_t q(t,\mathbf{x}) &= \displaystyle\sum_{i=1}^n \frac{\mu_i^2}{2}\partial_{x_i x_i} q(t,\mathbf{x}) + (m(\mathbf{x}) - \overline{m}(t))q(t,\mathbf{x}), \quad t \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \\ q(0,\mathbf{x}) &= q_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n, \end{cases}$$

*telle que $q \in L^\infty((0,T) \times \mathbb{R}^n)$ pour tout $T > 0$, et telle que la fonction :*

$$t \mapsto \overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})q(t,\mathbf{x})\mathrm{d}\mathbf{x},$$

*soit bien définie et continue sur $\mathbb{R}_+$. De plus, pour tout $t \geq 0$, la fonction $q(t,\cdot)$ vérifie :*

$$\int_{\mathbb{R}^n} q(t,\mathbf{x})\mathrm{d}\mathbf{x} = 1.$$

L'objectif étant d'obtenir des informations sur la distribution de fitness $p(t,m)$, nous allons plutôt utiliser cette distribution au lieu de $q$. Plus précisément, suite à l'anisotropie, nous allons utiliser le vecteur fitness $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{R}_-^n$ tel que le vecteur fitness associé à un phénotype $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ soit donné par :

$$\mathbf{m}(\mathbf{x}) = (m_1(\mathbf{x}), \ldots, m_n(\mathbf{x})), \qquad \text{avec} \qquad m_i(\mathbf{x}) = -\frac{x_i^2}{2}.$$

La distribution du vecteur fitness $\mathbf{p}(t,\mathbf{m})$ peut être reliée à la distribution $q(t,\mathbf{x})$ explicitement :

**Proposition 4.** *Pour tout $t \geq 0$, et tout vecteur fitness $\mathbf{m} = (m_1, \ldots, m_n) \in (\mathbb{R}_-^*)^n$, nous avons :*

$$\mathbf{p}(t, \mathbf{m}) = \frac{2^{-n/2}}{\sqrt{|m_1 \cdots m_n|}} \sum_{\varepsilon=(\varepsilon_1,\ldots,\varepsilon_n)\in\{\pm 1\}^n} q(t, \mathbf{x}^\varepsilon(\mathbf{m})),$$

*avec $\mathbf{x}^\varepsilon(\mathbf{m}) = (\varepsilon_1\sqrt{-2m_1}, \ldots, \varepsilon_n\sqrt{-2m_n}) \in \mathbb{R}^n$. De plus, en tout temps $t$, la distribution $\mathbf{p}$ vérifie :*

$$\int_{\mathbb{R}_-^n} \mathbf{p}(t, \mathbf{m}) \mathrm{d}\mathbf{m} = 1, \quad \text{et} \quad \overline{m}(t) = \sum_{i=1}^n \int_{\mathbb{R}_-^n} m_i \mathbf{p}(t, \mathbf{m}) \mathrm{d}\mathbf{m}.$$

*Les intégrales écrites ci-dessus sont convergentes.*

Cette relation permet d'obtenir une équation différentielle parabolique dégénérée satisfaite par la distribution $p(t, m)$.

**Théorème 3.**
*La distribution $\mathbf{p}$ du vecteur fitness est une solution de classe $C^{1,2}(\mathbb{R}_+ \times (\mathbb{R}_-^*)^n)$ de :*

$$\partial_t \mathbf{p}(t, \mathbf{m}) = \sum_{i=1}^n \mu_i^2 |m_i| \, \partial_{m_i m_i} \mathbf{p}(t, \mathbf{m}) - \frac{3}{2} \sum_{i=1}^n \mu_i^2 \, \partial_{m_i} \mathbf{p}(t, \mathbf{m}) + \left( \sum_{i=1}^n m_i - \overline{m}(t) \right) \mathbf{p}(t, \mathbf{m}),$$

*pour $t \geq 0$ et $\mathbf{m} \in (\mathbb{R}_-^*)^n$.*

Cependant peu de résultats analytiques semblent pouvoir découler de cette dernière. L'intérêt de ce changement de variable se trouve sur la modification du terme quadratique $m(\mathbf{x})$ en un terme linéaire $m$. Cela permet alors d'obtenir une équation différentielle linéaire de transport vérifiée par la fonction génératrice des cumulants $C_\mathbf{p}$ de $\mathbf{p}$. Remarquons de plus que la fitness moyenne peut se calculer à l'aide de $C_\mathbf{p}$ : $\overline{m}(t) = \sum_{i=1}^n \partial_{z_i} C_\mathbf{p}(t, \mathbf{O})$.

**Théorème 4.**
*La fonction génératrice des cumulants $C_{\mathbf{p}}$ de $\mathbf{p}$ est une solution de classe $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^n)$ :*

$$\begin{cases} \partial_t C_{\mathbf{p}}(t,\mathbf{z}) &= A(\mathbf{z}) \cdot \nabla C_{\mathbf{p}}(t,\mathbf{z}) - b(\mathbf{z}) - \overline{m}(t),\ t \geq 0,\ \mathbf{z} \in \mathbb{R}_+^n, \\ C_{\mathbf{p}}(0,\mathbf{z}) &= C_{\mathbf{p}_0}(\mathbf{z}),\ \mathbf{z} \in \mathbb{R}_+^n, \end{cases}$$

*où :*

$$A(\mathbf{z}) = (1-\mu_1^2 z_1^2, \ldots, 1-\mu_n^2 z_n^2) \in \mathbb{R}^n,\ \ b(\mathbf{z}) = \sum_{i=1}^n \frac{\mu_i^2}{2} z_i \in \mathbb{R},\ \ \overline{m}(t) = \sum_{i=1}^n \partial_i C_{\mathbf{p}}(t,\mathbf{O}),$$

*et $\nabla C_{\mathbf{p}}(t,\mathbf{z})$ représentent le gradient de $C_{\mathbf{p}}$ en fonction de la variable $\mathbf{z}$.*

Cette équation de transport admet une solution explicite en tout temps :

**Proposition 5.** *La fonction génératrice des cumulants $C_{\mathbf{p}}(t,\mathbf{z})$ de $\mathbf{p}$, évaluée au temps $t \geq 0$ et au point $\mathbf{z} = (z_1, \ldots, z_n) \in [0, 1/\mu_1) \times \cdots \times [0, 1/\mu_n)$, est donnée par :*

$$C_{\mathbf{p}}(t,\mathbf{z}) = \frac{1}{2}\sum_{i=1}^n \log\left[\frac{\cosh(\mu_i t)\,\cosh(\operatorname{atanh}(\mu_i z_i))}{\cosh(\mu_i t + \operatorname{atanh}(\mu_i z_i))}\right] + C_{\mathbf{p}_0}(\psi(t,\mathbf{z})) - C_{\mathbf{p}_0}(\psi(t,\mathbf{O})),$$

*avec :*

$$\psi(t,\mathbf{z}) = (\psi_1(t,\mathbf{z}), \ldots, \psi_n(t,\mathbf{z})), \quad \text{et} \quad \psi_j(t,\mathbf{z}) = \frac{1}{\mu_j}\tanh\left(\mu_j t + \operatorname{atanh}(\mu_j z_j)\right).$$

*De plus, la fitness moyenne au temps $t \geq 0$ vaut :*

$$\overline{m}(t) = \sum_{i=1}^n \left[(1-\tanh^2(\mu_i t))\,\partial_i C_{\mathbf{p}_0}(\psi(t,\mathbf{O})) - \frac{\mu_i}{2}\tanh(\mu_i t)\right].$$

Étant donné que nous avons obtenu une formule explicite pour la fitness moyenne $\overline{m}(t)$ en tout temps $t$, nous pouvons aussi étudier la limite en temps grand de celle-ci, ou encore de $C_{\mathbf{p}}$. Par ailleurs, $C_{\mathbf{p}}$ est relié à la distribution $\mathbf{p}$ par l'opérateur de transformation de Laplace. Cela implique aussi des résultats sur le comportement de la distribution $\mathbf{p}$ en temps long.

**Théorème 5.**
*Considérons la distribution* $\mathbf{p}$ *et la fitness moyenne* $\overline{m}(t)$*, définie précédemment. Alors :*

$(i)$ $\mathbf{p}(t,\cdot)$ *converge faiblement dans* $(\mathbb{R}_-^*)^n$ *vers* $\mathbf{p}_\infty$*, quand* $t \to +\infty$*, avec :*

$$\mathbf{p}_\infty(\mathbf{m}) = \frac{1}{\pi^{n/2}\sqrt{\mu_1\cdots\mu_n}\sqrt{|m_1\cdots m_n|}} \exp\left(\sum_{i=1}^n \frac{m_i}{\mu_i}\right) \quad \textit{pour tout } m \in (\mathbb{R}_-^*)^n,$$

*i.e.,* $\displaystyle\int_{(\mathbb{R}_-^*)^n} \mathbf{p}(t,\mathbf{m})\,\phi(\mathbf{m})\,\mathrm{d}\mathbf{m} \to \int_{(\mathbb{R}_-^*)^n} \mathbf{p}_\infty(\mathbf{m})\,\phi(\mathbf{m})\,\mathrm{d}\mathbf{m}$ *quand* $t \to +\infty$ *pour toute fonction test* $\phi \in C_c^\infty((\mathbb{R}_-^*)^n)$*;*

$(ii)$ $\overline{m}(t) \to \overline{m}_\infty = -\displaystyle\sum_{i=1}^n \frac{\mu_i}{2}$ *quand* $t \to +\infty$ *et* $\overline{m}_\infty = \displaystyle\sum_{i=1}^n \int_{\mathbb{R}_-^n} m_i \mathbf{p}_\infty(\mathbf{m})\mathrm{d}\mathbf{m}$*;*

$(iii)$ *la fonction* $\mathbf{p}_\infty$ *est une solution de classe* $C^2((\mathbb{R}_-^*)^n)$ *de :*

$$0 = \sum_{i=1}^n \mu_i^2 |m_i| \partial_{ii}\mathbf{p}_\infty(\mathbf{m}) - \frac{3}{2}\sum_{i=1}^n \mu_i^2 \partial_i \mathbf{p}_\infty(\mathbf{m}) + \left(\sum_{i=1}^n m_i - \overline{m}_\infty\right)\mathbf{p}_\infty(\mathbf{m}).$$

De plus, cette formule explicite montre l'intérêt de prendre en considération l'anistropie. En effet, contrairement au cas isotrope, la fitness moyenne $\overline{m}(t)$ peut présenter des plateaux évolutifs, dans le sens suivant :

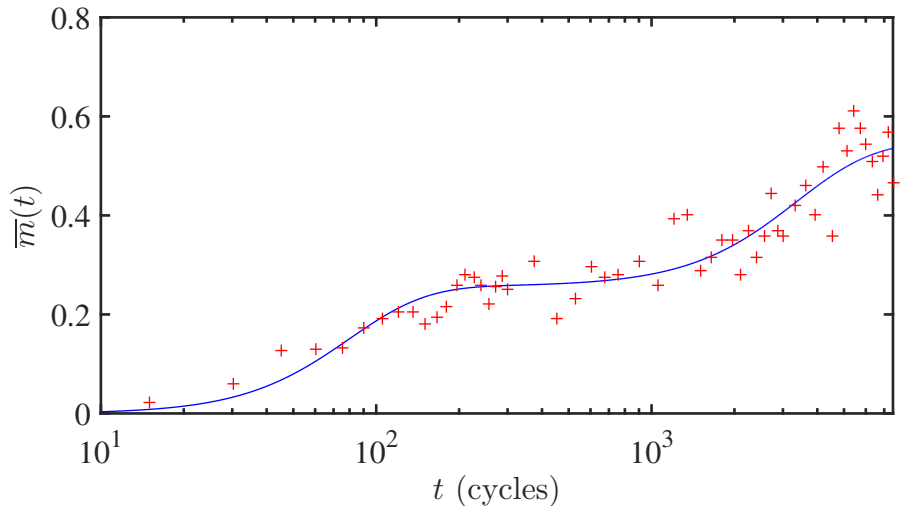
**Proposition 6.** *Supposons que les résultats restent encore vrais lorsque la distribution phénotypique initiale est de la forme* $\delta_{\mathbf{x}_0}$*, pour* $\mathbf{x}_0 = (x_{0,1},\dots,x_{0,n}) \in \mathbb{R}^n$*. Soit* $\overline{m}(t)$ *définie précédemment et pour* $1 \le k \le n-1$*,* $\overline{m}_{\infty,k}$ *par :*

$$\overline{m}_{\infty,k} := -\sum_{i=k+1}^n \frac{x_{0,i}^2}{2} - \sum_{i=1}^k \frac{\mu_i}{2}.$$

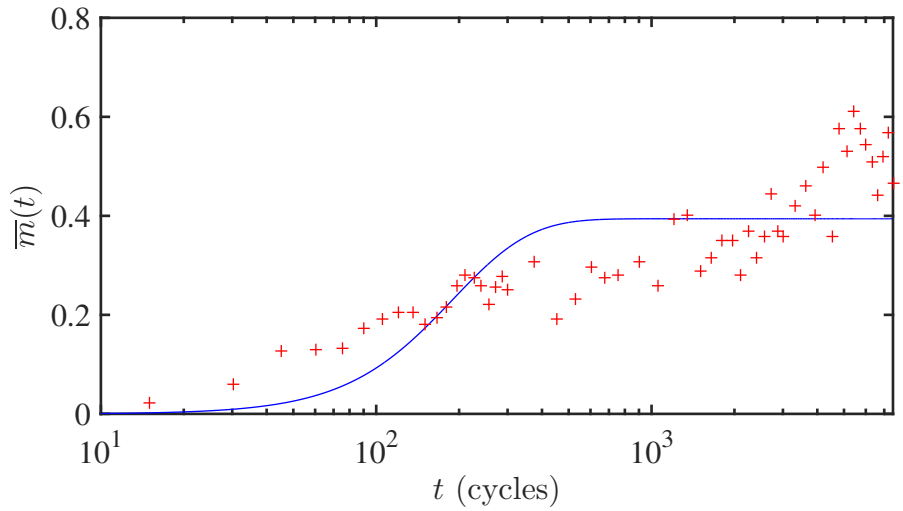
*Pour* $T > 0$*,* $\varepsilon > 0$ *et* $\mu_1 > 0$*, il existe* $\mu_2,\dots,\mu_n > 0$ *tels que, pour tout* $k \in [\![1, n-1]\!]$*, l'ensemble* $\left\{t \ge 0,\ |\overline{m}(t) - \overline{m}_{\infty,k}| \le \varepsilon\right\}$ *contienne un intervalle de longueur au moins égale à* $T$*.*

Ce comportement est important pour la compréhension de certaines données dues à des expériences, comme celle menée par R. Lenski (LTEE) sur *Escherichia coli* (Lenski 2017 ; Lenski ; Rose ; Simpson ; Tadler 1991 ; Lenski ; Travisano 1994). En effet, sur les 10 000 premières générations, les mesures de la fitness moyenne semblaient converger vers une valeur asymtotique. Cependant, celle-ci connaît depuis peu une nouvelle augmentation. Cette allure de courbe ne peut correspondre au cas évolutif de mutations isotropes, mais peut être expliquée par

des mutations anisotropes (*cf.* Fig. 1.15).




(a) Modèle anisotrope




(b) Modèle isotrope

FIGURE 1.15. – **Trajectoire adaptative donnée par les modèles isotrope et anisotrope *versus* les données de l'expérience LTEE.** La fonction $f(t) = \overline{m}(t+2) - \overline{m}(2)$, avec $\overline{m}(t)$ donnée par la Proposition 6. La fonction $f(t)$ avec les paramètres correspondant à la meilleure approximation est représentée en bleu. Les croix rouges correspondent aux données de LTEE (population Ara-1, $\ln[w(6.64t)]$, avec $w$ la donnée d'origine). Les valeurs donnant la meilleure approximation valent (a, modèle anistrope) $\mathbf{x}_0 = (0.73, 0.76)$, $\mu_1 = 1.3 \cdot 10^{-2}$ et $\mu_2 = 3.0 \cdot 10^{-4}$ ; (b, modèle isotrope) $\mathbf{x}_0 = (0.44, 0.78)$ et $\mu_1 = \mu_2 = 5.3 \cdot 10^{-3}$.

## Chapitre 3

Suite à des changements environnementaux, comme des modifications de la concentration d'antibiotiques ou de vitamines, l'adaptation de certaines populations peut être facilitée ou bien compromise.

Différentes recherches sur la dynamique de la distribution phénotypique ont déjà été menées dans le cas d'environnements invariants (DESAI ; FISHER 2011 ; GERRISH ; COLATO ; PERELSON ; SNIEGOWSKI 2007 ; GIL ; HAMEL ; MARTIN ; ROQUES 2017 ; GIL ; HAMEL ; MARTIN ; ROQUES 2019 ; HAMEL ; LAVIGNE ; MARTIN ; ROQUES 2020 ; SNIEGOWSKI ; GERRISH 2010 ; TSIMRING ; LEVINE ; KESSLER 1996). Le chapitre précédent nous a permis de mettre en lumière l'importance de l'anisotropie dans l'évolution d'une population (apparition potentielle de plateaux adaptatifs).

Ces différentes recherches peuvent aussi s'appliquer à des problèmes de changement environnementaux brutaux. Cependant, lorsque les paramètres abiotiques varient de façon continue, l'optimum phénotypique $\mathbf{O}$ donné par le modèle de Fisher se déplace au cours du temps. Ainsi la fitness $m(\mathbf{x})$ doit être modifiée :

$$m(t,\mathbf{x}) = -\frac{\|\mathbf{x}-\mathbf{O}(t)\|^2}{2},$$

où $\mathbf{O}(t) \in C(\mathbb{R}_+,\mathbb{R}^n)$ représente le phénotype optimal au temps $t$. Par soucis de clareté, nous ferons l'hypothèse que cet optimum se déplace selon l'axe :

$$\{\mathbf{x}=(x_1,\dots,x_n)\in\mathbb{R}^n, x_2=\cdots=x_n=0\}.$$

Plus précisément, la fonction vectorielle $\mathbf{O}(t)$ s'écrit :

$$\forall t\geq 0,\ \mathbf{O}(t)=(\Delta(t),0,\dots,0)\in\mathbb{R}^n.$$

Le cas isotrope a été traité récemment par ROQUES ; PATOUT ; BONNEFON ; MARTIN 2020, donnant une description analytique de la trajectoire de la fitness moyenne $\overline{m}(t)$. Le but de ce chapitre est d'étendre ces résultats au cas anisotrope :

$$\partial_t q(t,\mathbf{x}) = \sum_{i=1}^{n}\frac{\mu_i^2}{2}\partial_{x_ix_i}q(t,\mathbf{x}) + (m(t,\mathbf{x})-\overline{m}(t))q(t,\mathbf{x}),\quad t>0,\ \mathbf{x}\in\mathbb{R}^n,$$

avec :

$$\overline{m}(t)=\int_{\mathbb{R}^n} m(t,\mathbf{x})q(t,\mathbf{x})\mathrm{d}\mathbf{x}.$$

Nous ferons des hypothèses similaires sur la distribution initiale $q_0$ à celles faites dans le Chapitre 2 :

- ◇ Pour $\alpha\in(0,1)$, $q_0\in C^{2+\alpha}(\mathbb{R}^n)$ et $\int_{\mathbb{R}^n} q_0(\mathbf{x})\mathrm{d}\mathbf{x}=1$ ;
- ◇ Il existe une fonction $g:\mathbb{R}_+\to\mathbb{R}_+$ décroissante avec $0\leq q_0\leq g(\|\cdot\|)$ dans $\mathbb{R}^n$, telle que pour tout $b>0$, la fonction $\mathbf{x}\mapsto\exp(b\|\mathbf{x}\|)g(\|\mathbf{x}\|)$ soit bornée et intégrable sur $\mathbb{R}^n$.

**Théorème 6. (Aronson ; Besala 1967 ; Chabrowski 1970)**
*Il existe une unique solution $q \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ au problème de Cauchy :*

$$\left\{ \begin{array}{rcl} \partial_t q(t,\mathbf{x}) & = & \displaystyle\sum_{i=1}^{n} \frac{\mu_i^2}{2} \partial_{x_i x_i} q(t,\mathbf{x}) + (m(t,\mathbf{x}) - \overline{m}(t))\, q(t,\mathbf{x}), \quad t \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \\ q(0,\mathbf{x}) & = & q_0(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n, \end{array} \right.$$

*telle que $q \in L^\infty((0,T) \times \mathbb{R}^n)$ pour tout $T > 0$, et la fonction :*

$$t \mapsto \overline{m}(t) = \int_{\mathbb{R}^n} m(t,\mathbf{x})\, q(t,\mathbf{x})\, \mathrm{d}\mathbf{x},$$

*soit bien définie et continue sur $\mathbb{R}_+$. De plus, en tout temps $t \geq 0$, nous avons :*

$$\int_{\mathbb{R}^n} q(t,\mathbf{x})\, \mathrm{d}\mathbf{x} = 1.$$

Contrairement au Chapitre 2 (Hamel ; Lavigne ; Martin ; Roques 2020) qui pouvait utiliser le vecteur fitness $\mathbf{m}(\mathbf{x}) = (-x_1^2/2, \ldots, -x_n^2/2) \in \mathbb{R}_-^n$, nous avons besoin d'un autre vecteur $\mathbf{v} = (v_0, v_1, \ldots, v_n) \in \mathbb{R}^{n+1}$ avec :

$$v_0(\mathbf{x}) = x_1 \in \mathbb{R} \quad \text{et} \quad \forall 1 \leq i \leq n,\ v_i(\mathbf{x}) = -\frac{x_i^2}{2},$$

pour tout $\mathbf{x} \in \mathbb{R}^n$. Sa distribution jointe $\mathbf{p}(t,\mathbf{v})$ est la mesure image de $q(t,\mathbf{x})\mathrm{d}\mathbf{x}$ par l'application $\mathbf{x} \mapsto (x_1, -x_1^2/2, \ldots, -x_n^2/2)$ :

**Proposition 7.** *On définit la sous-variété :*

$$\Gamma_{n+1} = \{(v_0, \ldots, v_n) \in \mathbb{R} \times \mathbb{R}_-^n,\ v_1 = -v_0^2/2\}.$$

*Pour tout $t \geq 0$ et $\mathbf{v} = (v_0, \ldots, v_n) \in \Gamma_{n+1}$, nous avons :*

$$\mathbf{p}(t,\mathbf{v}) = \frac{2^{-(n-1)/2}}{\sqrt{|v_2 \cdots v_n|}} \sum_{\varepsilon \in \{\pm 1\}^{n-1}} q(t, \mathbf{x}^\varepsilon(t,\mathbf{v})),$$

*où le changement de variable $\mathbf{x}^\varepsilon(\mathbf{v})$ est défini par $\mathbf{x}^\varepsilon(\mathbf{v}) = (v_0, x_2^\varepsilon(\mathbf{v}), \ldots, x_n^\varepsilon(\mathbf{v}))$, avec $x_i^\varepsilon(\mathbf{v}) = \varepsilon_i \sqrt{-2v_i}$ et $\varepsilon = (\varepsilon_2, \ldots, \varepsilon_n) \in \{\pm 1\}^{n-1}$. De plus, nous avons :*

$$\forall\, t \geq 0,\ \int_{\Gamma_{n+1}} \mathbf{p}(t,\mathbf{v})\, \mathrm{d}\mathbf{v} = 1 \ \text{ et } \ \overline{m}(t) = -\frac{\Delta(t)^2}{2} + \int_{\Gamma_{n+1}} \boldsymbol{\alpha}(t) \cdot \mathbf{v}\, \mathbf{p}(t,\mathbf{v}) \mathrm{d}\mathbf{v},$$

*avec $\boldsymbol{\alpha}(t) = (\Delta(t), 1, \ldots, 1) \in \mathbb{R}^{n+1}$. Les intégrales ci-dessus sont convergentes.*

Comme dans le cas d'un optimum fixe, cette distribution vérifie une équation parabolique dégénérée (sur une sous-variété), tandis que sa fonction génératrice des cumulants satisfait une équation de transport à coefficients dépendant du temps :

**Theorem 8.**
*La fonction génératrice des cumulants $C_{\mathbf{p}}$ de $\mathbf{p}$ est bien définie et est une solution de classe $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^{n+1})$ du système :*

$$\begin{cases} \partial_t C_{\mathbf{p}}(t,\mathbf{z}) &= \boldsymbol{\alpha}(t)\cdot\nabla C_{\mathbf{p}}(t,\mathbf{z}) - \overline{v}(t) + V(\mathbf{z})\cdot\nabla C_{\mathbf{p}}(t,\mathbf{z}) + \gamma(\mathbf{z}),\ t\geq 0,\ \mathbf{z}\in\mathbb{R}_+^n, \\ C_{\mathbf{p}}(0,\mathbf{z}) &= C_{\mathbf{p}_0}(\mathbf{z}),\ \mathbf{z}\in\mathbb{R}_+^n, \end{cases}$$

*où :*

$$\boldsymbol{\alpha}(t) = (\Delta(t),1,\dots,1),\ V(\mathbf{z}) = \left(-\mu_1^2 z_0 z_1, -\mu_1^2 z_1^2, \dots, -\mu_n^2 z_n^2\right),$$
$$\gamma(\mathbf{z}) = \frac{\mu_1^2}{2} z_0^2 - \sum_{i=1}^n \frac{\mu_i^2}{2} z_i, \text{ et } \overline{v}(t) = \boldsymbol{\alpha}(t)\cdot\nabla C_{\mathbf{p}}(t,0),$$

*et $\nabla C_{\mathbf{p}}(t,\mathbf{z})$ représente le gradient de $C_{\mathbf{p}}$ en fonction de la variable $\mathbf{z}$.*

Celle-ci peut se résoudre grâce à la méthode des caractéristiques présentée en Section 1.5.2.1 :

**Proposition 9.** *En tout temps $t\geq 0$, la fitness moyenne $\overline{m}(t)$ est donnée par :*

$$\begin{aligned} \overline{m}(t) = &\left[\frac{\Delta(t)}{\cosh[\mu_1 t]} - \mu_1\int_0^t \Delta(s)\frac{\sinh[\mu_1 s]}{\cosh^2[\mu_1 t]}\mathrm{d}s\right]\partial_{z_0} C_{\mathbf{p}}(0,\boldsymbol{\psi}(t)) \\ &- \frac{1}{2}\left[\Delta(t) - \mu_1\int_0^t \Delta(s)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s\right]^2 \\ &+ \sum_{i=1}^n \left[\left[1-\tanh^2(\mu_i t)\right]\partial_{z_i} C_{\mathbf{p}}\left(0,\boldsymbol{\psi}(t)\right) - \frac{\mu_i}{2}\tanh(\mu_i t)\right], \end{aligned}$$

*avec :*

$$\boldsymbol{\psi}(t) = \left(\int_0^t \Delta(s)\frac{\cosh[\mu_1(t-s)]}{\cosh[\mu_1 t]}\mathrm{d}s, \frac{\tanh(\mu_1 t)}{\mu_1}, \dots, \frac{\tanh(\mu_n t)}{\mu_n}\right) \in \mathbb{R}^{n+1}.$$

En particulier, si nous connaissons explicitement le mouvement de l'optimum, la dynamique de la fitness moyenne $\overline{m}(t)$ est elle-aussi connue :

◇ Si l'optimum se déplace à vitesse constante, *i.e.*, $\Delta(t) = ct$, pour $c\in\mathbb{R}$,

alors :

$$\overline{m}(t) = \frac{c\tanh(\mu_1 t)}{\mu_1 \cosh(\mu_1 t)}\,\partial_{z_0} C_{\mathbf{p}}(0, \boldsymbol{\psi}(t)) - \frac{c^2}{2\mu_1^2}\,\tanh^2[\mu_1 t] + \sum_{i=1}^{n}\left[[1-\tanh^2(\mu_i t)]\,\partial_{z_i} C_{\mathbf{p}}\,(0, \boldsymbol{\psi}(t)) - \frac{\mu_i}{2}\tanh(\mu_i t)\right],$$

avec :

$$\boldsymbol{\psi}(t) = \left(\frac{c}{\mu_1^2}\left(1 - \frac{1}{\cosh[\mu_1 t]}\right), \frac{\tanh(\mu_1 t)}{\mu_1}, \ldots, \frac{\tanh(\mu_n t)}{\mu_n}\right).$$

converge vers :

$$\overline{m}(\infty) = -\sum_{i=1}^{n}\frac{\mu_i}{2} - \frac{c^2}{2\mu_1^2},$$

quand $t$ tend vers $+\infty$ ;

- ◇ Si l'optimum converge vers un point $\mathbf{O}(\infty)$, dans ce cas la fitness moyenne converge vers :

$$\overline{m}(\infty) = -\sum_{i=1}^{n}\frac{\mu_i}{2},$$

valeur correspondant à la limite de $\overline{m}(t)$ dans le cas d'un optimum fixe ;

- ◇ Si $\Delta$ est une fonction concave croissante, la fitness moyenne converge vers :

$$\overline{m}(\infty) = -\sum_{i=1}^{n}\frac{\mu_i}{2}\,;$$

- ◇ Si $\Delta$ est une fonction convexe croissante, la fitness moyenne diverge vers $-\infty$. La population ne s'adapte pas, et finit par s'éteindre.

# Partie II

## Chapitre 4

Plus tôt dans ce chapitre, nous avons vu les différentes notions d'environnements sources et environnements puits. Dans la littérature, l'un des systèmes migratoires entre source et puits est le système à effet trou noir : un environnement source envoie continuellement des individus dans l'environnement puits, sans modification de la structure interne (démographique et phénotype) de la source (GOMULKIEWICZ ; HOLT ; BARFIELD 1999). Ce modèle est l'un des premiers permettant de comprendre des phénomènes d'invasion (COLAUTTI ; ALEXANDER ; DLUGOSCH ; KELLER et al. 2017) ou de changement d'hôtes (JANSEN ; COORS ; STOKS ; DE MEESTER 2011 ; SOKURENKO ; GOMULKIEWICZ ; DYKHUIZEN 2006). En particulier, nous pouvons étudier dans de telles situations la réussite (ou l'échec)

d'une invasion et connaître le temps qu'il faut à la population pour s'établir dans ce nouvel environnement.

Plusieurs modèles d'intéraction source-puits (DRURY ; DRAKE ; LODGE ; DWYER 2007 ; GARNIER ; ROQUES ; HAMEL 2012) se sont intéressés à la dynamique démographique sans considérer l'évolution de la population. Cependant, celle-ci peut altérer la finalité de l'adaptation, par le biais de la sélection, des mutations, de la dérive génétique, *etc.* Elle peut soit favoriser l'adaptation locale, soit l'empêcher (DÉBARRE ; RONCE ; GANDON 2013 ; GOMULKIEWICZ ; HOLT ; BARFIELD ; NUISMER 2010 ; LENORMAND 2002).

Le Chapitre 4 (LAVIGNE ; MARTIN ; ANCIAUX ; PAPAIX et al. 2020) permet d'étudier la dynamique d'un tel système source-puits à effet trou noir, avec un taux de migration constant. De plus, le puits sera supposé initiallement vide. Dans ce cas, la taille de population dans le puits vérifie l'équation :

$$\begin{cases} N'(t) = \overline{r}(t)N(t) + d,\ t > 0 \\ N(0) = 0, \end{cases}$$

où $d > 0$ correspond au nombre d'individus partant de la source et arrivant dans le puits, par unité de temps. En absence d'adaptation, *i.e.*, dans le cas où $\overline{r}$ est constant, il existe un équilibre démographique donné par $N = d/(-\overline{r})$, possible seulement dans le cas où $\overline{r} < 0$. Dans le cas où les individus évoluent, nous devons faire certaines hypothèses sur l'évolution de $\overline{r}(t)$ dans le puits.

Nous ferons l'hypothèse que la source est à l'état d'équilibre mutation-sélection donné par le modèle de Wright-Fisher. Dans ce cas, les fitness de la population dans la source sont distribuées selon une loi Gamma :

$$m_{source} \sim -\Gamma(n/2, \mu).$$

Cependant, les optima phénotypiques des deux environnements ne sont pas les mêmes. Nous faisons l'hypothèse que celui du puits se trouve à l'origine du repère phénotypique, tandis que celui de la source est en $\mathbf{x}^* \in \mathbb{R}^n$. La différence entre ces deux habitats est quantifiée par :

$$m_D = -\|\mathbf{x}^*\|^2/2.$$

Nous remarquons ici que la fitness d'un individu (en particulier l'optimum de la source) change une fois qu'il a migré. Cela modifie la distribution de la distribution des fitness des migrants, une fois qu'ils sont arrivés dans le puits, en :

$$p_{migr}(m) = \begin{cases} \frac{1}{\mu} \left(\frac{|m|}{m_D}\right)^{\frac{1}{2}\left(\frac{n}{2}-1\right)} e^{\frac{m-m_D}{\mu}} I_{\frac{n}{2}-1}\left[\frac{2\sqrt{m_D|m|}}{\mu}\right], \text{ si } m < 0, \\ 0, \text{ si } m \geq 0, \end{cases}$$

où $I_\nu$ est la fonction de Bessel modifiée du premier ordre.

Ainsi, nous modéliserons l'évolution de la population grâce à l'EDP de transport vérifiée par la fonction génératrice des cumulants de la fitness :

$$\partial_t C(t,z) = \left[1-\mu^2 z^2\right]\partial_z C(t,z) - \partial_z C(t,0) - \frac{\mu^2 n}{2} z + \frac{d}{N(t)}\left(e^{\Phi(z)-C(t,z)} - 1\right),$$

où $\Phi$ est la fonction génératrice des cumulants de la distribution des fitness des migrants dans le puits.

Cette équation admet une unique solution régulière, ce qui permet d'obtenir une formule pour la fitness moyenne dans le puits en tout temps :

$$\overline{r}(t) = \frac{f(t)-1}{\int_0^t f(\tau)\mathrm{d}\tau}, \quad \text{avec } f(t) = \exp\left[\left(r_{\max} - \mu\frac{n}{2}\right)t + \frac{m_D}{2\mu}(e^{-2\mu t}-1)\right].$$

Remarquons que cette formule ne fait pas intervenir le taux de migration $d$ : cela est dû au choix du modèle démographique Malthusien. Cela ne sera plus vrai si par exemple, nous prenons un modèle démographique logistique. Cette indépendence sera vérifiée grâce à des simulations IBM, dans une gamme de fort taux de migrations $d$.

Cette formule permet de capter quatre phases évolutives plus ou moins marquées selon les paramètres. Celles-ci se traduisent principalement au niveau de la distribution phénotypique et de la dynamique démographique :

1. Augmentation de la taille de population pour se stabiliser vers une certaine valeur ;
2. Équilibre migration-sélection, avec un faible impact des mutations et apparition d'un second mode autour de l'optimum du puits (généré à partir de l'effet combiné des rares migrants adaptés et des mutations *de novo* dans le puits) ;
3. Croissance rapide de la distribution phénotypique autour de l'optimum du puits et écrasement du mode autour de l'optimum de la source ;
4. Stabilisation d'une distribution gaussienne autour de l'optimum du puits (équilibre mutations-sélection).

Enfin cette formule permet d'avoir le comportement asymptotique d'un tel système évolutif. En effet, il existe deux scenarii possibles pour la population, selon les paramètres :

- ◇ si $r_{\max} - \mu n/2 \geq 0$, alors $\overline{r}(\infty) = r_{\max} - \mu n/2$ et $N(t) \to +\infty$, quand $t \to +\infty$. Le puits devient alors une source et l'invasion a lieu ;
- ◇ si $r_{\max} - \mu n/2 < 0$, alors $\overline{r}(\infty) = r_{\max} - \mu n/2 - \delta(m_D)$ ($\delta(m_D)$ étant une fonction vérifiant $m_D > \delta(m_D) > m_D/8$ quand $\mu$ est suffisamment grand) et $N$ converge. L'invasion échoue et la population n'est maintenue que par la migration. Il s'agit en fait d'un cas de mutagénèse létale.

Dans le cas où la population s'adapte, le temps d'établissement de la population dans le puits est alors donné par :

$$t_0 = \frac{1}{2\mu}\left[c + W_0\left(-c\,e^{-c}\right)\right], \qquad \text{avec } c = \frac{m_D}{r_{\max} - \mu\, n/2},$$

et $W_0$ la branche principale de la fonction de Lambert. Comme pour la formule de $\overline{r}$, nous remarquons que ce temps $t_0$ ne dépend pas du taux de migration. Cette indépendance a été vérifiée par des simulations IBM, dans un régime de fort taux de migration (*cf.* Fig. 1.16 a).

Étudions l'effet du coefficient $m_D$. Grâce à des approximations de la fonction de Lambert, le temps $t_0$ peut être approximé par une fonction linéaire en $m_D$ :

$$t_0 \approx \begin{cases} \frac{c}{2\mu}, & \text{pour } m_D \text{ faible,} \\ \frac{c}{\mu}, & \text{pour } m_D \text{ fort.} \end{cases}$$

Ainsi plus les individus de la source seront mal adaptés au puits, plus il faudra de temps pour qu'ils puissent s'adapter (cf. Fig. 1.16 c).

Il reste l'effet du paramètre mutationnel $\mu$, qui est plus difficile à décrire : quand $\mu$ croît, le temps $t_0$ décroît, jusqu'à atteindre un miminum global, pour ensuite croître. Là encore, la mutagénèse létale agît. En augmentant le taux de mutation, l'adaptation de la population est favorisée. Cependant, si le taux de mutation devient trop élevé, nous nous rapprochons du cas critique de mutagénèse : ainsi le temps $t_0$ augmente jusqu'à exploser en $\mu$ fini (*cf.* Fig 1.16 b).

Comme dit en discussion de HAMEL ; LAVIGNE ; MARTIN ; ROQUES 2020, cette étude permettrait d'étudier l'effet d'un troisième environnement intermédiaire, entre la source et le puits (*cf.* Fig 1.16 c).

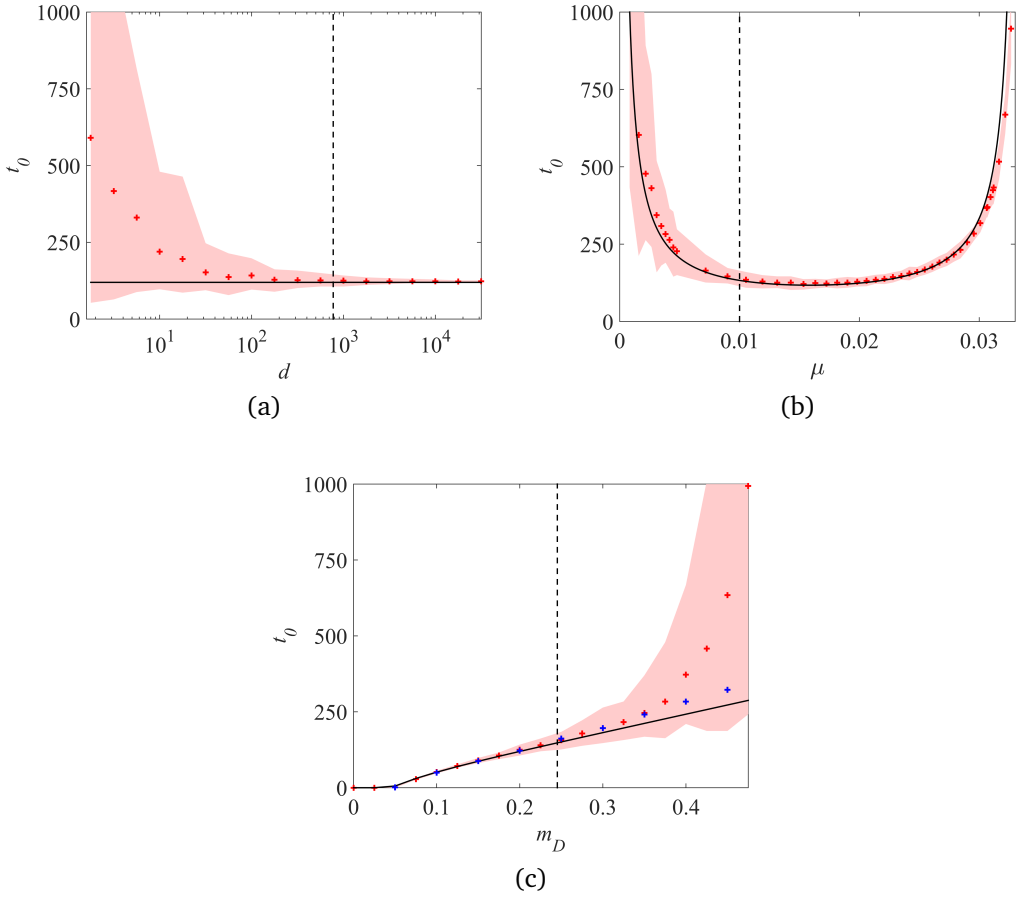



FIGURE 1.16. – **Dépendance du temps $t_0$ d'établissement de la population, par rapport au taux d'immigration $d$, du paramètre mutationnel $\mu$ et de la différence entre les habitats $m_D$.** Valeur théorique de $t_0$ (courbe noire) vs valeur obtenue par IBM (croix rouges) et l'intervalle de confiance à $95\%$. Les paramètres fixés sont (a) $m_D = 0.2$, $U = 0.1$ (b) $m_D = 0.2$, $d = 10^3$ et (c) $d = 10^3$, $U = 0.1$. La ligne verticale en pointillé correspond à la valeur de $d$, $\mu$ et $m_D$ telles que $-dU/\overline{r}(0) = 500$ et $U = U_c$. Les croix bleues (c) correspondent au temps d'établissement $t_0^I(m_D)$, obtenues par IBM, en présence d'un habitat intermédiaire avec un optimum phénotypique $\mathbf{x}^I$ tel que $\|\mathbf{x}^* - \mathbf{x}^I\|^2/2 = \|\mathbf{x}^I\|^2/2 = m_D/2$. Dans tous les cas, nous fixons $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$.

## Chapitre 5

Ayant traité le système source-puits à effet trou noir dans le Chapitre 4, nous nous penchons maintenant sur une migration symétrique.

Plus précisément, nous considérons deux environnements, chacun associé à un optimum phénotypique noté $\mathbf{O}_1$ ou $\mathbf{O}_2$. Nous pouvons faire l'hypothèse que ces deux optima se trouvent sur l'axe $\{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n, x_2 = \cdots = x_n = 0\}$, *i.e.*,

qu'il existe $\beta > 0$ avec :

$$\mathbf{O}_1 = (-\beta, 0, \dots, 0) \qquad \text{et} \qquad \mathbf{O}_2 = (\beta, 0, \dots, 0).$$

Le but de ce chapitre est de comprendre l'effet de la migration et de la différence entre les habitats, mesurée par $m_D = \|\mathbf{O}_1 - \mathbf{O}_2\|^2/2$, sur l'évolution démographique des habitats (persistence ou extinction).

Ce type de modèle a été étudié par MIRRAHIMI ; GANDON 2020, en utilisant une méthode basée sur des équations de type Hamilton-Jacobi, dans le cas $n = 1$. Les auteurs ont pu montrer l'existence de deux équilibres distincts pour la distribution phénotypique, dans le cadre de faibles mutations. Il existe une valeur critique pour le taux de migration, sous lequel les deux sous-populations restent distinctes, entrainant une densité bimodale des phénotypes pour la population globale : les individus se spécialisent. Au delà de cette valeur, la population se mélange complètement, pour créer une densité unimodale : les individus deviennent généralistes.

Nous étudierons un modèle déterministe basé sur l'approximation diffusive :

$$\forall t > 0,\ \forall \mathbf{x} \in \mathbb{R}^n, \begin{cases} \partial_t u_1(t,\mathbf{x}) = \dfrac{\mu^2}{2}\Delta u_1(t,\mathbf{x}) + f_1(t,\mathbf{x},u_1) + \delta[u_2(t,\mathbf{x}) - u_1(t,\mathbf{x})], \\ \partial_t u_2(t,\mathbf{x}) = \dfrac{\mu^2}{2}\Delta u_2(t,\mathbf{x}) + f_2(t,\mathbf{x},u_2) + \delta[u_1(t,\mathbf{x}) - u_2(t,\mathbf{x})], \end{cases} \tag{1.19}$$

avec $u_i$ la densité des phénotypes dans l'environnement $i \in \{1, 2\}$, $\delta > 0$ le taux de migration et $\mu > 0$ un paramètre mutationnel. Remarquons que nous faisons aussi l'hypothèse d'égalité des paramètres mutationnels.

Contrairement aux autres chapitres, nous nous intéresserons ici à des questions de persistance et d'extinction, et nous étudirons donc des densités de populations (notées $u$ tout au long de ce manuscrit) plutôt que des distributions de phénotypes (notées $q$). De plus, nous étudierons deux types de fonctions de croissance $f_i$ :

◇ Premier type – Cas Malthusien : $f_i(t,\mathbf{x},u_i) = r_i(\mathbf{x})u_i(t,\mathbf{x})$, avec :

$$r_i(\mathbf{x}) = r_{\max} - \frac{\|\mathbf{x} - \mathbf{O}_i\|^2}{2},$$

revient au système démographique :

$$\forall t > 0, \begin{cases} N_1'(t) = \overline{r}_1(t)N_1(t) + \delta[N_2(t) - N_1(t)], \\ N_2'(t) = \overline{r}_2(t)N_2(t) + \delta[N_1(t) - N_2(t)], \end{cases}$$

où $N_i(t) = \int_{\mathbb{R}^n} u_i(t,\mathbf{x})\mathrm{d}\mathbf{x}$ est la taille de population dans l'environnement $i$ ;

◇ Second type – Cas logistique : Le choix :

$$f_i(t,\mathbf{x},u_i) = \left(r_i(\mathbf{x}) - \int_{\mathbb{R}^n} u_i(t,\mathbf{y})\mathrm{d}\mathbf{y}\right) u_i(t,\mathbf{x}) = [r_i(\mathbf{x}) - N_i(t)]u_i(t,\mathbf{x}),$$

donne le système démographique suivant :

$$\forall t>0,\ \begin{cases} N_1'(t)=\overline{r}_1(t)N_1(t)-N_1^2(t)+\delta[N_2(t)-N_1(t)],\\ N_2'(t)=\overline{r}_2(t)N_2(t)-N_2^2(t)+\delta[N_1(t)-N_2(t)],\end{cases}$$

Tout comme dans la Partie I, nous devons faire des hypothèses sur la condition initiale $(u_1^0,u_2^0)$ :

(i) $u_1^0(x_1,\dots,x_n)=u_2^0(-x_1,x_2,\dots,x_n)$ pour tout $(x_1,\dots,x_n)\in\mathbb{R}^n$ ;

(ii) $u_1^0\in C^2,\alpha(\mathbb{R}^n)$ pour $\alpha\in(0,1)$ ;

(iii) $N^0:=N(0)=\int_{\mathbb{R}^n}u_1^0(\mathbf{x})\mathrm{d}\mathbf{x}>0$ et $N^0<+\infty$ ;

(iv) Il existe une fonction décroissante $g:\mathbb{R}_+\to\mathbb{R}$ telle que pour tout $\mathbf{x}\in\mathbb{R}^n$, $0\le u_1^0(\mathbf{x})\le g(\|\mathbf{x}-\mathbf{O}_1\|)$ et la fonction $r\mapsto r^{n+1}g(r)$ est intégrable et tend vers 0 quand $r\to+\infty$.

**Théorème 7.**
*Supposons que $f_1,f_2$ sont soit du premier type (cas Malthusien) soit du second type (cas logistique). Il existe alors une unique solution $\mathbf{u}=(u_1,u_2)\in C^{1,2}(\mathbb{R}_+\times\mathbb{R}^n)$ de (1.19), telle que $\mathbf{u}\in L^\infty((0,T)\times\mathbb{R}^n)$ pour tout $T>0$, $u_i(t,\mathbf{x})\to 0$ quand $\|\mathbf{x}\|\to+\infty$ uniformément localement en $t\in\mathbb{R}_+$, $u_i>0$ dans $(0,+\infty)\times\mathbb{R}^n$, les tailles de population $N_i:\mathbb{R}_+\to(0,+\infty)$ sont de classe $C^1$, les fitness moyennes $\overline{r}_i:\mathbb{R}_+\to\mathbb{R}$ sont continues, $\mathbf{u}$ est symétrique :*

$$u_1(t,x_1,\dots,x_n)=u_2(t,-x_1,x_2,\dots,x_n),\quad \text{pour tout } t\ge0 \text{ et } \mathbf{x}=(x_1,\dots,x_n)\in\mathbb{R}^n,$$

*et :*

$$N_1(t)=N_2(t)=:N(t)\quad\text{et}\quad\overline{r}_1(t)=\overline{r}_2(t)=:\overline{r}(t),\quad\text{pour tout } t\ge0.$$

*Les fonctions $u_i$ satisfont l'équation parabolique non locale :*

$$\partial_t u_i(t,\mathbf{x})=\frac{\mu^2}{2}\Delta u_i(t,\mathbf{x})+f_i(t,\mathbf{x},u_i)+\delta[u_i(t,\iota(\mathbf{x}))-u_i(t,\mathbf{x})],$$

*pour tout $t\ge0$ et $\mathbf{x}\in\mathbb{R}^n$, où la fonction $\iota$ est définie par :*

$$\forall(x_1,\dots,x_n)\in\mathbb{R}^n,\quad \iota(x_1,\dots,x_n)=(-x_1,x_2,\dots,x_n)\in\mathbb{R}^n.$$

Pour tout $R>0$, l'opérateur différentiel auto-adjoint $\mathcal{A}$ :

$$\mathcal{A}:=-\frac{\mu^2}{2}\Delta-\begin{pmatrix} r_1(\mathbf{x})-\delta & \delta\\ \delta & r_2(\mathbf{x})-\delta\end{pmatrix},$$

agit sur l'ensemble fonctionnel $[W^{2,n}_{loc}(B(0,R))\cap C_0(\overline{B(0,R)})]^2$, avec $B(0,R)$ la boule ouverte euclidienne de $\mathbb{R}^n$ de centre 0 et de rayon $R>0$, et $C_0(\overline{B(0,R)})$ l'en-

semble des fonctions continues sur $\overline{B(0,R)}$ qui s'annulent sur $\partial B(0,R)$. D'après SWEERS 1992, Théorème 1.1, il existe une unique valeur propre $\lambda^{\delta,R} \geq -r_{\max}$ et un unique couple (à multiplication près par un scalaire) de fonctions propres positives (dans $B(0,R)$) $(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) \in [W^{2,n}_{loc}(B(0,R)) \cap C_0(\overline{B(0,R)})]^2$ telles que :

$$\mathcal{A}\,(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) = \lambda^{\delta,R}\,(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) \text{ dans } B(0,R).$$

La suite de valeurs propres $(\lambda^{\delta,R})_{R>0}$ converge quand $R \to +\infty$, vers une quantité $\lambda^\delta \geq -r_{\max}$. Cette valeur limite est à relier avec le comportement asymptotique de la population :

**Théorème 8. CROISSANCE MALTHUSIENNE : EXPLOSION VS EXTINCTION**
*Supposons que $f_1, f_2$ sont du premier type. Soit $\delta > 0$ et $\lambda^\delta$ définie précédemment. Soit $\mathbf{u}$ la solution de (1.19), et $N(t) = N_1(t) = N_2(t)$ la taille de population dans chaque habitat.*

*(i) Si $\lambda^\delta < 0$, alors $N(t) \to +\infty$ quand $t \to +\infty$ (explosion population).*

*(ii) Si $\lambda^\delta = 0$ et si $u^0$ est à support compact, alors :*

$$\limsup_{t\to+\infty} N(t) < +\infty \quad \text{(majoration de la population)}.$$

*(iii) Si $\lambda^\delta > 0$ et si $u^0$ est à support compact, alors $N(t) \to 0$ quand $t \to +\infty$ (extinction de la population).*

**Théorème 9. CROISSANCE LOGISTIQUE : PERSISTANCE VS EXTINCTION**
*Supposons que $f_1, f_2$ sont du second type. Soit $\delta > 0$ et $\lambda^\delta$ définie précédemment. Soit $\mathbf{u}$ la solution de (1.19), et $N(t) = N_1(t) = N_2(t)$ la taille de population dans chaque habitat.*

*(i) Si $\lambda^\delta < 0$, alors :*

$$0 < \liminf_{t\to+\infty} N(t) \leq \limsup_{t\to+\infty} N(t) < +\infty \quad \text{(persistance de la population)},$$

*pour certaines conditions initiales $\mathbf{u}^0$.*

*(ii) Si $\lambda^\delta \geq 0$ et si $u^0$ est à support compact, alors $N(t) \to 0$ quand $t \to +\infty$ (extinction de la population).*

Grâce à une étude de la fonction $\delta \mapsto \lambda^\delta$ (concavité et limite quand $\delta \to 0$ et $\delta \to +\infty$), nous finissons par obtenir un critère sur les paramètres du modèle :

**Corollaire 1.**
*Supposons que $\lambda^0 < 0$, i.e., $r_{\max} > \mu\, n/2$. Soit $\delta > 0$, $\mathbf{u}$ la solution de (1.19) et $N(t) = N_1(t) = N_2(t)$ la taille de population dans chaque habitat.*

*(i) Si $\lambda^\infty \leq 0$, i.e., si $m_D \leq 4\,(r_{\max} - \mu\, n/2)$, alors $\lim_{t\to+\infty} N(t) = +\infty$ dans le cas Malthusien, tandis que :*

$$0 < \liminf_{t\to+\infty} N(t) \leq \limsup_{t\to+\infty} N(t) < +\infty \quad (\textit{persistance de la population}),$$

*dans le cas logistique, pour certaines conditions initiales $\mathbf{u}^0$.*

*(ii) Si $\lambda^\infty > 0$, i.e., si $m_D > 4\,(r_{\max} - \mu\, n/2)$, alors il existe un taux de migration critique $\delta_{crit} > 0$, indépendant de la donnée initiale $\mathbf{u}^0$, telle que :*

*(ii-a) si $\delta < \delta_{crit}$, alors $\lim_{t\to+\infty} N(t) = +\infty$ dans le cas Malthusien, , tandis que :*

$$0 < \liminf_{t\to+\infty} N(t) \leq \limsup_{t\to+\infty} N(t) < +\infty \quad (\textit{persistance de la population}),$$

*dans le cas logistique, pour certaines conditions initiales $\mathbf{u}^0$ ;*

*(ii-b) si $\delta = \delta_{crit}$ et si $u^0$ est à support compact, alors $\limsup_{t\to+\infty} N(t) < +\infty$ dans le cas Malthusien, et $N(t) \to 0$ quand $t \to +\infty$ dans le cas logistique ;*

*(ii-c) si $\delta > \delta_{crit}$ et si $u^0$ est à support compact, alors $\lim_{t\to+\infty} N(t) = 0$ pour les cas Malthusien et logistique.*

## Chapitre 6

Comme vu dans les chapitres précédents, la sélection, les mutations et la migration ont un impact joint complexe sur la dynamique (évolutive et démographique) de la population, et donc sur celle de la fitness moyenne de chaque sous-population. En effet, les mutations et la migration augmentent la variance locale des phénotypes, mais génèrent des fardeaux évolutifs (mutationnels et migratoires), suite aux mutations délétères et à l'arrivée d'individus maladaptés (Blows ; Hoffmann 2005 ; Lenormand 2002).

Suite à ces différentes difficultés, il semble nécessaire de faire certaines simplifications du modèle. Certaines recherches actuelles se sont intéressées par exemple aux états d'équilibres (Gandon ; Mirrahimi 2017b ; Meszéna ; Czibula ; Geritz 1997 ; Mirrahimi ; Gandon 2020), ou à l'évolution démographique sans évolution (Brown ; Pavlovic 1992 ; Freedman ; Waltman 1977 ; Hastings 1983 ; Holt 1985 ; Holt ; Gomulkiewicz 1997 ; Kawecki 1995 ; Kawecki 2000 ; Kawecki ; Holt 2002 ; Levin 1974 ; Pulliam 1988). Nous continuerons ici à utiliser

certaines hypothèses utilisées par ces travaux : générations non chevauchantes, croissance Malthusienne, deux habitats connectés par une migration symétrique.

Dans la continuité du Chapitre 5, nous étudierons le système (1.19). À partir de là (comme dans le cas d'un seul environnement, *cf.* Partie I), nous pourrons obtenir une équation différentielle satisfaite par les densités $\mathbf{p}_1$ et $\mathbf{p}_2$, des vecteurs fitness définis par :

$$\mathbf{m} = (m_1, m_2), \qquad \text{avec } m_i(\mathbf{x}) = -\frac{\|\mathbf{x} - \mathbf{O}_i\|^2}{2},$$

pour tout $\mathbf{x} \in \mathbb{R}^n$ et $i \in \{1, 2\}$. Ainsi la transformée de Laplace de cette densité (généralisation de la fonction génératrice des cumulants) est solution d'une équation différentielle d'ordre 1 :

**Proposition 10.** *Les transformées de Laplace de $\mathbf{p}_1$ et $\mathbf{p}_2$ :*

$$\forall i \in \{1,2\},\ \forall t \geq 0,\ \forall (z_1, z_2) \in \mathbb{R}_+^2,\ M_i(t, z_1, z_2) = \int e^{z_1 m_1 + z_2 m_2} \mathbf{p}_i(t, m_1, m_2) \mathrm{d}m_1 \mathrm{d}m_2,$$

*sont bien définies. La fonction vectorielle $\mathbf{M} = (M_1, M_2)$ est une solution $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^2)$ de :*

$$\forall t > 0,\ \forall (z_1, z_2) \in \mathbb{R}_+^2,\ \partial_t \mathbf{M}(t, \mathbf{z}) = S_1(\mathbf{z}) \partial_{z_1} \mathbf{M}(t, \mathbf{z}) + S_2(\mathbf{z}) \partial_{z_2} \mathbf{M}(t, \mathbf{z}) + A(\mathbf{z}) \mathbf{M}(t, \mathbf{z}),$$

*où les matrices sont :*

$$S_1(\mathbf{z}) = \begin{pmatrix} 1 - \mu^2 (z_1 + z_2) z_1 & 0 \\ 0 & -\mu^2 (z_1 + z_2) z_1 \end{pmatrix}, \quad S_2(\mathbf{z}) = \begin{pmatrix} -\mu^2 (z_1 + z_2) z_2 & 0 \\ 0 & 1 - \mu^2 (z_1 + z_2) z_2 \end{pmatrix},$$

$$A(\mathbf{z}) = \begin{pmatrix} r_{\max} - \mu_1^2 \left[ m_D z_1 z_2 + \frac{n}{2} (z_1 + z_2) \right] - \delta & \delta \\ \delta & r_{\max} - \mu_2^2 \left[ m_D z_1 z_2 + \frac{n}{2} (z_1 + z_2) \right] - \delta \end{pmatrix},$$

*et les paramètres $\mu^2 = \lambda U$ et $m_D = 2\beta^2$.*

Dans le cas d'habitats isolés ($\delta = 0$), nous pouvons obtenir une solution explicite, qui mène aux même résultats que Lavigne ; Martin ; Anciaux ; Papaix et al. 2020 ; Martin ; Roques 2016, sur le fardeau mutationnel :

$$\mathrm{Load}_{\mathrm{mut}}(\mu_i, n) = \lim_{t \to \infty} \overline{m}_i(t) = -\frac{\mu_i n}{2}.$$

Cependant, une formule dans le cadre général $\delta > 0$ semble compliquée à obtenir. Nous allons alors utiliser un système approchant celui de la Proposition 10. Pour ce faire, nous remarquons que ce système différentiel peut se simplifier en

une seule équation différentielle d'ordre 1 non locale :

$$\begin{aligned}\partial_t M_1(t,\mathbf{z}) &= [1-\mu^2(z_1+z_2)z_1]\partial_{z_1}M_1(t,\mathbf{z}) - \mu^2(z_1+z_2)z_2\partial_{z_2}M_1(t,\mathbf{z})\\ &\quad + \left[r_{\max} - \mu^2\left(m_D z_1 z_2 + \frac{n}{2}(z_1+z_2)\right)\right] M_1(t,\mathbf{z}) + \delta[M_1(t,z_2,z_1) - M_1(t,\mathbf{z})],\end{aligned}$$

dans le cas où les distributions initiales sont symétriques :

$$\forall (x_1,\dots,x_n)\in\mathbb{R}^n,\ u_1^0(x_1,\dots,x_n) = u_2^0(-x_1,x_2,\dots,x_n),$$

avec $M_2(t,z_1,z_2) = M_1(t,z_2,z_1)$. Les formules de Taylor nous permettent alors d'approcher cette équation par l'équation suivante :

$$\begin{aligned}\partial_t M_1(t,\mathbf{z}) \approx [1+\delta(z_2-z_1)-\mu^2(z_1+z_2)z_1]\partial_{z_1}M_1(t,\mathbf{z})\\ + \left[\delta(z_1-z_2) - \mu^2(z_1+z_2)z_2\right]\partial_{z_2}M_1(t,\mathbf{z})\\ + \left[r_{\max} - \mu^2\left(m_D z_1 z_2 + \frac{n}{2}(z_1+z_2)\right)\right] M_1(t,\mathbf{z}),\end{aligned}$$

qui admet une solution explicite. Celle-ci implique une étude de la dynamique de la population en tout temps, mais aussi une approximation de la valeur critique $\delta_{crit}$ décrite dans le Chapitre 5, ainsi que du fardeau de migration :

**Conjecture 11.** *Pour $m_D \leq 4\left(r_{\max} - \frac{\mu n}{2}\right)$, la population explose, alors que, si $m_D > 4\left(r_{\max} - \frac{\mu n}{2}\right)$, il existe un taux de migration critique $\delta^* > 0$ :*

$$\delta^* = \delta^*(r_{\max},\mu,m_D,n) := \frac{\mu}{2}\left[-1+\sqrt{\frac{m_D}{m_D - 4\left(r_{\max}-\frac{\mu n}{2}\right)}}\right],$$

*tel que :*

- ◇ *pour $\delta < \delta^*$, la population croît exponentiellement ;*
- ◇ *pour $\delta > \delta^*$, la population s'éteint.*

*La fitness moyenne converge vers :*

$$\overline{m}(\infty) = -\frac{\mu n}{2} - \frac{m_D}{4}\left(1+\mu^2 \underset{\delta\to+\infty}{o}(1)\right),$$

*quand $t\to+\infty$.*

# Partie III

## Chapitre 7

Ce chapitre se concentre sur une équation de réaction-diffusion, modélisant une population dans un environnement continu hétérogène. De plus, la dispersion est influencée par son phénotype $\theta \in \mathbb{R}$. Bon nombre de résultats sont déjà connus sur la vitesse de propagation asymptotique d'une population asexuée (BERESTYCKI ; MOUHOT ; RAOUL 2015 ; BOUIN ; HENDERSON ; RYZHIK 2017 ; CALVEZ ; HENDERSON ; MIRRAHIMI ; TURANOVA et al. 2018).

Nous nous concentrerons donc sur l'équation :

$$\partial_t f(t,x,\theta) = r_1 f(t,x,\theta)[1 - K^{-1}\varrho(t,x)] + \theta\Delta_x f(t,x,\theta) + B[f](t,x,\theta), \quad (1.20)$$

où $f(t,x,\theta)$ est la densité des individus présentant le phénotype $\theta \in (\theta_{\min}, +\infty)$, à la position $x \in \mathbb{R}$ au temps $t > 0$. La quantité $\varrho(t,x) = \int_{\theta_{\min}}^{+\infty} f(t,x,\theta)\mathrm{d}\theta$ correspond au nombre d'individus présents en $x \in \mathbb{R}$ au temps $t \geq 0$. Ce modèle suppose que chaque individu peut envahir un nouvel environnement, selon son phénotype. Nous ferons l'hypothèse que la densité de population initiale est à support compact.

Le terme $r_1 f[1-K^{-1}\varrho]$ représente la sélection avec densité dépendance, comme dans les modèles précédents, où le taux de sélection dépend de la densité $f$ (au travers de $\varrho$). Le terme $\theta\Delta_x f$ modélise la diffusion spatiale, se produisant selon un taux valant le trait phénotypique $\theta$. Ainsi plus $\theta$ sera élevé, plus les individus de phénotype $\theta$ ont une mobilité importante.

Il reste à décrire le terme de reproduction $B[f]$. Le Chapitre 7 étudie le cas de populations asexuées. Ainsi la reproduction étant clonale, le descendant reçoit à sa naissance le même trait que son parent, à modifications prêts dues aux mutations. En faisant l'hypothèse que ces variations sont faibles, ce phénomène peut finalement être approché en l'EDP suivante (BENICHOU ; CALVEZ ; MEUNIER ; VOITURIEZ 2012) :

$$\partial_t f(t,x,\theta) = r_1 f(t,x,\theta)[1 - K^{-1}\varrho(t,x)] + \theta\Delta_x f(t,x,\theta) + \mu\Delta_\theta f(t,x,\theta),$$

où $\mu > 0$ est une constante dépendant du taux de mutation et de la variance mutationnelle. Quitte à étudier la fonction $\theta_{\min}K^{-1}f(\theta_{\min}^2 t/\mu, \sqrt{\theta_{\min}^3/\mu}x, \theta_{\min}\theta)$, on se ramène à l'EDP :

$$\partial_t f(t,x,\theta) = r f(t,x,\theta)[1 - \varrho(t,x)] + \theta\Delta_x f(t,x,\theta) + \Delta_\theta f(t,x,\theta),$$

pour $r > 0$, $\theta \in (1, +\infty)$ et $\varrho(t,x) = \int_1^{+\infty} f(t,x,\theta)\mathrm{d}\theta$. Nous ajouterons à cette équation, et à la condition initiale, une condition de bord :

$$\forall t \geq 0,\ \forall x \in \mathbb{R},\ \partial_\theta f(t,x,1) = 0.$$

Différentes recherches (BERESTYCKI ; MOUHOT ; RAOUL 2015 ; BOUIN ; HENDERSON ; RYZHIK 2017 ; CALVEZ ; HENDERSON ; MIRRAHIMI ; TURANOVA et al. 2018) ont pu montrer pour certaines valeurs de $r$ que la solution possède un front s'accélérant à la vitesse $3/2$, *i.e.*, que la position du front $X(t)$ est de la forme $c^* t^{3/2}$. Nous présenterons des résultats numériques dans ce chapitre, étant une base de comparaison pour le chapitre suivant.

### Chapitre 8

Le dernier chapitre porte sur l'invasion de population sexuée dans un environnement hétérogène. Ainsi le modèle considéré se base sur l'EDP (1.20). Cependant, le terme $B[f]$ sera modifié. En effet, lors de la naissance d'un individu, son phénotype sera supposé être la moyenne des phénotypes de ses parents, à modification prêt, suite à la recombinaison et à la ségrégation. Ainsi le modèle considéré sera :

$$\partial_t f(t,x,\theta) = r\left[\iint_{(\theta_{\min},+\infty)^2} \mathcal{G}_\lambda\left(\theta - \frac{\theta_1+\theta_2}{2}\right)\frac{f(t,x,\theta_1)f(t,x,\theta_2)}{\varrho(t,x)}\mathrm{d}\theta_1\mathrm{d}\theta_2 - K^{-1}\varrho(t,x)\right] + \theta\Delta_x f(t,x,\theta),$$

pour $r>0$ taux de reproduction et $\mathcal{G}_\lambda$ densité gaussienne normalisée de variance $\lambda^2>0$. En faisant un changement d'échelle similaire au cas asexué du Chapitre 7, nous pouvons nous ramener au cas où $r=1$, $K=1$ et $\theta_{\min}=1$.

Ce modèle a été numériquement étudié dans CALVEZ ; CREVAT ; DEKENS ; FABRÈGES et al. 2019. Ici, nous présenterons ces résultats, en ajoutant une conjecture, s'appuyant sur une démonstration formelle, sur la vitesse de propagation du front :

**Conjecture 12.** *Définissons la constante :*

$$y_c = 4r^{3/4}\sqrt{\frac{\lambda}{3}}.$$

*Ainsi la densité $f$ en temps grand $t\geq 0$ peut être approchée par :*

$$f_{approx}(t,x,\theta) = \frac{K}{\theta_{\min}}\begin{cases} \exp\left[-\frac{1}{4\lambda^2}\left[\theta-\lambda^{4/5}\left(6rx^2\right)^{1/5}\right]^2\right], & \text{pour } x\leq y_c\, t^{5/4},\\ \exp\left[rt-\left(\frac{9x^4}{256\,\lambda^2t^2}\right)^{1/3}\right]\exp\left[-\frac{1}{4\lambda^2}\left[\theta-\left(\frac{3\lambda^2x^2}{2t}\right)^{1/3}\right]^2\right], & \\ & \text{pour } x\geq y_c\, t^{5/4}. \end{cases}$$

Ce résultat nous permet de mieux comprendre la différence des invasions de populations sexuées et asexuées. Par ailleurs, cette conjecture, qui semble être

validée par des solutions numériques de l'équation, permet de prédire l'évolution de différentes invasions d'espèces sexuées qui ont lieu actuellement (crapaud buffle en Australie, Martin triste en Afrique du Sud, sauterelles en Angleterre, *etc.*)

## Perspectives

1. Dans le Chapitre 3, l'optimum bien que supposé mobile, ne peut se mouvoir que dans une direction précise (à savoir un des axes du repère). Cependant, les modifications de paramètres abiotiques de l'environnement peuvent agir sur différents traits phénotypiques. Ainsi, l'extension naturelle de la Partie I est d'étudier un optimum $\mathbf{O}(t) \in \mathbb{R}^n$, pouvant bouger dans n'importe quelle direction. Nous devrons alors prendre un autre vecteur fitness, pour pouvoir espérer capturer l'évolution analytique de la fitness moyenne. L'idée première serait de définir le vecteur fitness $\mathbf{v}(\mathbf{x}) \in \mathbb{R}^{2n}$ étant :

$$\forall \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n,\ \mathbf{v}(\mathbf{x}) = \left(x_1, \dots, x_n, -\frac{x_1^2}{2}, \dots, -\frac{x_n^2}{2}\right) \in \mathbb{R}^{2n}.$$

   Le même genre d'étude, que celle faite dans le Chapitre 3, semble pouvoir se réaliser. Seulement, les calculs semblent plus complexes.

2. L'étude de deux habitats connectés par des phénomènes migratoires, faite dans la Partie II, mène à différents scenarii possibles : le puits à effet trou noir et les puits réciproques symétriques. Pour ce dernier, l'évolution de la transformée de Laplace a été approchée, sans avoir pu étudier l'erreur commise. Celle-ci semble importante pour comprendre l'effet de cette approximation sur le coefficient d'adaptation locale $LA$.

   De plus, cette étude semble permettre d'approximer le cadre général asymétrique (voir la discussion du Chapitre 6). Il reste encore à justifier ce phénomène de simplification d'étude du cas asymétrique au cas symétrique.

3. La Partie III permet de généraliser le cas de migration de la Partie II dans le cas d'un environnement continu. Cependant, il reste à étudier le cas intermédiaire d'îles dénombrables connectées par migration.

   En outre, l'invasion d'une population sexuée dans un milieu continu n'a été que partiellement traitée. En effet, les résultats restent numériques et formels. Une preuve rigoureuse est encore à fournir pour l'étude du front de propagation.

4. La démographie a été supposée principalement suivre une croissance exponentielle. Cependant, comme vu dans l'article LAVIGNE ; MARTIN ; ANCIAUX ; PAPAIX et al. 2020 (ou Chapitre 4), l'hypothèse de croissance exponentielle peut donner des résultats inattendus, comme l'indépendance du temps d'adaptation à un nouvel habitat par rapport au taux de migration, dans le

cas du système à effet trou noir. Nous aurions pu aussi étudier ce problème en prenant en compte la possibilité d'une capacité de charge $K > 0$, et donc en considérant une croissance logistique. Ce modèle démographique semble *a priori* donner une évolution de la fitness moyenne dépendante de $d$ mais pas de $d/K$.

5. Le modèle évolutif utilisé lors de cette thèse repose sur le modèle de Wright-Fisher, couplé avec le Modèle Géométrique de Fisher. Bien que celui-ci a donné bon nombre d'applications biologiques, nous aurions pu appliquer d'autres modèles évolutifs. En effet, cette approche déterministe n'est valide que dans le cadre de forte mutation et faible sélection – régime appelé WSSM. Pour pallier cela, nous pourrions étudier le modèle "House of Cards" (TURELLI 1984), donnant une bonne approximation déterministe de l'évolution d'une population asexuée si les mutations sont faibles.

   Ce modèle fait l'hypothèse que les mutants possèdent une fitness absolue, distribuée selon une distribution unique, indépendante de celle du parent. La dépendance au patrimoine génétique intervient dans la distribution des effets des mutations :

$$f(s|m) = g(s+m), \qquad \text{pour } m \text{ la fitness du parent.}$$

   Ainsi la fitness de l'enfant mutant est distribuée selon la fonction $g$ (*cf.* KRYAZHIMSKIY ; TKAČIK ; PLOTKIN 2009 et MCCANDLISH ; STOLTZFUS 2014 dans le cas $g$ exponentielle ou gaussienne). Cela implique une modification des équations différentielles vérifiées par la distribution des fitness et par sa fonction génératrice (MARTIN ; ROQUES 2016).

6. Dans le cas de faible mutation, nous pourrions aussi utiliser des outils probabilistes, comme dans ANCIAUX ; CHEVIN ; RONCE ; MARTIN 2018. Cela permettrait d'exprimer des quantités comme des probabilités de survie, d'extinction, d'invasion, *etc.*

| Notation | Description |
|---|---|
| $\mathbf{x}$ | Phénotype |
| $r$ | Fitness Malthusienne absolue |
| $r_{\max}$ | Maximum de la fitness Malthusienne absolue |
| $m$ | Fitness Malthusienne relative |
| $\mathbf{m}$ | Vecteur de composantes de fitness |
| $u(t, \mathbf{x})$ | Densité phénotypique de la population au temps $t$ |
| $q(t, \mathbf{x})$ | Distribution phénotypique de la population au temps $t$ |
| $p(t, m)$ | Distribution de fitness au temps $t$ |
| $\mathbf{p}(t, \mathbf{m})$ | Distribution du vecteur des composantes de fitness au temps $t$ |
| $N, N(t)$ | Taille de la population au temps $t$ |
| $\overline{r}_t$, $\overline{r}(t)$ | Fitness moyenne (absolue) au temps $t$ |
| $\overline{m}_t$, $\overline{m}(t)$ | Fitness moyenne (relative) au temps $t$ |
| $\overline{X}$ | Valeur moyenne d'une variable aléatoire $X(m)$, moyennée sur la distribution phénotypique dans une population |
| $\langle X \rangle$ | Moyenne de la variable $X$ sur un nombre (fini) de replicats de population |
| $M_t(\cdot)$, $M(t, \cdot)$ | MGF (ou transformée de Laplace) au temps $t$ |
| $C_t(\cdot)$, $C(t, \cdot)$ | CGF au temps $t$ |
| $U$ | Taux de mutation |
| $\mathbf{x}^*$, $\mathbf{O}$ | Optimum phénotypique |
| $\delta$ | Taux de migration par individus |
| $d$ | Taux de migation (puits à effet trou noir) |
| $m^\star$ | Fitness Malthusienne des migrants dans le puits |
| $m_D$ | Différence entre deux habitats, Stress environnemental |
| $n$ | Dimension de l'espace phénotypique |
| $\lambda$ | Variance mutationnelle de chaque trait (cas isotrope) |
| $\Lambda$ | Matrice de variance–covariance de l'effet des mutations |

TABLE 1.1. – **Principales notations utilisées tout au long de ce manuscrit**.

# I

# Adaptation into a unique environment

2

# Dynamics of adaptation in an anisotropic phenotype-fitness landscape

**F. HAMEL [a], F. LAVIGNE[a,b,c], G. MARTIN [c] and L. ROQUES[a]**

[a] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
[b] BioSP, INRA, 84914, Avignon, France
[c] ISEM (UMR 5554), CNRS, 34095, Montpellier, France

**Abstract**

We study the dynamics of adaptation of a large asexual population in a $n$-dimensional phenotypic space, under anisotropic mutation and selection effects. When $n = 1$ or under isotropy assumptions, the 'replicator-mutator' equation is a standard model to describe these dynamics. However, the $n$-dimensional anisotropic case remained largely unexplored.

We prove here that the equation admits a unique solution, which is interpreted as the phenotype distribution, and we propose a new and general framework to the study of the quantitative behavior of this solution. Our method builds upon a degenerate nonlocal parabolic equation satisfied by the distribution of the 'fitness components', and a nonlocal transport equation satisfied by the cumulant generating function of the joint distribution of these components. This last equation can be solved analytically and we then get a general formula for the trajectory of the mean fitness and all higher cumulants of the fitness distribution, over time. Such mean fitness trajectory is the typical outcome of empirical studies of adaptation by experimental evolution, and can thus be compared to empirical data.

In sharp contrast with the known results based on isotropic models, our results show that the trajectory of mean fitness may exhibit $(n-1)$ plateaus before it converges. It may thus appear 'non-saturating' for a transient but possibly long time, even though a phenotypic optimum exists. To illustrate the empirical relevance of these results, we show that the anisotropic model leads to a very good fit of *Escherichia coli* long-term evolution experiment, one of the most famous experimental dataset in experimental evolution. The two 'evolutionary epochs' that have been observed in this experiment have long puzzled the community: we propose that the pattern may simply stem from a climbing hill process, but in an anisotropic fitness landscape.

# Sommaire

## 2.1. Introduction

### Biological motivation

Understanding the adaptation of asexual organisms (such as viruses, bacteria or cancer cells) under the combined effects of selection and mutation is a fundamental issue in population genetics.

In parallel, the development of experimental evolution (especially in microbes) has made it possible to compare observed dynamics, in the lab, with alternative models or infer various model parameters from the data (for a recent special issue on this subject see Rosenzweig; Sherlock 2014). Still, the problem of modelling asexual evolutionary dynamics is inherently complex (discussed in e.g. Gerrish; Sniegowski 2012): recurrent mutation changes the key parameters of the dynamical system constantly, competition between numerous, ever changing types must be described, and both mutational and demographic (birth/death) events are stochastic in nature.

Recent models of asexual adaptation seek to follow the dynamics of the full distribution of fitness - the expected reproductive output of a lineage - within populations. Contrarily to other approaches, such as 'origin fixation models' which only follow the expected mean fitness $\overline{m}(t)$ (Sniegowski; Gerrish 2010), these models do not make a low polymorphism assumption, but in exchange for ignoring or simplifying the stochastic components of the dynamics. The resulting outcome is a partial differential equation (PDE) or an integro-differential equation (IDE), that typically describes the dynamics of the distribution of a single 'trait'. In some cases, this trait may be fitness itself as in Alfaro; Carles 2014a; Gil; Hamel; Martin; Roques 2017; Tsimring; Levine; Kessler 1996, leading to equations of the form:

$$\partial_t p(t,m) = \mathcal{M}_1[t,m,p(t,m)] + (m - \overline{m}(t))\, p(t,m),$$

where the variable $m \in \mathbb{R}$ is the fitness and $\mathcal{M}_1$ is a differential or an integral operator describing the effect of mutations on the distribution $p(t,\cdot)$ of fitness. Other models describe the distribution of a given trait $x \in \mathbb{R}$ determining fitness (birth rate, phenotype), as in Alfaro; Carles 2017; Champagnat; Ferrière; Méléard 2006:

$$\partial_t q(t,x) = \mathcal{M}_2[t,x,q(t,x)] + (m(x) - \overline{m}(t))\, q(t,x),$$

where $m(x)$ is a function which describes the relationship between the trait $x$ and the fitness. Here, $\mathcal{M}_2$ is a differential or an integral operator describing the effect of mutations on the distribution $q(t,\cdot)$ of the trait $x$. In these two equations, the last term $(m - \overline{m})\, p(t,m)$ (resp. $(m(x) - \overline{m}(t))\, q(t,x)$) corresponds to the effects of selection and will be explained later. Broadly, it emerges when the growth rate $m(x)$ (in the absence of competition) of a genotype with phenotype $x$ depends only on its own phenotype. This corresponds to the classic frequency-independent model of selection.

One established finding from empirical fitness trajectories is that epistasis, namely the fact that the distribution of a mutation's effect on fitness depends on its ancestor's genetic background, must be accounted for to explain the data (*e.g.* Kryazhimskiy; Tkačik; Plotkin 2009). More precisely, fitness trajectories tend to decelerate over time, implying *a priori* that beneficial mutations become weaker and/or rarer as the population adapts (ignoring deleterious ones, which is of course debatable). The question is then: which particular form of epistasis does explain/predict the data best? A common metaphor to explain this decelerating pattern in fitness trajectories is to invoke some form of 'fitness landscape' connecting genotypes or phenotypes with fitness, as in the function $m(x)$ above, with one or several adaptive peak(s) where fitness is maximal (discussed in Elena; Sanjuán 2003).

In this view, deceleration in fitness trajectories stems from the hill climbing process of adaptation up a fitness peak. This view is appealing because of its intuitive/visual illustration, but also because of a particular form of landscape, Fisher's geometrical model (FGM; a single peak phenotype-fitness landscape), has shown promising potential when compared to empirical measures of fitness epistasis. In the FGM, a multivariate phenotype at a set of $n$ traits (a vector $\mathbf{x} \in \mathbb{R}^n$) determines fitness. The most widely used version assumes a quadratic form of the Malthusian fitness function $m(\mathbf{x})$, which decreases away from a single optimum $\mathbf{O} \in \mathbb{R}^n$ (Martin; Lenormand 2015; Tenaillon 2014). Without loss of generality, the traits can be defined such that the optimum sets the origin of phenotype space:

$$m(\mathbf{x}) = -\frac{\|\mathbf{x}\|^2}{2},$$

with $\|\cdot\|$ the Euclidian norm in $\mathbb{R}^n$. To describe the mutation effects on phenotypes, the standard 'isotropic Gaussian FGM' uses a normal distribution $\mathcal{N}(0, \lambda\, I_n)$ with $\lambda > 0$ the phenotypic mutational variance at each trait and $I_n$ the identity matrix.

Given the potential complexity of the relationship between genotype, phenotype and fitness, the FGM may appear highly oversimplified. However, several arguments suggest that it can be more than a mere heuristic and may be a relevant, empirically sound null model of adaptation (also discussed in a recent review of the FGM Tenaillon 2014). First (Martin 2014), the Gaussian FGM can be retrieved from a much less constrained model of phenotype-to-fitness landscape: Random Matrix theory arguments show that it emerges as a limit of highly integrative phenotypic networks, near an optimum (with an arbitrary fitness function and distribution of the original phenotypes). Second, several assumptions of the classic FGM have been tested quite extensively on empirical distributions of mutation effects on fitness, showing its power to fit or predict the observed patterns. For example, the FGM has shown to accurately predict observed distributions of epistasis among random and beneficial mutations in a virus and a bacterium (Martin; Elena; Lenormand 2007), to accurately fit patterns of epistasis among beneficial mutations in a fungus (Schoustra; Hwang;

Krug; Visser 2016) or the pattern of re-adaptation (compensation) from different deleterious mutant backgrounds in a 'recovery' experiment with a bacterium (Perfeito; Sousa; Bataillon; Gordo 2014). It also accurately predicted how stressful conditions (lowering the parental fitness) affect the mean and variance of mutation fitness effects (Martin; Lenormand 2006b). The assumption of a quadratic fitness function is perhaps the most critical to the patterns tested so far: various empirical tests showed that a deviation from the quadratic would actually lead to a lower fit of the data (see Fraïsse; Welch 2019 and fig. 5 in Martin; Lenormand 2006b). Overall, this suggests that, in spite of its apparent simplicity, the FGM with a quadratic fitness function is not so oversimplified: it is a natural limit of a diversity of more complex models and is consistent with various empirical patterns of mutation fitness effects.

On the other hand, this 'fitness peak' view is challenged by the observation (Wiser; Ribeck; Lenski 2013) that, in the longest evolution experiment ever undertaken (Long Term Evolution Experiment LTEE, in the bacterium *Escherishia coli*), fitness has not reached any maximum after more than 70,000 generations. It has been suggested (Good; Desai 2015) that this experiment actually shows a 'two epochs' dynamics with an initial (potentially saturating) fitness trajectory and a later and distinct non-saturating dynamics. A similar pattern could be invoked in another mid-term experiment with an RNA virus (Novella; Duarte; Elena; Moya, et al. 1995). Several other experiments did seem to yield a 'true' saturation (plateau) in fitness trajectories over several thousands of generations in *E. coli* (*e.g.* Crécy-Lagard; Bellalou; Mutzel; Marlière 2001; Fong; Palsson 2004; LaCroix; Sandberg; O'Brien; Utrilla, et al. 2015), but they may simply be on a too short timescale to identify subsequent fitness increases. In fact, the LTEE itself did seem to show a strongly saturating pattern after 10,000 generations (Lenski; Travisano 1994).

Overall, these different insights on epistasis and adaptation trajectories are difficult to reconcile under a single theory: why would a single peak model show a good fit to mutation epistasis over single mutations, or short term fitness trajectories, yet be invalid over the longer term? Is there really a two 'epochs' dynamics in some long-term trajectories, when and why? A desirable model should reconcile both timescales, describe both epistasis among random single mutations and in long term fitness trajectories. Proposed models that do accurately fit non-saturating trajectories (Wiser; Ribeck; Lenski 2013) have the drawback that they only focus on beneficial mutations and do not yield any prediction on the distribution of epistasis among random mutations. On the other hand, simple fitness landscapes like the FGM do predict some short term patterns and the full distribution of epistasis among random mutants, but cannot yield a never-ending fitness trajectory (a fitness plateau must be reached at the optimum).

### Aim of the paper

In this article, we explore the possibility that an extended version of the FGM be able to fit multiple epochs fitness trajectories. The central extension proposed here is to consider anisotropy in the landscape, in that different underlying traits affect fitness differently and mutate differently. Namely, the fitness function associated with a trait $\mathbf{x} = (x_1, \dots, x_n)$ is:

$$m(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i \, m_i(\mathbf{x}), \text{ where } m_i(\mathbf{x}) = -\frac{x_i^2}{2}, \tag{2.1}$$

where the coefficients $\alpha_i$ are positive. The mutation effects on phenotypes are described with an anisotropic Gaussian distribution $\mathcal{N}(0, \Lambda)$, where $\Lambda = \text{diag}\,(\lambda_1, \dots, \lambda_n)$ is any given positive diagonal matrix.

Using central limit and random matrix theory arguments (Martin 2014), the FGM can be obtained as a limit of a much less constrained model where high-dimensional phenotypes integrate into a phenotypic network to a smaller set of traits that directly determine fitness, with an optimum. The resulting FGM, however, is not necessarily isotropic, it can also show a single dominant direction in phenotype space (that affects fitness much more) with all other directions remaining approximately equivalent. Our initial intuition is that, in these conditions, adaptation along the dominant direction will drive the early dynamics while a second 'epoch' will be visible when adaptation only proceeds along the remaining dimensions. Therefore, it seems natural to explore how such a model would fare when compared to the fitness trajectories of the LTEE. Yet, this requires deriving the fitness dynamics resulting from such a complex mutational model, in the presence of mutation and competition between asexual lineages. This is the aim of the present paper.

The existence of a phenotype optimum has been taken into account in recent PDE and IDE approaches (second class of models alluded to above). For instance, in Alfaro; Carles 2017; Alfaro; Veruete 2018, the fitness depends on a single phenotypic trait $x \in \mathbb{R}$, through a function $m(x)$ which admits an optimum. Extending these works to take into account the dependence of the fitness on several traits appears as a natural question; especially since we know that the number of traits affected by mutation and subject to selection critically affects the evolutionary dynamics (Martin; Roques 2016; Waxman; Peck 1998). So far, and to the best of our knowledge, such mathematical models that take into account $n-$dimensional phenotypes together with the existence of a phenotype optimum always assume an isotropic dependence between the traits $\mathbf{x} \in \mathbb{R}^n$ and the fitness, and an isotropic effect of mutations on phenotypes (see Gil; Hamel; Martin; Roques 2019 for an IDE approach and Martin; Roques 2016 for an approach based on the analysis of a PDE satisfied by a moment generating function of the fitness distribution). Our goal here is to extend these approaches in order to characterize the dynamics of adaptation in a $n$-dimensional phenotypic space,

without making an isotropy assumption. For the sake of simplicity here, we ignore any stochastic component of the dynamics, but note that the same equations can be obtained explicitly as a deterministic limit of a stochastic model of mutation and growth (Martin; Roques 2016).

### Model assumptions and definitions

We assume that a phenotype is a set of $n$ biological traits, which is represented by a vector $\mathbf{x} \in \mathbb{R}^n$. To describe the evolution of the phenotypic distribution we begin with an intuitive description of the concept of fitness. In an asexual population of $K$ types of phenotypes $\mathbf{x}_1, \ldots, \mathbf{x}_K$, with $r_{\max} > 0$ the growth rate of the optimal phenotype, the *Malthusian fitness, relative to the optimal phenotype* is denoted $m_j = m(\mathbf{x}_j) \leq 0$ for a phenotype $\mathbf{x}_j \in \mathbb{R}^n$. It is defined by the following relation: $N_j'(t) = (r_{\max} + m_j) N_j(t)$, with $N_j(t)$ the abundance of the phenotype $\mathbf{x}_j$ at time $t$. When we sum these equations over all indexes $j = 1, \ldots, K$, we obtain an ordinary differential equation for the total population size $N(t) = \sum_{j=1}^K N_j(t)$ at time $t$: $N'(t) = (r_{\max} + \overline{m}(t)) N(t)$, where the quantity $\overline{m}(t) = \sum_{j=1}^K m_j N_j(t)/N(t)$ is the mean relative fitness in the population at time $t$. Now if we turn to the distribution of the phenotype frequencies $q(t, \mathbf{x}_j) = N_j(t)/N(t)$, we get the partial differential equation:

$$\partial_t q(t, \mathbf{x}_j) = (m(\mathbf{x}_j) - \overline{m}(t)) q(t, \mathbf{x}_j). \tag{2.2}$$

This equation can be generalized to a continuous distribution of phenotype frequencies (see *e.g.* Tsimring; Levine; Kessler 1996), as we assume in the sequel, with:

$$\overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})\, q(t, \mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{2.3}$$

If $m(\mathbf{x}) < \overline{m}(t)$, then (2.2) implies that the frequency of the phenotype $\mathbf{x}$ decreases, whereas if $m(\mathbf{x}) > \overline{m}(t)$, then the frequency increases.

We recall that mutation's effects on phenotypes are assumed to follow an anisotropic Gaussian distribution $\mathcal{N}(0, \Lambda)$, with $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ a positive diagonal matrix. Assuming a small mutational variance $\max \lambda_i \ll 1$, and a mutation rate $U$, these mutational effects can be approximated by an elliptic operator $\sum_{i=1}^n (\mu_i^2/2)\partial_{ii}$, where $\mu_i = \sqrt{U \lambda_i} > 0$ and $\partial_{ii}$ denotes the second order partial derivative with respect to the $i$-th coordinate of $\mathbf{x}$ (or later $\mathbf{m}$ as in (2.17) below). We refer to Appendix for further details on the derivation of this diffusion approximation. Intuitively, the regime where it applies is the 'Weak Selection Strong Mutation' (WSSM) regime (high rate of small effect mutations) where a wide diversity of lineages accumulate mutations and co-segregate at all times. The extreme alternative is the Strong Selection Weak Mutation (SSWM) regime where few mutations of large effect invade successively and the population is effectively close to monomorphic at all times. This latter regime will not be treated here.

Overall, the corresponding PDE describing the dynamics of the phenotype

distribution $q(t,\mathbf{x})$, under the combined effects of selection and mutation, is:

$$\partial_t q(t,\mathbf{x}) = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\,\partial_{ii} q(t,\mathbf{x}) + (m(\mathbf{x}) - \overline{m}(t))\,q(t,\mathbf{x}),\ t>0,\ \mathbf{x}\in\mathbb{R}^n, \tag{2.4}$$

with $\overline{m}(t)$ defined by (2.3). Recent studies (Alfaro; Carles 2017) have already treated the case $n=1$ (see also Alfaro; Veruete 2018 for more general fitness functions). In this paper, we consider the general case $n\geq 1$. Without loss of generality, we may assume in the sequel that $\alpha_1=\ldots=\alpha_n=1$ in (2.1), up to a scaling of the phenotype space ($\tilde{q}(t,\mathbf{x}) = (\prod_{i=1}^n \alpha_i^{1/2})\ q(t, x_1/\sqrt{\alpha_1},\ldots,x_n/\sqrt{\alpha_n})$). This leads to:

$$m\left(\mathbf{x}\right) = -\frac{\|\mathbf{x}\|^2}{2}, \tag{2.5}$$

with $\|\cdot\|$ the standard Euclidian norm in $\mathbb{R}^n$. Similarly, we could remove the anisotropy in the mutation effects, up to an other scaling of the phenotype space (defined this time by $\tilde{q}(t,\mathbf{x}) = (\prod_{i=1}^n \mu_i^{-1}) q(t,\mu_1 x_1,\ldots,\mu_n x_n)$), but in this case the coefficients $\alpha_i$ would be replaced by $\alpha_i\sqrt{\mu_i}$ and hence cannot be taken all equal to 1. Note that, for the sake of simplicity, we assumed a diagonal covariance matrix $\Lambda$ and that $m(\mathbf{x})$ was a linear combination of the fitness components $m_i(\mathbf{x}) = -x_i^2/2$. More general covariance matrices $\Lambda$ and quadratic forms $m(\mathbf{x})$ could have been considered as well using the transformations of the phenotype space presented in Martin; Lenormand 2006a.

In the one-dimensional case, an equation of the form (2.4) with a general nonlocal reaction term of the form $m(x) - \int_{\mathbb{R}} m(z)q(t,z)\mathrm{d}z$ has been studied in Alfaro; Veruete 2018. Under the assumption that $m(x)$ tends to $-\infty$ as $x$ tends to $\pm\infty$, the authors have established a formula for $q$ involving the eigenelements of the operator $\mathcal{H}_1 = \frac{\mu^2}{2}\mathrm{d}^2/\mathrm{d}x^2 + m(x)$. The formula can be made more explicit for our choice of $m$, $m(x) = -x^2/2$ (Alfaro; Carles 2017). However, the method used in Alfaro; Carles 2017, which consists in reducing the Eq. (2.4) to the heat equation thanks to changes of variables based on Avron-Herbst formula and generalized Lens transform, cannot be directly applied in our $n-$dimensional anisotropic framework. Recently, another method based on constrained Hamilton-Jacobi equations has been developed to study the evolution of phenotypically structured populations, with integral or differential mutation operators (*e.g.*, Barles; Mirrahimi; Perthame 2009; Diekmann; Jabin; Mischler; Perthame 2005; Gandon; Mirrahimi 2017b; Lorz; Mirrahimi; Perthame 2011; Perthame; Barles 2008). This method assumes a small mutation parameter $\mu^2/2$ (the terms $\mu_i^2/2$ in (2.4)) of order $\varepsilon \ll 1$, and is based on a scaling $t\to t/\varepsilon$. Thus, it typically describes asymptotic evolutionary dynamics, at large times and in a 'small mutation' regime. Note however that, as $\mu^2 = U\,\lambda$, with $U$ the mutation rate and $\lambda$ the mutational variance (at each trait), it encompasses the cases where the mutation rate is not small, provided that $\lambda \ll 1/U$. This method should apply with anisotropic mutation operators, as in (2.4). However, to the best of our knowledge, it cannot

lead to explicit transient trajectories of adaptation, which is the real objective of our work. We use here a completely different approach.

In order to understand the dynamics of adaptation, an important quantity is of course the *fitness distribution* $p(t,m)$, such that $p(t,m)\,\mathrm{d}m$ is the pushforward measure of the measure $q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ by the map $\mathbf{x}\mapsto -\|\mathbf{x}\|^2/2$, and the *mean fitness* $\overline{m}(t)$. In the isotropic case (Gil; Hamel; Martin; Roques 2019), the authors directly focused on the equation satisfied by $p(t,m)$ and not by $q(t,\mathbf{x})$. Whereas (2.4) is quadratic into the variables $x_1,\ \ldots,\ x_n$, the equation satisfied by $p(t,m)$ is linear with respect to $m$, which makes possible the derivation of some PDEs satisfied by generating functions of $p(t,m)$. Here, due to the anisotropy, the dynamics of $p(t,m)$ is not summarized by a single one-dimensional PDE. Instead, we define the fitness components $\mathbf{m}=(m_1,\ldots,m_n)$ and the *joint distribution of the fitness components* $\mathbf{p}(t,\mathbf{m})$ such that $\mathbf{p}(t,\mathbf{m})\,\mathrm{d}\mathbf{m}$ is the pushforward measure of the measure $q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ by the map $\mathbf{x}\mapsto(-x_1^2/2,\cdots,-x_n^2/2)$ (see Proposition 14 for more details). As fitness is the sum of the fitness components, the distributions $\mathbf{p}$ and $p$ are linked by:

$$p(t,m)=\int_{\mathbb{R}_-^{n-1}\cap\{m\leq\sum_{i=1}^{n-1}m_i\}}\mathbf{p}\left(t,m_1,\ldots,m_{n-1},m-\sum_{i=1}^{n-1}m_i\right)\mathrm{d}m_1\ldots\mathrm{d}m_{n-1}, \tag{2.6}$$

where $\mathbb{R}_-=(-\infty,0]$.

Thus the mean fitness $\overline{m}(t)$ can be easily connected with these distributions:

$$\overline{m}(t)=\int_{\mathbb{R}^n}m(\mathbf{x})q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}=\sum_{i=1}^{n}\int_{\mathbb{R}_-^n}m_i\mathbf{p}(t,\mathbf{m})\,\mathrm{d}\mathbf{m}=\int_{\mathbb{R}_-}mp(t,m)\,\mathrm{d}m,$$

see also (2.15).

This paper is organized as follows. The results are presented in Section 2.2, and their proofs are presented in Section 2.4. More precisely, Section 2.2.1 is dedicated to the analysis of the time-dependent problem (2.4). We begin with the existence and uniqueness of the solution of the Cauchy problem. We then give an explicit formula for $q(t,\mathbf{x})$ in the particular case of a Gaussian initial distribution of phenotypes $q_0(\mathbf{x})$. Then, we derive a nonlocal degenerate parabolic equation satisfied by $\mathbf{p}(t,\mathbf{m})$, and the equation satisfied by its cumulant generating function (CGF). Solving the equation satisfied by the CGF, we derive an explicit formula for $\overline{m}(t)$. Then, in Section 2.2.2, we study the long time behavior and the stationary states of (2.4). We shall see that the distribution of the fitness components $\mathbf{p}(t,\mathbf{m})$ converges towards a distribution $\mathbf{p}_\infty(\mathbf{m})$ as $t\to+\infty$, and we give an explicit formula for $\mathbf{p}_\infty(\mathbf{m})$. Lastly, in Section 2.2.3 we show that including anisotropy in the models may help to understand experimental trajectories of fitnesses, such as those obtained in the famous experiment of Richard Lenski (Lenski; Rose; Simpson; Tadler 1991; Lenski; Travisano 1994; Wiser; Ribeck; Lenski 2013).

## 2.2. Main results

### 2.2.1. The time-dependent problem

#### Solution of the Cauchy problem associated with Eq. (2.4) for $q(t,\mathrm{x})$

We first show that the Cauchy problem admits a unique solution. We need the following assumption on the initial distribution $q_0$:

$$q_0 \in C^{2+\alpha}(\mathbb{R}^n), \tag{2.7}$$

for some $\alpha \in (0,1)$, that is, $\|q_0\|_{C^{2+\alpha}(\mathbb{R}^n)} < +\infty$. Additionally, we assume that:

$$\int_{\mathbb{R}^n} q_0(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1, \tag{2.8}$$

and there exists a function $g : \mathbb{R}_+ \to \mathbb{R}_+$ (with $\mathbb{R}_+ = [0,+\infty)$) such that:

$$g \text{ is non-increasing}, \;\; 0 \le q_0 \le g(\|\cdot\|) \text{ in } \mathbb{R}^n, \;\; \mathbf{x} \mapsto m(\mathbf{x})g(\|\mathbf{x}\|) \text{ is bounded in } \mathbb{R}^n, \\ \text{and } \int_{\mathbb{R}^n} |m(\mathbf{x})|\, g(\|\mathbf{x}\|)\,\mathrm{d}\mathbf{x} < +\infty. \tag{2.9}$$

These assumptions are made throughout the paper, and are therefore not repeated in the statements of the results below.

We can now state an existence and uniqueness result for the distribution of phenotypes.

**Theorem 13.**
*There exists a unique nonnegative solution $q \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ of:*

$$\begin{cases} \partial_t q(t,\mathbf{x}) &= \displaystyle\sum_{i=1}^{n} \frac{\mu_i^2}{2}\partial_{ii} q(t,\mathbf{x}) + (m(\mathbf{x}) - \overline{m}(t))\, q(t,\mathbf{x}), \quad t \ge 0,\ \mathbf{x} \in \mathbb{R}^n, \\ q(0,\mathbf{x}) &= q_0(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n, \end{cases} \tag{2.10}$$

*such that $q \in L^\infty((0,T)\times\mathbb{R}^n)$ for all $T > 0$, and the function:*

$$t \mapsto \overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})\, q(t,\mathbf{x})\,\mathrm{d}\mathbf{x},$$

*is real-valued and continuous in $\mathbb{R}_+$. Moreover, we have:*

$$\forall\, t \ge 0, \;\; \int_{\mathbb{R}^n} q(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = 1.$$

The next result gives an explicit solution of (2.10) in the particular case where the phenotypes are initially Gaussian-distributed.

**Corollary 1.**
*Assume that the initial distribution of phenotype frequencies is Gaussian, that is,*

$$\forall\, \mathbf{x} \in \mathbb{R}^n,\ q_0(\mathbf{x}) = (2\pi)^{-n/2}\Big(\prod_{i=1}^{n}(s_i^0)^{-1/2}\Big)\ \exp\left(-\sum_{i=1}^{n}\frac{(x_i-\overline{q}_i^0)^2}{2s_i^0}\right), \tag{2.11}$$

*for some parameters $\overline{q}_i^0 \in \mathbb{R}$ and $s_i^0 > 0$. Then the solution $q(t,\mathbf{x})$ of the Cauchy problem (2.10) is Gaussian at all times:*

$$\forall\, t \geq 0,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\ q(t,\mathbf{x}) = (2\pi)^{-n/2}\Big(\prod_{i=1}^{n}(s_i(t))^{-1/2}\Big)\ \exp\left(-\sum_{i=1}^{n}\frac{(x_i-\overline{q}_i(t))^2}{2s_i(t)}\right), \tag{2.12}$$

*with:*

$$\overline{q}_i(t) = \frac{\mu_i \overline{q}_i^0}{\mu_i\cosh(\mu_i t) + s_i^0\sinh(\mu_i t)}, \quad \textit{and} \quad s_i(t) = \mu_i\,\frac{\mu_i\sinh(\mu_i t) + s_i^0\cosh(\mu_i t)}{\mu_i\cosh(\mu_i t) + s_i^0\sinh(\mu_i t)}. \tag{2.13}$$

*Moreover, we have:*

$$\overline{m}(t) = -\sum_{i=1}^{n}\frac{\overline{q}_i^2(t) + s_i(t)}{2}.$$

We also note, in Corollary 1, that the distribution $q(t,\mathbf{x})$ converges to a Gaussian distribution with mean $\overline{q}_\infty = 0$ and variances $s_{\infty,i} = \mu_i$, as $t \to +\infty$.

The determination of an explicit formula for $q(t,\mathbf{x})$ becomes more involved when the initial distribution $q_0$ is not a Gaussian. In this case, we study the equation satisfied by the distribution of the fitness components $\mathbf{p}(t,\mathbf{m})$ (see Introduction) and, as a by-product, we derive an explicit formula for $\overline{m}(t)$.

### A degenerate parabolic PDE satisfied by $\mathbf{p}(t,\mathbf{m})$

Our objective here is to derive an equation for $\mathbf{p}(t,\mathbf{m})$ that only involves linear dependencies with respect to the coefficients $m_i$, and holds for general initial phenotype distributions $q_0$, which may not be Gaussian.

First, we express the distribution of the fitness components $\mathbf{p}(t,\mathbf{m})$ in terms of the distribution of phenotypes $q(t,\mathbf{x})$ given in Theorem 13.

**Proposition 14.** *For all $t \geq 0$ and $\mathbf{m} = (m_1, \dots, m_n) \in (\mathbb{R}_-^*)^n$, there holds:*

$$\mathbf{p}(t, \mathbf{m}) = \frac{2^{-n/2}}{\sqrt{|m_1 \cdots m_n|}} \sum_{\varepsilon=(\varepsilon_1,\dots,\varepsilon_n)\in\{\pm 1\}^n} q(t, \mathbf{x}^\varepsilon(\mathbf{m})), \tag{2.14}$$

*with $\mathbf{x}^\varepsilon(\mathbf{m}) = (\varepsilon_1\sqrt{-2m_1}, \dots, \varepsilon_n\sqrt{-2m_n}) \in \mathbb{R}^n$. Furthermore, we have:*

$$\forall\, t \geq 0, \ \int_{\mathbb{R}_-^n} \mathbf{p}(t, \mathbf{m})\, \mathrm{d}\mathbf{m} = 1, \quad \text{and} \quad \overline{m}(t) = \sum_{i=1}^n \int_{\mathbb{R}_-^n} m_i \mathbf{p}(t, \mathbf{m})\, \mathrm{d}\mathbf{m}, \tag{2.15}$$

*where all above integrals are convergent.*

It also turns out that the expression (2.14) becomes simpler when $q$ satisfies some symmetry properties. In the sequel, for a given function $f \in C(\mathbb{R}^n)$, we define its $\#-$symmetrization $f^\# \in C(\mathbb{R}^n)$ by:

$$\forall\, \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n, \ f^\#(\mathbf{x}) = 2^{-n} \sum_{\varepsilon=(\varepsilon_1,\dots,\varepsilon_n)\in\{\pm 1\}^n} f(\varepsilon_1 x_1, \dots, \varepsilon_n x_n).$$

From the symmetries inherent to (2.4), it is easy to check that, if $q$ is the solution of (2.4) with initial condition $q_0$ satisfying the conditions of Theorem 13, then $q^\#$ is the solution of (2.4) with initial condition $q_0^\#$, and $q(t, \cdot)$ and $q^\#(t, \cdot)$ have the same mean fitness $\overline{m}(t)$ at every time $t \geq 0$. Furthermore, using the expression (2.14), we observe that $\mathbf{p}(t, \mathbf{m})$ can be described in terms of the $\#-$symmetrization of $q(t, \cdot)$:

$$\mathbf{p}(t, \mathbf{m}) = \frac{2^{n/2}}{\sqrt{|m_1 \cdots m_n|}} q^\#(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m})), \tag{2.16}$$

for every $t \geq 0$ and $\mathbf{m} \in (\mathbb{R}_-^*)^n$, with:

$$\mathbf{x}^{\mathbf{1}}(\mathbf{m}) = (\sqrt{-2m_1}, \dots, \sqrt{-2m_n}) \in \mathbb{R}_+^n.$$

This function $\mathbf{p}(t, \mathbf{m})$ satisfies a degenerate parabolic equation:

**Theorem 15.**
*The distribution function of the fitness components* $\mathbf{p}$ *is a classical* $C^{1,2}(\mathbb{R}_+ \times (\mathbb{R}_-^*)^n)$ *solution of:*

$$\partial_t \mathbf{p}(t,\mathbf{m}) = \sum_{i=1}^{n} \mu_i^2 |m_i|\, \partial_{ii}\mathbf{p}(t,\mathbf{m}) - \frac{3}{2}\sum_{i=1}^{n} \mu_i^2\, \partial_{m_i}\mathbf{p}(t,\mathbf{m}) + \left(\sum_{i=1}^{n} m_i - \overline{m}(t)\right) \mathbf{p}(t,\mathbf{m}), \quad (2.17)$$

*for* $t \geq 0$ *and* $\mathbf{m} \in (\mathbb{R}_-^*)^n$*, with initial condition:*

$$\mathbf{p}_0(\mathbf{m}) = \frac{2^{n/2}}{\sqrt{|m_1 \cdots m_n|}}\, q_0^{\#}(\mathbf{x}^{\mathbf{1}}(\mathbf{m})). \quad (2.18)$$

Eq. (2.17) is degenerate in the sense that the operator on the right-hand side is not uniformly elliptic, as at least one of the coefficients $\mu_i^2|m_i|$ in front of the second order differential terms vanishes at the boundary of $(\mathbb{R}_-^*)^n$.

In the isotropic case ($\mu_i = \mu > 0$ for all $i$), it is also possible to derive a scalar equation for the distribution of fitness $p(t,m)$, defined by (2.6).

**Theorem 16.** (ISOTROPIC CASE)
*If* $\mu_i = \mu$ *for all* $1 \leq i \leq n$*, then the fitness distribution* $p(t,m)$ *is a classical* $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}_-^*)$ *solution of:*

$$\partial_t p(t,m) = -\mu^2\, m\, \partial_{mm} p(t,m) + \mu^2 \left(\frac{n}{2} - 2\right) \partial_m p(t,m) + (m - \overline{m}(t))\, p(t,m), \quad (2.19)$$

*for* $t \geq 0$ *and* $m < 0$*, with initial condition,*

$$p_0(m) = \int_{\mathbb{R}_-^{n-1} \cap \{m \leq \sum_{i=1}^{n-1} m_i\}} \mathbf{p}_0\left(m_1, \ldots, m_{n-1}, m - \sum_{i=1}^{n-1} m_i\right) \mathrm{d}m_1 \ldots \mathrm{d}m_{n-1}. \quad (2.20)$$

As expected, Eqs. (2.17) and (2.19) only involve linear combinations of the coefficients $m_i$. This allows us to derive simpler equations satisfied by the generating functions of $\mathbf{p}(t,\mathbf{m})$ and $p(t,m)$.

### Generating functions

We define the *moment generating functions* (MGFs) $M_{\mathbf{p}}$ and $M_p$ of $\mathbf{p}$ and $p$ and their logarithms – the *cumulant generating functions* (CGFs) – $C_{\mathbf{p}}$ and $C_p$ by:

$$M_{\mathbf{p}}(t,\mathbf{z}) = \int_{\mathbb{R}_-^n} e^{\mathbf{z}\cdot\mathbf{m}} \mathbf{p}(t,\mathbf{m})\, \mathrm{d}\mathbf{m}, \quad M_p(t,z) = \int_{-\infty}^{0} e^{zm} p(t,m)\, \mathrm{d}m, \quad (2.21)$$

and:

$$C_{\mathbf{p}}(t,\mathbf{z}) = \log M_{\mathbf{p}}(t,\mathbf{z}), \quad C_p(t,z) = \log M_p(t,z), \tag{2.22}$$

for $t \geq 0$, $\mathbf{z} \in \mathbb{R}_+^n$ and $z \geq 0$. The integrals are well defined because the components $m_i$ are all negative. Furthermore, it follows from (2.15) and the nonnegativity of $\mathbf{p}$ that, for each $t \geq 0$, the functions $M_{\mathbf{p}}(t,\cdot)$ and $C_{\mathbf{p}}(t,\cdot)$ are of class $C^\infty((\mathbb{R}_+^*)^n) \cap C^1(\mathbb{R}_+^n)$, while the functions $M_p(t,\cdot)$ and $C_p(t,\cdot)$ are of class $C^\infty(\mathbb{R}_+^*) \cap C^1(\mathbb{R}_+)$.

The following result gives the equation satisfied by $C_{\mathbf{p}}$.

**Theorem 17.**
*The cumulant generating function $C_{\mathbf{p}}$ of $\mathbf{p}$ is of class $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^n)$ [a] and it solves:*

$$\begin{cases} \partial_t C_{\mathbf{p}}(t,\mathbf{z}) &= A(\mathbf{z}) \cdot \nabla C_{\mathbf{p}}(t,\mathbf{z}) - b(\mathbf{z}) - \overline{m}(t),\ t \geq 0,\ \mathbf{z} \in \mathbb{R}_+^n, \\ C_{\mathbf{p}}(0,\mathbf{z}) &= C_{\mathbf{p}_0}(\mathbf{z}),\ \mathbf{z} \in \mathbb{R}_+^n, \end{cases} \tag{2.23}$$

*where:*

$$A(\mathbf{z}) = (1-\mu_1^2 z_1^2, \ldots, 1-\mu_n^2 z_n^2) \in \mathbb{R}^n,\ b(\mathbf{z}) = \sum_{i=1}^n \frac{\mu_i^2}{2} z_i \in \mathbb{R},\ \overline{m}(t) = \sum_{i=1}^n \partial_{z_i} C_{\mathbf{p}}(t,\mathbf{O}), \tag{2.24}$$

*and $\nabla C_{\mathbf{p}}(t,\mathbf{z})$ denotes the gradient of $C_{\mathbf{p}}$ with respect to the variable $\mathbf{z}$.*

a. This means that the partial derivatives of $C_{\mathbf{p}}$ with respect to the variables $t$ and $\mathbf{z}$ exist and are continuous in $\mathbb{R}_+ \times \mathbb{R}_+^n$.

As a corollary of this theorem, we derive an equation satisfied by $C_p$ in the isotropic case.

**Corollary 2.** (*Isotropic Case*)
*If $\mu_i = \mu$ for all $1 \leq i \leq n$, then the cumulant generating function $C_p$ of the fitness distribution $p$ is a $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+)$ solution of:*

$$\begin{cases} \partial_t C_p(t,z) &= (1-\mu^2 z^2)\partial_z C_p(t,z) - \frac{n}{2}\mu^2 z - \overline{m}(t),\ t \geq 0,\ z \in \mathbb{R}_+, \\ C_p(0,z) &= C_{p_0}(z),\ z \in \mathbb{R}_+, \end{cases} \tag{2.25}$$

*where $p_0$ is the initial fitness distribution given in (2.20), and $\overline{m}(t) = \partial_z C_p(t,0)$.*

Note that Eq. (2.25) is directly obtained as a by-product of (2.23), without using Eq. (2.19) satisfied by $p(t,m)$. Each of these two equations (2.23) and (2.25) has a unique solution which can be computed explicitly, leading to a formula for $\overline{m}(t)$. We refer to Section 2.2.3 for an application of this result.

In the Introduction, we mentioned that the coefficients $\mu_i^2$ can be interpreted as

the product between the mutation rate $U$ and the variance $\lambda_i$ at the $i$-th trait. In the isotropic case, $\mu_i^2 = \mu^2 = U\,\lambda$. Thus we have retrieved the equation mentioned in Appendix E in Martin; Roques 2016 (*cf.* Eq. (E5)).

The last two results of this section provide some explicit expressions of $C_{\mathbf{p}}(t, \mathbf{z})$ and $C_p(t, z)$ when $\mathbf{z}$ and $z$ are close enough to $\mathbf{O}$ and $0$.

**Proposition 18.** *The cumulant generating function $C_{\mathbf{p}}$ of $\mathbf{p}$ is given by, for all $t \geq 0$ and $\mathbf{z} = (z_1, \ldots, z_n) \in [0, 1/\mu_1) \times \cdots \times [0, 1/\mu_n)$,*

$$C_{\mathbf{p}}(t, \mathbf{z}) = \frac{1}{2} \sum_{i=1}^{n} \log \left[ \frac{\cosh(\mu_i t)\, \cosh(\operatorname{atanh}(\mu_i z_i))}{\cosh(\mu_i t + \operatorname{atanh}(\mu_i z_i))} \right] + C_{\mathbf{p}_0}(\psi(t, \mathbf{z})) - C_{\mathbf{p}_0}(\psi(t, \mathbf{O})), \quad (2.26)$$

*with:*

$$\psi(t, \mathbf{z}) = (\psi_1(t, \mathbf{z}), \ldots, \psi_n(t, \mathbf{z})), \quad \text{and} \quad \psi_j(t, \mathbf{z}) = \frac{1}{\mu_j} \tanh\left(\mu_j t + \operatorname{atanh}(\mu_j z_j)\right). \quad (2.27)$$

*Moreover, for all $t \geq 0$, we have:*

$$\overline{m}(t) = \sum_{i=1}^{n} \left[ (1 - \tanh^2(\mu_i t))\, \partial_{z_i} C_{\mathbf{p}_0}(\psi(t, \mathbf{O})) - \frac{\mu_i}{2} \tanh(\mu_i t) \right]. \quad (2.28)$$

**Corollary 3.** (*Isotropic case*)
*If $\mu_i = \mu$ for all $1 \leq i \leq n$, then the cumulant generating function $C_p$ of $p$ is given by:*

$$C_p(t, z) = \frac{n}{2} \log \left( \frac{\cosh(\mu\, t) \cosh(\operatorname{atanh}(\mu\, z))}{\cosh(\mu\, t + \operatorname{atanh}(\mu\, z))} \right) + C_{p_0}(\varphi(t, z)) - C_{p_0}(\varphi(t, 0)), \quad (2.29)$$

*for $t \geq 0$ and $0 \leq z < 1/\mu$, with $\varphi(t, z) = (1/\mu) \tanh\left(\mu\, t + \operatorname{atanh}(\mu\, z)\right)$. Moreover, we have:*

$$\overline{m}(t) = (1 - \tanh^2(\mu t))\, \partial_z C_{p_0}\left( \frac{\tanh(\mu t)}{\mu} \right) - \frac{n\mu}{2} \tanh(\mu t). \quad (2.30)$$

## 2.2.2. Long time behavior and stationary states

We are here interested in the long time behaviour of the solutions of (2.4) and (2.17). We begin with a result on the convergence of the solution of (2.17) at $t \to +\infty$.

**Theorem 19.**
*Let* $\mathbf{p}$ *and* $\overline{m}(t)$ *be as in the previous section. Then:*

$(i)$ $\mathbf{p}(t,\cdot)$ *weakly converges in* $(\mathbb{R}_-^*)^n$ *to* $\mathbf{p}_\infty$ *as* $t \to +\infty$*, where:*

$$\mathbf{p}_\infty(\mathbf{m}) = \frac{1}{\pi^{n/2}\sqrt{\mu_1 \cdots \mu_n}\sqrt{|m_1 \cdots m_n|}} \exp\left(\sum_{i=1}^n \frac{m_i}{\mu_i}\right) \quad \text{for all } m \in (\mathbb{R}_-^*)^n, \tag{2.31}$$

*in the sense that* $\int_{(\mathbb{R}_-^*)^n} \mathbf{p}(t,\mathbf{m})\,\phi(\mathbf{m})\,\mathrm{d}\mathbf{m} \to \int_{(\mathbb{R}_-^*)^n} \mathbf{p}_\infty(\mathbf{m})\,\phi(\mathbf{m})\,\mathrm{d}\mathbf{m}$ *as* $t \to +\infty$ *for every test function* $\phi \in C_c^\infty((\mathbb{R}_-^*)^n)$*;*

$(ii)$ $\overline{m}(t) \to \overline{m}_\infty = -\sum_{i=1}^n \frac{\mu_i}{2}$ *as* $t \to +\infty$ *and* $\overline{m}_\infty = \sum_{i=1}^n \int_{\mathbb{R}_-^n} m_i \mathbf{p}_\infty(\mathbf{m})\mathrm{d}\mathbf{m}$*;*

$(iii)$ *the function* $\mathbf{p}_\infty$ *is a classical* $C^2((\mathbb{R}_-^*)^n)$ *solution of:*

$$0 = \sum_{i=1}^n \mu_i^2 |m_i| \partial_{ii}\mathbf{p}_\infty(\mathbf{m}) - \frac{3}{2}\sum_{i=1}^n \mu_i^2 \partial_{m_i}\mathbf{p}_\infty(\mathbf{m}) + \left(\sum_{i=1}^n m_i - \overline{m}_\infty\right)\mathbf{p}_\infty(\mathbf{m}). \tag{2.32}$$

In the isotropic case, we retrieve the result of Martin; Roques 2016 in the WSSM case (Weak Selection and Strong Mutation), which says that the fitnesses are asymptotically distributed according to the symmetrized Gamma distribution $-\Gamma(n/2,\mu)$, with $\mu = \sqrt{U\lambda}$:

**Corollary 4.** (*Isotropic case*)
*If* $\mu_i = \mu$ *for all* $1 \le i \le n$*, then* $p(t,\cdot)$ *weakly converges in* $\mathbb{R}_-^*$ *to* $p_\infty$ *as* $t \to +\infty$*, where:*

$$p_\infty(m) = \frac{|m|^{\frac{n}{2}-1}}{\Gamma(n/2)\,\mu^{n/2}} \exp\left(\frac{m}{\mu}\right) \quad \text{for all } m < 0,$$

*and* $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}\mathrm{d}t$ *is the standard Gamma function.*

Thanks to the previous two results, we get the asymptotic behavior of the phenotype distribution in the symmetric case.

**Corollary 5.**
*If $q_0$ is #-symmetric in the sense that $q_0 = q_0^{\#}$, then $q(t,\cdot) = q^{\#}(t,\cdot)$ weakly converges to $q_\infty$ as $t \to +\infty$ on $\mathbb{R}^n$ where, for all $x \in \mathbb{R}^n$,*

$$q_\infty(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\mu_1 \cdots \mu_n}} \exp\left(-\sum_{i=1}^{n} \frac{x_i^2}{2\mu_i}\right), \tag{2.33}$$

*and $\overline{m}(t) \to \overline{m}_\infty = \int_{\mathbb{R}^n} m(\mathbf{x})\, q_\infty(\mathbf{x})\, \mathrm{d}\mathbf{x}$ as $t \to +\infty$.*

We note that $q_\infty$ is a classical positive stationary state of (2.4), *i.e.*, it satisfies the following equation:

$$\sum_{i=1}^{n} \frac{\mu_i^2}{2}\, \partial_{ii} q_\infty(\mathbf{x}) + m(\mathbf{x})\, q_\infty(\mathbf{x}) = \overline{m}_\infty\, q_\infty(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

Thus, the distribution $q(t,\mathbf{x})$ and the mean fitness $\overline{m}(t)$ converge to the principal eigenfunction (resp. eigenvalue) of the operator $\mathcal{H}_n = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\partial_{ii} + m(\mathbf{x})$.

In the 1D case ($x \in \mathbb{R}$), the results in Alfaro; Veruete 2018 imply that $q(t,x)$ can be written in terms of the eigenelements $(\lambda_k, \phi_k)$ of the operator $\mathcal{H}_1 = -\frac{\mu^2}{2}\,\mathrm{d}^2/\mathrm{d}x^2 - m(x)$. Namely,

$$q(t,x) = K(t) \sum_{k=0}^{\infty} (q_0, \phi_k)_{L^2(\mathbb{R})}\ \phi_k(x)\, e^{-\lambda_k t},$$

where $K(t)$ is such that $q(t,\cdot)$ sums to 1, and $(\cdot,\cdot)_{L^2(\mathbb{R})}$ is the standard scalar product in $L^2(\mathbb{R})$. As a corollary, the authors of Alfaro; Veruete 2018 obtained the convergence of the distribution $q(t,x)$ and of the mean fitness $\overline{m}(t)$ to the principal eigenfunction (resp. eigenvalue) of $\mathcal{H}_1$, namely $\phi_0$ (resp. $\lambda_0$). Thus, in the 1D case, their result is stronger than that of Corollary 5, as the convergence of $q(t,x)$ occurs in $L^p(\mathbb{R}^n)$, for all $1 \le p \le +\infty$, and does not require $q_0$ to be #-symmetric. Based on the spectral properties of the operator $\mathcal{H}_n$ (Hislop; Sigal 2012), we conjecture that their results can be extended to the anisotropic multidimensional framework considered here, meaning that the convergence result in Corollary 5 remains true when the initial condition $q_0$ is not #-symmetric.

### 2.2.3. Effect of anisotropy: numerical computations and connection with *Escherichia coli* long-term evolution experiment

The objective of this section is to illustrate the importance of taking anisotropy into account when modelling adaptation trajectories. Isotropic models (Martin;

Roques 2016) lead to regularly saturating trajectories of $\overline{m}(t)$ with a plateau, *i.e.* a single 'epoch'. Here, we show that, in the presence of anisotropy, the trajectory of $\overline{m}(t)$ can exhibit several plateaus before reaching a stable level close to $\overline{m}_\infty$. Thus, the dynamics of adaptation can show several evolutionary 'epochs', as those observed in the *E. coli* long-term evolution experiment (Good; Desai 2015), corresponding to different time-scales at which adaptation occurs.

For the sake of simplicity of the computations, and although the existence and uniqueness results of Section 2.2.1 were only obtained with continuous initial distributions of phenotypes, we assume a Dirac initial distribution of the phenotypes. Namely, we assume that $q_0 = \delta_{\mathbf{x}_0}$ with $\mathbf{x}_0 = (x_{0,1}, \dots, x_{0,n}) \in \mathbb{R}^n$. This corresponds to an initially clonal population. This hypothesis implies that the initial distribution of the fitness components and the initial CGF are respectively given by:

$$\mathbf{p}_0 = \delta_{\mathbf{m}(\mathbf{x}_0)}, \quad \text{and} \quad C_{\mathbf{p}_0}(\mathbf{z}) = \exp\left(\mathbf{m}(\mathbf{x}_0) \cdot \mathbf{z}\right),$$

with $\mathbf{m}(\mathbf{x}_0) = (-x_{0,1}^2/2, \dots, -x_{0,n}^2/2)$. In this case, the expression (2.28) in Proposition 18 simplifies to:

$$\overline{m}(t) = \sum_{i=1}^{n} \left( \frac{x_{0,i}^2}{2} \left( \tanh^2(\mu_i t) - 1 \right) - \frac{\mu_i}{2} \tanh(\mu_i t) \right). \tag{2.34}$$

### Trajectory of adaptation in the presence of anisotropy: an illustrative example

We take $n = 3$ and $\mu_1 > \mu_2 > \mu_3$. The corresponding trajectory of $\overline{m}(t)$ is depicted in Fig. 2.1. After a brief initial decay which was already observed in the isotropic case (Martin; Roques 2016), $\overline{m}(t)$ rapidly increases and reaches a first plateau (of value close to $\overline{m}_{\infty,1} := -x_{0,2}^2/2 - x_{0,3}^2/2 - \mu_1/2$). Then, $\overline{m}(t)$ rapidly increases again to reach a second plateau (of value close to $\overline{m}_{\infty,2} := -x_{0,3}^2/2 - \mu_1/2 - \mu_2/2$). Finally, $\overline{m}(t)$ increases again and stabilises around $\overline{m}_\infty = -\mu_1/2 - \mu_2/2 - \mu_3/2$. Interestingly, although the ultimate value $\overline{m}_\infty$ does not depend on the initial phenotype, the intermediate plateaus depend on $\mathbf{x}_0$. Their values approximately correspond to the fitnesses associated with the successive projections of $\mathbf{x}_0$ on the hyperplanes $\{x_1 = 0\}$ and $\{x_1 = x_2 = 0\}$ minus the mutation load (we recall that the optimal phenotype was fixed at $\mathbf{x} = (0, \dots, 0)$, see Introduction).

More generally speaking, we prove in Section 2.4.5 that, for $n \geq 2$ and $\mu_1 > \dots > \mu_n$, the trajectory exhibits $(n-1)$ plateaus (before the final one of value $\overline{m}_\infty$), of respective values:

$$\overline{m} \approx \overline{m}_{\infty,k} := -\sum_{i=k+1}^{n} \frac{x_{0,i}^2}{2} - \sum_{i=1}^{k} \frac{\mu_i}{2}, \tag{2.35}$$

for $k = 1, \dots, (n-1)$. More precisely, we obtain the following proposition.

**Proposition 20.** *Let $\overline{m}(t)$ and $\overline{m}_{\infty,k}$ be defined by (2.34) and (2.35), respectively. Given $\mathbf{x}_0 = (x_{0,1}, \dots, x_{0,n}) \in \mathbb{R}^n$, $T > 0$, $\varepsilon > 0$ and $\mu_1 > 0$, there exist $\mu_2, \dots, \mu_n > 0$ such that, for each $k \in [\![1, n-1]\!]$, the set:*

$$\Big\{ t \geq 0,\ |\overline{m}(t) - \overline{m}_{\infty,k}| \leq \varepsilon \Big\},$$

*contains an interval of length at least equal to $T$.*

Note that the values $\overline{m}_{\infty,k}$ may not be ordered by increasing order, depending on the parameter values, possibly leading to nonmonotone trajectories of $\overline{m}(t)$. The plateaus are more visible when the $\mu_i$'s have different orders of magnitude. More precisely, we show in Section 2.4.5 that, given $T > 0$, $\overline{m}(t)$ remains around each plateau of value $\overline{m}_{\infty,k}$ at least during a time interval of duration $T$, for a good choice of the parameters $\mu_i$.

**Remark 1.** *The proof of Proposition 20 does not lead to explicit expressions for the intervals at which the plateaus occur. In the particular case $n = 2$, we can obtain a more precise characterization of these intervals, see Section 2.4.5. In particular, taking $\mu_1 = 1$, $\mu_2 = 10^{-k}$ for some $k \geq 1$ and $x_{0,1} = x_{0,2} = 1$, we show that:*

*(i)* $\overline{m}'(t) \to -1/2 - 10^{-2\,k}/2$ *as* $t \to 0$*;*

*(ii) at* $t = t_0 = \ln(3)$, $\overline{m}'(t_0) > 27/250 - \mu_2/2 > 0.1 - 10^{-k}/2$*;*

*(iii) for all* $t$ *in the interval* $I := \left(\ln\left(2 \cdot 10^k\right), 10^{k/2}\right)$, $\overline{m}'(t)$ *is of order* $10^{-3\,k/2}$*;*

*(iv) at* $t = t_1 = 10^k \ln[(\sqrt{2} + \sqrt{6})/2] > 10^{k/2}$, $\overline{m}'(t)$ *is of order* $10^{-k}$*.*

*At larger times, we already know that $\overline{m}'(t) \to 0$. Finally, this shows that $\overline{m}(t)$ remains stable within the interval $I$, corresponding to a part of the plateau, relatively to the times $t_0$ (before the interval $I$) and $t_1$ (after the interval $I$).*

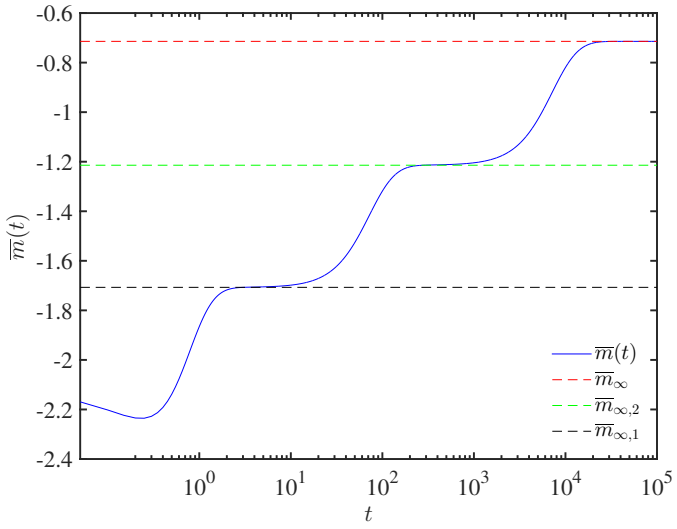



Figure 2.1. – **Trajectory of adaptation in the presence of anisotropy, with $n = 3$.** The function $\overline{m}(t)$ is given by formula (2.34), and the approximations of the values of the intermediate plateaus are $\overline{m}_{\infty,1} = -x_{0,2}^2/2 - x_{0,3}^2/2 - \mu_1/2$ and $\overline{m}_{\infty,2} = -x_{0,3}^2/2 - \mu_1/2 - \mu_2/2$. The mutational parameters are $\mu_1 = \sqrt{2}$, $\mu_2 = \sqrt{2} \cdot 10^{-2}$ and $\mu_3 = \sqrt{2} \cdot 10^{-4}$. The other parameter values are $\mathbf{x}_0 = (3/2, 1, 1)$.

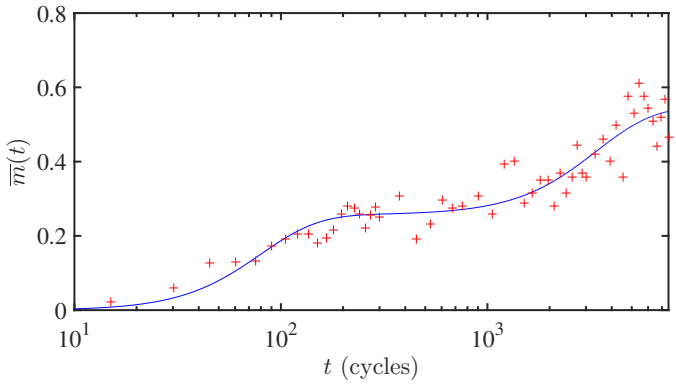
### Long term evolution experiment with *Escherichia coli*

The long term evolution experiment (LTEE) has been carried by Lenski and his collaborators since 1988 (Lenski; Rose; Simpson; Tadler 1991). Twelve populations of *E. coli* have been founded from a single common ancestor, and are still evolving after more than 70,000 generations. The fitness evolved rapidly during the first 2000 generations, and then remained nearly static between generations $5000$ and $10,000$ (Lenski; Travisano 1994), which would at least phenomenologically advocate for the existence of a phenotype optimum. However, more recent data (after generation 10,000) indicate that the mean fitness seems to increase without bounds (Wiser; Ribeck; Lenski 2013). Our goal here is not to propose a new explanation of the LTEE data, but simply to check whether the anisotropic model (2.4) leads to a better fit than an isotropic model.

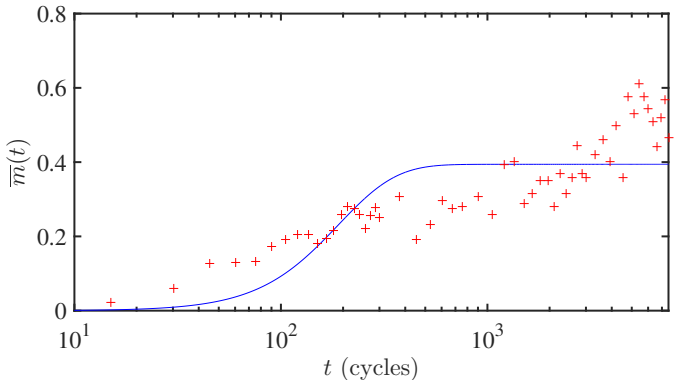
The interpretation of the fitness data from the LTEE is quite subtle (see the comments in Richard Lenski's Experimental Evolution website `http://myxo.css.msu.edu/ecoli/srvsrf.html`). For the sake of simplicity, as our objective is to check if the trajectories given by (2.34) can be qualitatively consistent with the fitness data from the LTEE, we make the following simplifying assumptions: (1) a time unit $t$ in our model corresponds to 1 cycle of the experiment (1 day), corresponding to $\approx 6.64$ generations (Lenski; Rose; Simpson; Tadler 1991); (2) the link between $\overline{m}(t)$ and the mean fitness values given by the LTEE data, which are of the form $w(\text{gen})$ (gen is the generation number and $w$ the Darwinian

mean fitness, which is related to the Malthusian fitness through an exponential) is made by assuming that $\overline{m}(t+2) - \overline{m}(2) = \ln[w(6.64t)]$. Thus, we assume (arbitrarily) that $2$ cycles were necessary to build the founding colony from the single ancestor (hence the term $t+2$). Additionally, in the data, the fitness is a relative fitness against the ancestor, which implies that $w(0) = 1$; this is why the quantity $\overline{m}(2)$ was subtracted to $\overline{m}(t+2)$. As mentioned above, the data are available for 12 populations. Here, we only use the data from one population (the population Ara-1, see Wiser; Ribeck; Lenski 2013), for which measurements were obtained at 100-generation intervals during the first 2000 generations, and then at 500-generation intervals.

We carried out a fit (using Matlab® Curve Fitting Toolbox®, with a nonlinear least squares method) with the function $f(t) = \overline{m}(t+2) - \overline{m}(2)$. For the sake of simplicity, we assumed a two-dimensional phenotypic space ($n = 2$). The only parameters to estimate are $\mu_1,\ \mu_2$ and $\mathbf{x}_0$. We compared the result of this fit with a fit of the isotropic model ($\mu_1 = \mu_2$). The results are depicted in Fig. 2.2. A graphical analysis shows that the anisotropic model gives a far better fit. This is confirmed by the adjusted $R^2$: $0.89$ for the anisotropic model *v*ersus $0.57$ for the isotropic model ($R^2$=1 indicates that the fitted model explains all of the variability in the data). In the anisotropic model, the fitted mutational parameters have two different orders of magnitude: $\mu_1 = 1.3 \cdot 10^{-2}$ and $\mu_2 = 3.0 \cdot 10^{-4}$. This leads to a first plateau until about 1000 cycles ($6640$ generations) followed by a second increase of the fitness. As expected, the isotropic model cannot explain this type of pattern.




(a) Anisotropic model




(b) Isotropic model

Figure 2.2. – **Trajectory of adaptation, anisotropic and isotropic model *versus* LTEE data.** The function $f(t) = \overline{m}(t+2) - \overline{m}(2)$, with $\overline{m}(t)$ given by (2.34) was fitted to the LTEE data. The functions $f(t)$ with the parameter values corresponding to the best fit are depicted in blue. The red crosses correspond to the LTEE data (population Ara-1, $\ln[w(6.64t)]$, with $w$ the original data). The values leading to the best fit are (a, anisotropic model) $\mathbf{x}_0 = (0.73, 0.76)$, $\mu_1 = 1.3 \cdot 10^{-2}$ and $\mu_2 = 3.0 \cdot 10^{-4}$; (b, isotropic model) $\mathbf{x}_0 = (0.44, 0.78)$ and $\mu_1 = \mu_2 = 5.3 \cdot 10^{-3}$.

## 2.3. Discussion

We considered a natural $n$-dimensional extension of the standard diffusive 'replicator-mutator' equation, to describe the dynamics of a phenotype distribution under anisotropic mutation and selection effects, and in the presence of a phenotype optimum. We proved that the corresponding Cauchy problem was well-posed (existence and uniqueness of the solution) and we proposed a new and general framework to the study of the quantitative behaviour of the solution $q(t, \mathbf{x})$. This framework enabled us to derive a formula for the mean fitness in

the population at all times (Eq. (2.28)), an important quantity to describe the dynamics of adaptation.

The case of an initially Gaussian distribution $q_0$ of the phenotypes is simpler, as the phenotype distribution $q(t, \mathbf{x})$ remains Gaussian (though anisotropic) at all times. However, when $q_0$ is not Gaussian, this result obviously breaks down. In this case, the method that we proposed consists in two steps: (1) to derive a degenerate nonlocal parabolic equation satisfied by the distribution $\mathbf{p}(t, \mathbf{m})$ of the fitness components, *i.e.*, the joint distribution of the $m_i = -x_i^2/2$. This equation is degenerate at $0$, but has the advantage of being linear with respect to $\mathbf{m}$; (2) to derive a nonlocal transport equation satisfied by the cumulant generating function of $\mathbf{p}(t, \mathbf{m})$. This last equation can be solved analytically, and its solution leads to explicit formulae for all the moments of $\mathbf{p}$.

Conversely, the methods that are developed in this paper could be applied to solve more general degenerate parabolic PDEs. The idea would be firstly to transform the degenerate equation into a non-degenerate equation, through a change of function of the type (2.16) (rewriting $q^\#$ in terms of $\mathbf{p}$) to get an existence result, and secondly to consider PDEs satisfied by moment generating functions in order to obtain uniqueness and more quantitative properties of the solution of the degenerate PDE.

A natural idea to solve (2.4) could be to consider directly the cumulant generating function associated with $q$:

$$C_q(t, \mathbf{z}) = \ln\left(\int_{\mathbb{R}^n} e^{\mathbf{z}\cdot\mathbf{x}} q(t, \mathbf{x})\, \mathrm{d}\mathbf{x}\right). \tag{2.36}$$

One would then have to see for which $\mathbf{z}$ this quantity makes sense, since now one integrates with respect to $\mathbf{x} \in \mathbb{R}^n$. Due to the nonlinear term $m(\mathbf{x})$ in (2.4), the equation satisfied by $C_q$ would then be a nonlocal second-order viscous Hamilton-Jacobi type equation:

$$\partial_t C_q(t, \mathbf{z}) = -\frac{1}{2}\,\Delta C_q(t, \mathbf{z}) - \frac{1}{2}\,\|\nabla C_q(t, \mathbf{z})\|^2 + \sum_{i=1}^{n} \frac{\mu_i^2}{2}\, z_i^2 - \overline{m}(t),$$

whereas Eq. (2.23) for $C_{\mathbf{p}}$ was a nonlocal first-order transport equation.

From an applied perspective, the results of Section 2.2.3 illustrate the importance of taking anisotropy into account, as it can open up further explanations of experimental data. In sharp contrast with the known results based on isotropic models, our results show that the trajectory of adaptation may exhibit $(n-1)$ plateaus before it reaches the final one. In particular, using our analytic formulae for the dynamics of $\overline{m}(t)$, we obtained a very good fit of one of the most famous experimental dataset in experimental evolution, for which several evolutionary 'epochs' had already been observed (Good; Desai 2015). This suggests that the FGM, in its anisotropic form, can generate a diversity of behaviours that may reconcile the various outcomes observed in long *versus* short term experimental

evolution. However, whether this versatility may lead to overparameterization remains an open statistical question.

### Model extensions

The simple formalism of model (2.4) can easily be extended to more general situations. A first generalization corresponds to the case where the environment changes over time, which would be described by a moving optimum changing smoothly with some external environmental factor (*e.g.*, a drug dose, a temperature, *etc.*). Such a moving optimum could be taken into account by setting:

$$m(t, \mathbf{x}) = -\frac{\|\mathbf{x} - \mathbf{O}(t)\|^2}{2},$$

with $\mathbf{O}(t) \in \mathbb{R}^n$ the optimum at time $t$, instead of (2.5). In this case, we conjecture that Gaussian solutions of the form (2.12) still exist, suggesting that most our results could be extended to this situation.

Another generalization would be to explicitly consider birth-dependent mutation. Indeed, in many cases, mutations occur at birth. For example, in unicellulars undergoing a birth-death process, divisions without mutation occur at per capita rate $b\,(1-u)$ and divisions with mutation occur at rate $b\,u$ where $u$ is a per division mutation probability. If the birth rate varies across genotypes or over time, then so does the mutation rate, and we can no longer consider a constant per capita mutation rate $U$. First, let us note that this complication can be ignored in several biologically relevant situations:

- In multicellulars, evolving over a discrete time demographic process (*e.g.* annual plants), the process of birth and mutations are decorrelated as what determines the mutation rate is the number of divisions in the germline, which does not a priori correlate with fitness (fecundity or survival). Hence the model with constant $U$ can be applied here, our model being a continuous time approximation to the exact discrete time model.
- In many microbial evolution experiments (like the LTEE presented in Section 2.2.3), cells undergo rounds of a pure birth process separated by regular dilutions by a constant factor. In this situation, if cells grow to a constant density, then remain 'dormant' before dilution (so-called stationary phase in microbiology) then each cycle corresponds to a constant number of divisions. With a faster birth rate, these divisions simply occur earlier on during the cycle, but are not more numerous. This process becomes approximately equivalent to a discrete time life cycle at the scale of growth dilution cycles. The mutation rate per cycle is then also constant (independent of the birth rate). Note that this heuristics is only approximate, as it ignores variation in the birth rate between genotypes, relative to its mean, among co-segregating lineages, over one growth cycle. However, that this heuristics applies to

the LTEE may explain why models that ignore the coupling of birth and mutation rate can provide a good fit to data from the LTEE. More detailed individual-based stochastic models for the LTEE can be found in *e.g.* Baake; Casanova; Probst; Wakolbinger 2019; Casanova; Kurt; Wakolbinger; Yuan 2016.

- Under 'viability selection', selection only acts to reduce the death rate $d$, while the birth rate $b$ shows limited evolution: mutation rates are then also roughly constant over time ($U \approx \bar{b}\, u$, where $\bar{b}$ is the mean birth rate, stable over time). This happens if (i) the environmental challenge mainly affects the death rate and (ii) birth and death rates are uncorrelated so that evolution of $d$ has little impact on $b$. In the context of phenotype-fitness landscape models, one situation where this applies is when fitness only depends on phenotype through the death rate, $m(\mathbf{x}) = b - d(\mathbf{x})$: the mutation rate is then $U = b\, u$, stable over time and across genotypes/phenotypes.

Overall there are several situations where one could ignore the connection between birth and mutation rates. Yet, in a more general context where the birth rate $b(\mathbf{x}) \geq 0$ does depend on phenotype $\mathbf{x}$, then the mutation rate *per capita* per unit time also depends on $\mathbf{x}$: $U(\mathbf{x}) = u\, b(\mathbf{x})$, with $u$ the probability of mutation per birth event. The model (2.4) can then be extended, to take into account this effect:

$$\partial_t q(t,\mathbf{x}) = \sum_{i=1}^{n} \frac{1}{2}\, \partial_{ii}[\mu_i^2(\mathbf{x}) q(t,\mathbf{x})] + (m(\mathbf{x}) - \overline{m}(t))\, q(t,\mathbf{x}),\ t > 0,\ \mathbf{x} \in \mathbb{R}^n,$$

with $\mu_i^2(\mathbf{x}) = \lambda_i\, U(\mathbf{x}) = \lambda_i\, u\, b(\mathbf{x})$, with $\lambda_i$ the mutational variance at each trait, see Appendix. However, the mathematical results of our paper cannot be straightforwardly extended to this case, even in the isotropic case. In particular, as soon as $b(\mathbf{x})$ is not constant, this equation has no Gaussian solution. Still, characterization of the steady states should be possible as a principal eigenvalue problem.

## 2.4. Proofs

This section is devoted to the proofs of the results announced in Section 2.2. Section 2.4.1 is concerned with problem (2.10) satisfied by the distribution of phenotype frequencies $q(t,\mathbf{x})$, while Section 2.4.2 deals with the fitness frequencies $\mathbf{p}(t,\mathbf{m})$ and the proofs of Proposition 14 and Theorem 15. In Section 2.4.3, we carry out the proofs of Theorem 17 and Proposition 18 on the cumulant generating functions, and their corollaries (Theorem 16 and Corollaries 2 and 3) in the isotropic case. Lastly, Sections 2.4.4 and 2.4.5 are concerned with the stationary states and the existence of plateaus for the mean fitness.

### 2.4.1. Proofs of Theorem 13 and Corollary 1 on the Cauchy problem (2.10)

Before considering the nonlocal problem (2.10), we begin with the local problem:

$$\left\{\begin{array}{rcl} \partial_t v(t,\mathbf{x}) & = & \displaystyle\sum_{i=1}^{n} \frac{\mu_i^2}{2}\partial_{ii} v(t,\mathbf{x}) + m(\mathbf{x})\, v(t,\mathbf{x}),\ t \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \\ v(0,\mathbf{x}) & = & q_0(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n, \end{array}\right. \tag{2.37}$$

where $q_0$ satisfies (2.7)-(2.9). As the fitness function:

$$m(\mathbf{x}) = -\frac{\|\mathbf{x}\|^2}{2},$$

is unbounded, standard parabolic theory with bounded coefficients does not apply. However, some properties of (2.37) can be for instance be found in Aronson; Besala 1967; Lunardi 1998; Protter; Weinberger 1967, which in particular lead to the following result.

**Theorem 21.**
*The problem (2.37) admits a unique bounded solution $v \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$. Additionally, we have:*

$$\forall T > 0,\ \exists\, S > 0,\ \forall\, t \in [0,T],\ \ \|v(t,\cdot)\|_{C^{2+\alpha}(\mathbb{R}^n)} \leq S\, \|q_0\|_{C^{2+\alpha}(\mathbb{R}^n)}, \tag{2.38}$$

*and:*

$$0 < v(t,\mathbf{x}) < (K_t * q_0)(\mathbf{x}),\ \textit{for all}\ (t,\mathbf{x}) \in \mathbb{R}_+^* \times \mathbb{R}^n, \tag{2.39}$$

*with:*

$$K_t(\mathbf{x}) = \frac{1}{(2\pi t)^{n/2}\,\mu_1 \cdots \mu_n} \exp\left[-\sum_{i=1}^{n} \frac{x_i^2}{2\mu_i^2 t}\right]. \tag{2.40}$$

*Proof.* Let us fix a time $T > 0$. Theorem 2 in Lunardi 1998 implies that (2.37) admits a unique bounded solution $v \in C^{1,2}([0,T] \times \mathbb{R}^n)$ and this solution satisfies (2.38). Theorem III in Aronson; Besala 1967 further implies that this solution is nonnegative. As $T$ was chosen arbitrarily, these existence, uniqueness and nonnegativity results extend to $t \in (0, +\infty)$, with local boundedness in $t$.

Let us set $h(t,\mathbf{x}) := (K_t * q_0)(\mathbf{x})$ for $t > 0$, and $h(0,\mathbf{x}) = q_0(\mathbf{x})$. The function $h$ satisfies:

$$\partial_t h(t,\mathbf{x}) = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\partial_{ii} h(t,\mathbf{x}),$$

for all $t > 0$ and $\mathbf{x} \in \mathbb{R}^n$. Let $\psi(t, \mathbf{x}) := v(t, \mathbf{x}) - h(t, \mathbf{x})$. We see that, for all $t > 0$ and $\mathbf{x} \in \mathbb{R}^n$,

$$\partial_t \psi(t, \mathbf{x}) - \sum_{i=1}^{n} \frac{\mu_i^2}{2} \partial_{ii} \psi(t, \mathbf{x}) = m(\mathbf{x}) v(t, \mathbf{x}) \leq 0, \tag{2.41}$$

and $\psi(0, \mathbf{x}) = 0$. By the Phragmèn-Linderlöf principle (Protter; Weinberger 1967, Theorem 10, Chapter 3), we get that $\psi \leq 0$ in $\mathbb{R}_+ \times \mathbb{R}^n$, *i.e.*, $v(t, \mathbf{x}) \leq h(t, \mathbf{x}) = (K_t * q_0)(\mathbf{x})$. We therefore infer that $0 \leq v(t, \mathbf{x}) \leq (K_t * q_0)(\mathbf{x})$ in $(0, +\infty) \times \mathbb{R}^n$. Since $q_0$ is bounded, this also implies that $v$ is bounded in $\mathbb{R}_+ \times \mathbb{R}^n$.

By the standard strong parabolic maximum principle, we conclude that the first inequality is strict, *i.e.*, $0 < v(t, \mathbf{x})$, for all $(t, \mathbf{x}) \in (0, +\infty) \times \mathbb{R}^n$, since $v(0, \cdot) = q_0$ is continuous, nonnegative and not identically equal to $0$. Furthermore, since the inequality in (2.41) is then strict for all $(t, \mathbf{x}) \in (0, +\infty) \times (\mathbb{R}^n \setminus \{\mathbf{O}\})$, we get that $\psi(t, \mathbf{x}) < 0$, *i.e.*, $v(t, \mathbf{x}) < (K_t * q_0)(\mathbf{x})$, for all $(t, \mathbf{x}) \in (0, +\infty) \times \mathbb{R}^n$. □

In order to connect (2.37) and (2.10), we need the following lemma.

**Lemma 1.** *The function:*

$$t \mapsto \overline{m}_v(t) := \int_{\mathbb{R}^n} m(\mathbf{x})\, v(t, \mathbf{x})\, \mathrm{d}\mathbf{x}, \tag{2.42}$$

*is real-valued and continuous in $\mathbb{R}_+$ and, for every $t \geq 0$, there holds:*

$$1 + \int_0^t \overline{m}_v(s)\, \mathrm{d}s = 1 + \int_0^t \int_{\mathbb{R}^n} m(\mathbf{x}) v(s, \mathbf{x})\, \mathrm{d}\mathbf{x}\, \mathrm{d}s = \int_{\mathbb{R}^n} v(t, \mathbf{x})\, \mathrm{d}\mathbf{x} > 0. \tag{2.43}$$

*Proof.* First of all, denote:

$$\underline{\mu} := \min(\mu_1, \ldots, \mu_n) > 0 \quad \text{and} \quad \overline{\mu} := \max(\mu_1, \ldots, \mu_n) > 0. \tag{2.44}$$

It follows from the assumptions on $q_0$ that $\overline{m}_v(0)$ is a nonpositive real number. Consider now any $t > 0$ and let us check that $\overline{m}_v(t)$ defined in (2.42) is a nonpositive real number. The function $\mathbf{x} \mapsto m(\mathbf{x})\, v(t, \mathbf{x})$ is nonpositive and continuous in $\mathbb{R}^n$. Furthermore, it follows from Theorem 21 and the assumptions

on $q_0$ that:

$$\begin{aligned}\int_{\mathbb{R}^n} |m(\mathbf{x})|\, v(t,\mathbf{x})\,\mathrm{d}\mathbf{x} &\le \frac{1}{2(2\pi t)^{n/2}\underline{\mu}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\overline{\mu}^2 t}\right) g(\|\mathbf{y}\|)\,\|\mathbf{x}\|^2\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x},\\ &= \frac{\overline{\mu}^n}{2\pi^{n/2}\underline{\mu}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \exp(-\|\mathbf{z}\|^2)\, g(\|\mathbf{x}-\overline{\mu}\sqrt{2t}\mathbf{z}\|)\,\|\mathbf{x}\|^2\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{x},\\ &\le \frac{\overline{\mu}^n}{2\pi^{n/2}\underline{\mu}^n}\int_{\mathbb{R}^n}\left(\int_{\|\mathbf{z}\|\le\|\mathbf{x}\|/(2\overline{\mu}\sqrt{2t})} \exp(-\|\mathbf{z}\|^2)\, g(\|\mathbf{x}\|/2)\,\mathrm{d}\mathbf{z}\right.\\ &\qquad\left. + \int_{\|\mathbf{z}\|>\|\mathbf{x}\|/(2\overline{\mu}\sqrt{2t})} \exp(-\|\mathbf{z}\|^2)\, g(0)\,\mathrm{d}\mathbf{z}\right)\|\mathbf{x}\|^2\,\mathrm{d}\mathbf{x},\\ &\le \frac{\overline{\mu}^n}{2\pi^{n/2}\underline{\mu}^n}\left(\pi^{n/2}\int_{\mathbb{R}^n} g(\|\mathbf{x}\|/2)\,\|\mathbf{x}\|^2\,d\mathbf{x} + g(0)\int_{\mathbb{R}^n} h_t(\|\mathbf{x}\|)\,\|\mathbf{x}\|^2\,\mathrm{d}\mathbf{x}\right),\end{aligned}$$

where:

$$h_t(r) := \int_{\|\mathbf{z}\|\ge r/(2\overline{\mu}\sqrt{2t})} \exp(-\|\mathbf{z}\|^2)\,\mathrm{d}\mathbf{z} = O(\exp(-r)), \quad \text{as } r\to+\infty.$$

Since the integral $\int_{\mathbb{R}^n}\|\mathbf{x}\|^2\, g(\|\mathbf{x}\|)\,\mathrm{d}\mathbf{x}$ converges by assumption, one concludes that the integral $\int_{\mathbb{R}^n}|m(\mathbf{x})|\, v(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ converges as well, hence $\overline{m}_v(t)$ is a nonpositive real number. Furthermore, since the quantities $h_t(r)$ are non-decreasing with respect to $t>0$, the same arguments together with Lebesgue's dominated convergence theorem imply straightforwardly that the function $\overline{m}_v$ is continuous in $\mathbb{R}_+$.

The convergence of the integral defining $\overline{m}_v(t)$ for every $t\ge 0$, together with the nonnegativity and continuity of $v$ and $\mathbf{x}\mapsto -m(\mathbf{x}) = \|\mathbf{x}\|^2/2$, immediately implies that the integral $\int_{\mathbb{R}^n} v(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ is a nonnegative real number for each $t\ge 0$. Moreover, since $v(t,\mathbf{x})>0$ for all $(t,\mathbf{x})\in\mathbb{R}_+^*\times\mathbb{R}^n$, we infer that:

$$\int_{\mathbb{R}^n} v(t,\mathbf{x})\,\mathrm{d}\mathbf{x} > 0 \quad \text{for all } t>0.$$

As in the previous paragraph, the function $t\mapsto\int_{\mathbb{R}^n} v(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ is also continuous in $\mathbb{R}_+$.

Fix now an arbitrary $t>0$. For $R>0$, denote $B_R=\{\mathbf{x}\in\mathbb{R}^n,\ \|\mathbf{x}\|<R\}$, $\nu$ the outward unit normal on $\partial B_R$ and $\mathrm{d}\sigma(\mathbf{x})$ the surface measure on $\partial B_R$. For every $\varepsilon\in(0,t)$ and every $R>0$, the integration of (2.37) over $[\varepsilon,t]\times B_R$ yields:

$$\begin{aligned}\int_{B_R} v(t,\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_{B_R} v(\varepsilon,\mathbf{x})\,\mathrm{d}\mathbf{x} &= \frac{1}{2}\int_\varepsilon^t\int_{\partial B_R}\sum_{i=1}^n \mu_i^2\,\nu_i\,\partial_{x_i} v(s,\mathbf{x})\,\mathrm{d}\sigma(\mathbf{x})\,\mathrm{d}s\\ &\qquad + \int_\varepsilon^t\int_{B_R} m(\mathbf{x})\, v(s,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}s.\end{aligned}$$

Since the first term of the right-hand side converges to $0$ as $R\to+\infty$ from

standard parabolic estimates (see Friedman 1964), one gets that:

$$\int_{\mathbb{R}^n} v(t,\mathbf{x})\,d\mathbf{x} - \int_{\mathbb{R}^n} v(\varepsilon,\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_\varepsilon^t \overline{m}_v(s)\,\mathrm{d}s,$$

by passing to the limit $R \to +\infty$. The limit $\varepsilon \to 0^+$ then yields:

$$\int_{\mathbb{R}^n} v(t,\mathbf{x})\,\mathrm{d}\mathbf{x} - 1 = \int_{\mathbb{R}^n} v(t,\mathbf{x})\,\mathrm{d}x - \int_{\mathbb{R}^n} q_0(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_0^t \overline{m}_v(s)\,\mathrm{d}s, \tag{2.45}$$

which gives the desired result (2.43) for $t > 0$. Formula (2.43) for $t = 0$ is trivial since $v(0,\cdot) = q_0$ has unit mass. The proof of Lemma 1 is thereby complete. □

*Proof of Theorem 13.* Let $v \geq 0$ be the unique classical solution of (2.37) given in Theorem 21. From Lemma 1, the function $\overline{m}_v$ defined by (2.42) is continuous in $\mathbb{R}_+$. Let us then set:

$$q(t,\mathbf{x}) = \frac{v(t,\mathbf{x})}{1 + \int_0^t \overline{m}_v(s)\,\mathrm{d}s}, \tag{2.46}$$

for every $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n$. We recall that the denominator on the right-hand side of (2.46) is positive from Lemma 1. From Theorem 21 and Lemma 1, the function $q$ is nonnegative and of class $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$. Furthermore, $q(0,\cdot) = v(0,\cdot) = q_0$ in $\mathbb{R}^n$ and it follows from Lemma 1 that:

$$\forall\, t \geq 0, \quad \int_{\mathbb{R}^n} q(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = 1, \quad \text{and} \quad \overline{m}(t) := \int_{\mathbb{R}^n} m(\mathbf{x})\,q(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{\overline{m}_v(t)}{1 + \int_0^t \overline{m}_v(s)\,\mathrm{d}s},$$

hence the function $\overline{m}$ is real-valued, nonpositive and continuous in $\mathbb{R}_+$. Lastly, it is immediate to see that $q$ obeys (2.10). Furthermore, since $v$ is bounded in $(0,T) \times \mathbb{R}^n$, the function $q$ is bounded in $(0,T) \times \mathbb{R}^n$ for every $T > 0$.

To show the uniqueness, assume now that we have two nonnegative classical solutions $q_1$ and $q_2$ of (2.10) in $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n) \cap L^\infty((0,T) \times \mathbb{R}^n)$ (for every $T > 0$), with the same initial condition $q_0$ satisfying (2.7)-(2.9), and such that the functions:

$$\overline{m}_1(t) := \int_{\mathbb{R}^n} m(\mathbf{x})\,q_1(t,\mathbf{x})\,\mathrm{d}\mathbf{x}, \quad \text{and} \quad \overline{m}_2(t) := \int_{\mathbb{R}^n} m(\mathbf{x})\,q_1(t,\mathbf{x})\,\mathrm{d}\mathbf{x},$$

are real-valued and continuous in $\mathbb{R}_+$. Define:

$$v_i(t,\mathbf{x}) = q_i(t,\mathbf{x})\,\exp\left(\int_0^t \overline{m}_i(s)\mathrm{d}s\right),$$

for $i = 1, 2$, and $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n$. The two functions $v_1$ and $v_2$ satisfy (2.37) and are bounded in $(0,T) \times \mathbb{R}^n$ for every $T > 0$. From Theorem 21, we get $v_1 \equiv v_2$ in $\mathbb{R}_+ \times \mathbb{R}^n$. Furthermore, for all $i = 1, 2$ and $t \in \mathbb{R}_+$, there holds:

$$\overline{m}_{v_i}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})\,v_i(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = \overline{m}_i(t)\,\exp\left(\int_0^t \overline{m}_i(s)\mathrm{d}s\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\exp\left(\int_0^t \overline{m}_i(s)\mathrm{d}s\right)\right].$$

Hence, for all $(t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n$, we get:

$$q_1(t, \mathbf{x}) = \frac{v_1(t, \mathbf{x})}{1 + \int_0^t \overline{m}_{v_1}(s)\, \mathrm{d}s} = \frac{v_2(t, \mathbf{x})}{1 + \int_0^t \overline{m}_{v_2}(s)\, \mathrm{d}s} = q_2(t, \mathbf{x}).$$

The proof of Theorem 13 is thereby complete. □

*Proof of Corollary 1.* It is a straightforward calculation to check that the function $q$ defined by (2.12)-(2.13) is a classical bounded solution of (2.10) with initial condition given by (2.11). The conclusion then follows from the uniqueness part of Theorem 13. □

### 2.4.2. A degenerate parabolic PDE satisfied by $\mathbf{p}(t, \mathbf{m})$

*Proof of Proposition 14.* We recall that the phenotypes are represented by $n$ traits, and so by a vector in $\mathbb{R}^n$ and that we have a constant optimum $\mathbf{O}$, which is fixed to $(0, \dots, 0)$ up to translation. We define, for each $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{\pm 1\}^n$, the subset:

$$Q_\varepsilon = \{\mathbf{x} \in \varepsilon_1 \mathbb{R}_+^* \times \cdots \times \varepsilon_n \mathbb{R}_+^*\} \subset \mathbb{R}^n.$$

For any time $t \geq 0$, we get from the law of total probability that:

$$\int_{\mathbb{R}_-^n} \mathbf{p}(t, \mathbf{m})\, \mathrm{d}\mathbf{m} = \int_{\mathbb{R}^n} q(t, \mathbf{x})\, \mathrm{d}\mathbf{x} = 1,$$

and that, for any $m \in \mathbb{R}_-^n$,

$$\mathbf{p}(t, \mathbf{m}) = \sum_{\varepsilon \in \{\pm 1\}^n} \mathbf{p}(t, \mathbf{m} \mid \mathbf{x}(\mathbf{m}) \in Q_\varepsilon) \int_{Q_\varepsilon} q(t, \mathbf{y}) \mathrm{d}\mathbf{y}, \tag{2.47}$$

with $\mathbf{p}(t, \mathbf{m} \mid \mathbf{x}(\mathbf{m}) \in Q_\varepsilon)$ the conditional density of the fitness vector $\mathbf{m}$, given that the associated phenotype $\mathbf{x}(\mathbf{m})$ is in $Q_\varepsilon$. In the above formula (2.47), we also use the fact that $\int_H q(t, \mathbf{x})\, \mathrm{d}\mathbf{x} = 0$ with:

$$H = \bigcup_{1 \leq i \leq n} \{\mathbf{x} \in \mathbb{R}^n,\ x_i = 0\},$$

since $q(t, \cdot)$ is continuous in $\mathbb{R}^n$.

As the fitness function $\mathbf{x} \in Q_\varepsilon \mapsto \mathbf{m}(\mathbf{x}) = (m_1(\mathbf{x}), \dots, m_n(\mathbf{x}))$ with $m_i(\mathbf{x}) = -x_i^2/2$ is one-to-one from $Q_\varepsilon$ to $(\mathbb{R}_-^*)^n$, with inverse:

$$\mathbf{m} \mapsto \mathbf{x}^\varepsilon(\mathbf{m}) = (\varepsilon_1 \sqrt{-2m_1}, \dots, \varepsilon_n \sqrt{-2m_n}),$$

we infer that, for every $t \geq 0$ and $\mathbf{m} \in (\mathbb{R}_-^*)^n$,

$$\mathbf{p}(t, \mathbf{m} \mid \mathbf{x}(\mathbf{m}) \in Q_\varepsilon) \int_{Q_\varepsilon} q(t, \mathbf{y}) \mathrm{d}\mathbf{y} = \frac{q(t, \mathbf{x}^\varepsilon(\mathbf{m}))}{|\det J_\varepsilon|} = \frac{2^{-n/2}}{\sqrt{|m_1 \cdots m_n|}} q(t, \mathbf{x}^\varepsilon(\mathbf{m})),$$

with $J_\varepsilon = \mathrm{diag}(-\varepsilon_1\sqrt{-2m_1}, \ldots, -\varepsilon_n\sqrt{-2m_n})$. Finally, we get:

$$\mathbf{p}(t, \mathbf{m}) = \frac{2^{-n/2}}{\sqrt{|m_1 \cdots m_n|}} \sum_{\varepsilon \in \{\pm 1\}^n} q(t, \mathbf{x}^\varepsilon(\mathbf{m})),$$

and:

$$\begin{aligned}
\overline{m}(t) &= \sum_{i=1}^n \int_{\mathbb{R}^n} -\frac{x_i^2}{2} q(t, \mathbf{x}) \mathrm{d}\mathbf{x} = \sum_{i=1}^n \int_{\mathbb{R}^n \setminus H} -\frac{x_i^2}{2} q(t, \mathbf{x}) \mathrm{d}\mathbf{x}, \\
&= \sum_{i=1}^n \sum_{\varepsilon \in \{\pm 1\}^n} \int_{\{\varepsilon_1 x_1 > 0, \ldots, \varepsilon_n x_n > 0\}} -\frac{x_i^2}{2} q(t, \mathbf{x}) \mathrm{d}\mathbf{x}, \\
&= \sum_{i=1}^n \sum_{\varepsilon \in \{\pm 1\}^n} \int_{(\mathbb{R}_-^*)^n} m_i q(t, \mathbf{x}^\varepsilon(\mathbf{m})) \frac{2^{-n/2}}{\sqrt{|m_1 \cdots m_n|}} \mathrm{d}\mathbf{m} = \sum_{i=1}^n \int_{(\mathbb{R}_-^*)^n} m_i \mathbf{p}(t, \mathbf{m}) \mathrm{d}\mathbf{m}.
\end{aligned}$$

Notice that all integrals in the last sum converge since all integrands are non-positive and the sum of these integrals is a real number. Observe lastly that $\int_{(\mathbb{R}_-^*)^n} m_i \mathbf{p}(t, \mathbf{m}) \mathrm{d}\mathbf{m} = \int_{\mathbb{R}_-^n} m_i \mathbf{p}(t, \mathbf{m}) \mathrm{d}\mathbf{m}$, for every $1 \leq i \leq n$, since $\mathbf{p}(t, \cdot)$ is an $L^1(\mathbb{R}_-^n)$ function. The proof of Proposition 14 is thereby complete. □

*Proof of Theorem 15.* Formula (2.16) implies that the function $\mathbf{p}$ is of class $C^{1,2}(\mathbb{R}_+ \times (\mathbb{R}_-^*)^n)$ with initial condition $\mathbf{p}_0$ given by (2.18). Furthermore, it is straightforward to check that, for all $t \geq 0$ and $\mathbf{m} \in (\mathbb{R}_-^*)^n$:

$$\begin{aligned}
2^{-n/2}\, \partial_{m_i} \mathbf{p}(t, \mathbf{m}) &= -\frac{q^\#(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m}))}{2m_i\sqrt{|m_1 \cdots m_n|}} - \frac{1}{\sqrt{2|m_i|}} \frac{\partial_{x_i} q^\#(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m}))}{\sqrt{|m_1 \cdots m_n|}}, \\
2^{-n/2}\, \partial_{ii} \mathbf{p}(t, \mathbf{m}) &= \frac{3}{4} \frac{q^\#(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m}))}{m_i^2\sqrt{|m_1 \cdots m_n|}} + \frac{3}{2} \frac{\partial_{x_i} q^\#(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m}))}{m_i\sqrt{2|m_i|}\sqrt{|m_1 \cdots m_n|}} - \frac{1}{2m_i} \frac{\partial_{ii} q^\#(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m}))}{\sqrt{|m_1 \cdots m_n|}},
\end{aligned}$$

with $\mathbf{x}^{\mathbf{1}}(\mathbf{m}) = (\sqrt{-2m_1}, \dots, \sqrt{-2m_n}) \in (\mathbb{R}_+^*)^n$. Hence, we have:

$$\begin{aligned}
&\sum_{i=1}^n \mu_i^2 \left( m_i \partial_{ii} \mathbf{p}(t,\mathbf{m}) + \frac{3}{2} \partial_{m_i} \mathbf{p}(t,\mathbf{m}) \right) \\
&\qquad = - \sum_{i=1}^n \frac{2^{n/2}}{\sqrt{|m_1 \cdots m_n|}} \frac{\mu_i^2}{2} \partial_{ii} q^{\#}(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m})), \\
&\qquad = -2^{n/2} \frac{\partial_t q^{\#}(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m})) - (m(\mathbf{x}^{\mathbf{1}}(\mathbf{m})) - \overline{m}(t))\, q^{\#}(t, \mathbf{x}^{\mathbf{1}}(\mathbf{m}))}{\sqrt{|m_1 \cdots m_n|}}, \\
&\qquad = -\partial_t \mathbf{p}(t,\mathbf{m}) + \left( \sum_{i=1}^n m_i - \overline{m}(t) \right) \mathbf{p}(t,\mathbf{m}).
\end{aligned}$$

Theorem 15 is thereby proven. □

*Proof of Theorem 16.* Since by definition $p(t,m)\,\mathrm{d}m$ is the pushforward measure, at each time $t \geq 0$, of the measure $q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ by the map $\mathbf{x} \mapsto -\|\mathbf{x}\|^2/2$, it follows from the layer-cake formula that:

$$p(t,m) = (2|m|)^{n/2-1} \int_{\mathbb{S}^{n-1}} q\Big(t, \sqrt{2|m|}\,\sigma\Big)\,\mathrm{d}\sigma = (2|m|)^{n/2-1}\, Q\Big(t, \sqrt{2|m|}\Big),$$

for all $(t,m) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, where $\mathbb{S}^{n-1}$ denotes the unit Euclidean sphere of $\mathbb{R}^n$ and:

$$Q(t,r) = \int_{\mathbb{S}^{n-1}} q(t, r\,\sigma)\,\mathrm{d}\sigma,$$

for $(t,r) \in \mathbb{R}_+ \times \mathbb{R}_+$. Since $q$ is of class $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$, it is easy to see that the function:

$$(t,\mathbf{x}) \mapsto \widetilde{q}(t,\mathbf{x}) := \int_{\mathbb{S}^{n-1}} q(t, \|x\|\,\sigma)\,\mathrm{d}\sigma,$$

is of class $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ too, hence the function $Q$ is of class $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}_+)$ and $p$ is of class $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}_-^*)$. Furthermore, $\widetilde{q}(0,\cdot)$ is of class $C^{2+\alpha}(\mathbb{R}^n)$ since $q(0,\cdot) = q_0$ is of class $C^{2+\alpha}(\mathbb{R}^n)$.

Since $q$ solves (2.10), which is invariant by rotation in the present isotropic case ($\mu_i = \mu$ for every $1 \leq i \leq n$) and since $\overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})\, q(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} m(\mathbf{x})\, \widetilde{q}(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$, for every $t \geq 0$, it follows that $\widetilde{q}$ solves (2.10) as well, with initial condition $\widetilde{q}(0,\cdot)$. As a consequence, $Q$ satisfies:

$$\partial_t Q(t,r) = \frac{\mu^2}{2} \left( \partial_{rr} Q(t,r) + \frac{n-1}{r} \partial_r Q(t,r) \right) + \left( -\frac{r^2}{2} - \overline{m}(t) \right) Q(t,r),$$

for all $(t,r) \in \mathbb{R}_+ \times \mathbb{R}_+^*$. But since:

$$Q(t,r) = r^{2-n}\, p\Big(t, -\frac{r^2}{2}\Big),$$

for all $(t,r) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, it is then straightforward to check that $p$ solves (2.19) in $\mathbb{R}_+ \times \mathbb{R}_-^*$. Lastly, the formula (2.20) is an immediate consequence of the definitions of $p$ and $\mathbf{p}$, and the proof of Theorem 16 is thereby complete. □

### 2.4.3. Generating functions

*Proof of Theorem 17.* Given $q_0$ satisfying (2.7)-(2.9), we call $v$ the unique bounded nonnegative $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ solution of (2.37) defined in Theorem 21 with initial condition $q_0^\#$. Notice that the function $q_0^\#$ satisfies the same assumptions (2.7)-(2.9) as $q_0$. By uniqueness and symmetry of (2.37) with respect to the change of variable $x_i$ into $-x_i$, for any $1 \le i \le n$, it follows that $v(t,\mathbf{x}) = v^\#(t,\mathbf{x})$ and, as in (2.46),

$$q^\#(t,\mathbf{x}) = \frac{v^\#(t,\mathbf{x})}{1+\int_0^t \overline{m}_v(s)\,\mathrm{d}s} = \frac{v(t,\mathbf{x})}{1+\int_0^t \overline{m}_v(s)\,\mathrm{d}s}, \tag{2.48}$$

for all $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}_+^n$.

As already noticed in Section 2.2.1, from (2.15) and the nonnegativity of $\mathbf{p}$, the functions $(t,\mathbf{z}) \mapsto M_{\mathbf{p}}(t,\mathbf{z})$ and $(t,\mathbf{z}) \mapsto C_{\mathbf{p}}(t,\mathbf{z})$ given in (2.21)-(2.22) are well defined in $\mathbb{R}_+ \times \mathbb{R}_+^n$. Let us start by proving the continuity of these functions $M_{\mathbf{p}}$ and $C_{\mathbf{p}}$ in $\mathbb{R}_+ \times \mathbb{R}_+^n$. Owing to the relations (2.16) and (2.48), we see that:

$$M_{\mathbf{p}}(t,\mathbf{z}) = \frac{2^{n/2}}{1+\int_0^t \overline{m}_v(s)ds} \int_{\mathbb{R}_-^n} \frac{e^{\mathbf{z}\cdot\mathbf{m}}}{\sqrt{|m_1\cdots m_n|}}\, v(t,\mathbf{x}^{\mathbf{1}}(\mathbf{m}))\,\mathrm{d}\mathbf{m}, \tag{2.49}$$

for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}_+^n$, with $\mathbf{x}^{\mathbf{1}}(\mathbf{m}) = (\sqrt{-2m_1},\ldots,\sqrt{-2m_n})$. Hence, we have:

$$\begin{aligned} M_{\mathbf{p}}(t,\mathbf{z}) &= \frac{2^n}{1+\int_0^t \overline{m}_v(s)ds} \int_{\mathbb{R}_+^n} \exp\Big(-\sum_{i=1}^n \frac{z_i x_i^2}{2}\Big)\, v(t,\mathbf{x})\,\mathrm{d}\mathbf{x}, \\ &= \frac{1}{1+\int_0^t \overline{m}_v(s)ds} \int_{\mathbb{R}^n} \exp\Big(-\sum_{i=1}^n \frac{z_i x_i^2}{2}\Big)\, v(t,\mathbf{x})\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{2.50}$$

Notice that the function $t \mapsto \int_0^t \overline{m}_v(s)\mathrm{d}s$ is continuous in $\mathbb{R}_+$. Furthermore, the function $(t,\mathbf{z}) \mapsto \exp(-\sum_{i=1}^n z_i x_i^2/2)\, v(t,\mathbf{x})$ is also continuous in $\mathbb{R}_+ \times \mathbb{R}_+^n$, for every $\mathbf{x} \in \mathbb{R}^n$. Lastly, as in the proof of Lemma 1, it follows from (2.9), (2.39)-(2.40) and (2.44) that, for any $\mathbf{z} \in \mathbb{R}_+^n$, $T > 0$, $t \in (0,T]$ and $\mathbf{x} \in \mathbb{R}^n$, there

holds:

$$
\begin{aligned}
0 \le \exp\Big(-\sum_{i=1}^n \frac{z_i x_i^2}{2}\Big)\, v(t,\mathbf{x}) &\le v(t,\mathbf{x}),\\
&\le \frac{1}{(2\pi t)^{n/2}\underline{\mu}^n} \int_{\mathbb{R}^n} \exp\Big(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\overline{\mu}^2 t}\Big)\, q_0^{\#}(\mathbf{y})\,\mathrm{d}\mathbf{y},\\
&\le \frac{\overline{\mu}^n}{\pi^{n/2}\underline{\mu}^n} \int_{\mathbb{R}^n} e^{-\|\mathbf{y}'\|^2}\, g(\|\mathbf{x}-\overline{\mu}\sqrt{2t}\mathbf{y}'\|)\,\mathrm{d}\mathbf{y}',\\
&\le \frac{\overline{\mu}^n}{\pi^{n/2}\underline{\mu}^n}\Big[\int_{\|\mathbf{y}'\|\le\|\mathbf{x}\|/(2\overline{\mu}\sqrt{2t})} e^{-\|\mathbf{y}'\|^2}\, g(\|\mathbf{x}\|/2)\,\mathrm{d}\mathbf{y}' + g(0)\int_{\|\mathbf{y}'\|>\|\mathbf{x}\|/(2\overline{\mu}\sqrt{2t})} e^{-\|\mathbf{y}'\|^2}\,\mathrm{d}\mathbf{y}'\Big],\\
&\le \frac{\overline{\mu}^n}{\pi^{n/2}\underline{\mu}^n}\Big[\pi^{n/2} g(\|\mathbf{x}\|/2) + g(0)\int_{\|\mathbf{y}'\|>\|\mathbf{x}\|/(2\overline{\mu}\sqrt{2T})} e^{-\|\mathbf{y}'\|^2}\,\mathrm{d}\mathbf{y}'\Big].
\end{aligned}
\tag{2.51}
$$

Call $h(\mathbf{x})$ the quantity given on the right-hand side of the last inequality. Since by (2.9) the function $g(\|\cdot\|)$ is in $L^\infty(\mathbb{R}^n)\cap L^1(\mathbb{R}^n)$, the function $h$ belongs to $L^1(\mathbb{R}^n)$, and is independent of $\mathbf{z}\in\mathbb{R}^n_+$ and $t\in(0,T]$. One then infers from (2.50) and Lebesgue's dominated convergence theorem that the function $M_{\mathbf{p}}$ is continuous in $\mathbb{R}_+\times\mathbb{R}^n_+$. As $C_{\mathbf{p}}=\log M_{\mathbf{p}}$, the cumulant generating function $C_{\mathbf{p}}$ is also continuous in $\mathbb{R}_+\times\mathbb{R}^n_+$.

Let us then check that $M_{\mathbf{p}}$ and $C_{\mathbf{p}}$ are $C^{0,1}(\mathbb{R}_+\times\mathbb{R}^n_+)$, meaning that the functions $\partial_{z_i}M_{\mathbf{p}}=\frac{\partial M_{\mathbf{p}}}{\partial z_i}$ and $\partial_{z_i}C_{\mathbf{p}}=\frac{\partial C_{\mathbf{p}}}{\partial z_i}$ exist and are continuous in $\mathbb{R}_+\times\mathbb{R}^n_+$, for every $1\le i\le n$. As a matter of fact, since $\mathbf{p}(t,\cdot)$ is a probability density function in $\mathbb{R}^n_-$ for any $t\ge 0$ and since the integral $\int_{\mathbb{R}^n_-}\|\mathbf{m}\|\,\mathbf{p}(t,\mathbf{m})\,\mathrm{d}\mathbf{m}$ converges by formula (2.15) in Proposition 14, it easily follows from Lebesgue's dominated convergence theorem that $\partial_{z_i}M_{\mathbf{p}}(t,\mathbf{z})$ exists for all $(t,\mathbf{z})\in\mathbb{R}_+\times\mathbb{R}^n_+$ and $1\le i\le n$, with:

$$
\partial_{z_i}M_{\mathbf{p}}(t,\mathbf{z}) = \int_{\mathbb{R}^n_-} m_i\, e^{\mathbf{z}\cdot\mathbf{m}}\mathbf{p}(t,\mathbf{m})\,\mathrm{d}\mathbf{m},
\tag{2.52}
$$

hence, as in (2.49)-(2.50),

$$
\begin{aligned}
\partial_{z_i}M_{\mathbf{p}}(t,\mathbf{z}) &= \frac{2^{n/2}}{1+\int_0^t \overline{m}_v(s)\mathrm{d}s}\int_{\mathbb{R}^n_-} m_i\,\frac{e^{\mathbf{z}\cdot\mathbf{m}}}{\sqrt{|m_1\cdots m_n|}}\, v(t,\mathbf{x}^{\mathbf{1}}(\mathbf{m}))\,\mathrm{d}\mathbf{m},\\
&= -\frac{1}{1+\int_0^t \overline{m}_v(s)\mathrm{d}s}\int_{\mathbb{R}^n}\frac{x_i^2}{2}\exp\Big(-\sum_{i=1}^n\frac{z_i x_i^2}{2}\Big)\, v(t,\mathbf{x})\,\mathrm{d}\mathbf{x},
\end{aligned}
\tag{2.53}
$$

from (2.16) and (2.48). On the one hand, the function $t\mapsto\int_0^t\overline{m}_v(s)\mathrm{d}s$ is continuous in $\mathbb{R}_+$ and so is the function $(t,\mathbf{z})\mapsto x_i^2\exp(-\sum_{i=1}^n z_i x_i^2/2)\,v(t,\mathbf{x})$ in $\mathbb{R}_+\times\mathbb{R}^n_+$, for every $\mathbf{x}\in\mathbb{R}^n$. On the other hand, as in the previous paragraph, it follows from (2.39)-(2.40) and (2.44) that, for any $1\le i\le n$, $\mathbf{z}\in\mathbb{R}^n_+$, $T>0$, $t\in(0,T]$

and $\mathbf{x} \in \mathbb{R}^n$, there holds:

$$0 \le x_i^2 \, \exp\Big( - \sum_{i=1}^n \frac{z_i x_i^2}{2} \Big) \, v(t,\mathbf{x}),$$
$$\le \|\mathbf{x}\|^2 \, v(t,\mathbf{x}) \le \frac{\overline{\mu}^n}{\pi^{n/2} \underline{\mu}^n} \Big[ \pi^{n/2} \|\mathbf{x}\|^2 g(\|\mathbf{x}\|/2) + g(0) \|\mathbf{x}\|^2 \int_{\|\mathbf{y}'\| > \|\mathbf{x}\|/(2\overline{\mu}\sqrt{2T})} e^{-\|\mathbf{y}'\|^2} \, \mathrm{d}\mathbf{y}' \Big].$$

Call $\widetilde{h}(\mathbf{x})$ the quantity given on the right-hand side of the last inequality. Since by (2.9) the function $\mathbf{x} \mapsto \|\mathbf{x}\|^2 g(\|\mathbf{x}\|)$ is in $L^1(\mathbb{R}^n)$, the function $\widetilde{h}$ belongs to $L^1(\mathbb{R}^n)$, and is independent of $\mathbf{z} \in \mathbb{R}^n_+$ and $t \in (0,T]$. One then infers from (2.53) and Lebesgue's dominated convergence theorem that the function $\partial_{z_i} M_{\mathbf{p}}$ is continuous in $\mathbb{R}_+ \times \mathbb{R}^n_+$, and so is the function $\partial_{z_i} C_{\mathbf{p}} = \partial_{z_i} M_{\mathbf{p}} / M_{\mathbf{p}}$.

In this paragraph, we are interested in the differentiation of $M_{\mathbf{p}}$ with respect to $t$. By (2.37), we already know that:

$$|\partial_t v(t,\mathbf{x})| \le \sum_{i=1}^n \frac{\mu_i^2}{2} |\partial_{ii} v(t,\mathbf{x})| + |m(\mathbf{x}) \, v(t,\mathbf{x})|, \tag{2.54}$$

for all $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n$. Fix $T > 0$ and let $S > 0$ be the constant given in (2.38). Thus, for all $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^n$, there holds:

$$|\partial_{ii} v(t,\mathbf{x})| \le \|v(t,\cdot)\|_{C^{2+\alpha}(\mathbb{R}^n)} \le S \|q_0^{\#}\|_{C^{2+\alpha}(\mathbb{R}^n)} \le S \|q_0\|_{C^{2+\alpha}(\mathbb{R}^n)}. \tag{2.55}$$

Let us now focus on the boundedness of the second term of the right-hand side of (2.54), that is, the boundedness of the function $(t,\mathbf{x}) \mapsto m(\mathbf{x}) v(t,\mathbf{x})$ in $[0,T] \times \mathbb{R}^n$. Since this function is continuous in $\mathbb{R}_+ \times \mathbb{R}^n$, let us show its boundedness in $(0,T] \times \mathbb{R}^n$. Thanks to (2.9) and (2.39)-(2.40), we get, as in (2.51), that:

$$|m(\mathbf{x}) v(t,\mathbf{x})| \le \int_{\mathbb{R}^n} \frac{\overline{\mu}^n \|x\|^2}{2\pi^{n/2} \underline{\mu}^n} \, q_0^{\#} \Big( \mathbf{x} - \overline{\mu} \sqrt{2t} \mathbf{y} \Big) \, e^{-\|\mathbf{y}\|^2} \, \mathrm{d}\mathbf{y},$$

for all $(t,\mathbf{x}) \in (0,T] \times \mathbb{R}^n$. Thus, we have:

$$|m(\mathbf{x}) v(t,\mathbf{x})| \le \frac{\overline{\mu}^n}{\pi^{n/2} \underline{\mu}^n} \int_{\mathbb{R}^n} \Big( \|\mathbf{x} - \overline{\mu}\sqrt{2t}\mathbf{y}\|^2 + 2t\overline{\mu}^2 \|\mathbf{y}\|^2 \Big) \, g\Big( \|\mathbf{x} - \overline{\mu}\sqrt{2t}\mathbf{y}\| \Big) \, e^{-\|\mathbf{y}\|^2} \, \mathrm{d}\mathbf{y},$$
$$\le \frac{C \, \overline{\mu}^n}{\pi^{n/2} \underline{\mu}^n} + \frac{2\overline{\mu}^{n+2} T \, g(0)}{\pi^{n/2} \underline{\mu}^n} \int_{\mathbb{R}^n} \|\mathbf{y}\|^2 \, e^{-\|\mathbf{y}\|^2} \, \mathrm{d}\mathbf{y},$$

where the constant $C$ is such that $\|\mathbf{x}'\|^2 q_0^{\#}(\mathbf{x}') \le \|\mathbf{x}'\|^2 g(\|\mathbf{x}'\|) \le C$, for all $\mathbf{x}' \in \mathbb{R}^n$. Therefore, the function $(t,\mathbf{x}) \mapsto m(\mathbf{x}) v(t,\mathbf{x})$ is bounded in $[0,T] \times \mathbb{R}^n$ for any $T > 0$, and so is $\partial_t v$ by (2.54)-(2.55). Together with (2.38), (2.48) and the continuity of $\overline{m}_v$ in $\mathbb{R}_+$, it follows that the function $\partial_t q^{\#}$ is bounded in $[0,T] \times \mathbb{R}^n$, for every $T > 0$. Finally, (2.16) implies that for all $(t,\mathbf{z}) \in [0,T] \times \mathbb{R}^n_+$ and

$\mathbf{m} \in (\mathbb{R}_-^*)^n$,

$$|e^{\mathbf{z}\cdot\mathbf{m}}\,\partial_t \mathbf{p}(t,\mathbf{m})| \le 2^{n/2}\|\partial_t q^{\#}\|_{L^\infty([0,T]\times\mathbb{R}^n)}\frac{e^{\mathbf{z}\cdot\mathbf{m}}}{\sqrt{|m_1\cdots m_n|}}.$$

Since the integrals:

$$\int_{\mathbb{R}_-^n}\frac{e^{\mathbf{z}\cdot\mathbf{m}}}{\sqrt{|m_1\cdots m_n|}}\,\mathrm{d}\mathbf{m} = 2^{n/2}\int_{\mathbb{R}_+^n}\exp\Big(-\sum_{i=1}^n\frac{z_i x_i^2}{2}\Big)\,\mathrm{d}\mathbf{x},$$

converge for all $\mathbf{z} \in (\mathbb{R}_+^*)^n$, it then easily follows from the previous estimates and from Lebesgue's dominated convergence theorem that the function $M_{\mathbf{p}}$ is differentiable with respect to $t$ in $\mathbb{R}_+ \times (\mathbb{R}_+^*)^n$, with:

$$\partial_t M_{\mathbf{p}}(t,\mathbf{z}) = \int_{\mathbb{R}_-^n} e^{\mathbf{z}\cdot\mathbf{m}}\,\partial_t \mathbf{p}(t,\mathbf{m})\,\mathrm{d}\mathbf{m}, \tag{2.56}$$

and that the function $\partial_t M_{\mathbf{p}}$ is itself continuous in $\mathbb{R}_+ \times (\mathbb{R}_+^*)^n$. So is the function $\partial_t C_{\mathbf{p}} = \partial_t M_{\mathbf{p}}/M_{\mathbf{p}}$. The continuity of the functions $\partial_t M_{\mathbf{p}}$ and $\partial_t C_{\mathbf{p}}$ in the closure $\mathbb{R}_+ \times \mathbb{R}_+^n$ of $\mathbb{R}_+ \times (\mathbb{R}_+^*)^n$ will be obtained as a consequence of the equations satisfied by these two functions, which shall be established below.

Let us then turn to find an equation satisfied by $M_{\mathbf{p}}$, in order to derive Eq. (2.23) satisfied by $C_{\mathbf{p}}$. Fix $(t,\mathbf{z}) \in \mathbb{R}_+ \times (\mathbb{R}_+^*)^n$. Thanks to (2.17), we have:

$$\begin{aligned} e^{\mathbf{z}\cdot\mathbf{m}}\,\partial_t \mathbf{p}(t,\mathbf{m}) = -\sum_{i=1}^n \mu_i^2 m_i e^{\mathbf{z}\cdot\mathbf{m}}\partial_{ii}\mathbf{p}(t,\mathbf{m}) - \frac{3}{2}\sum_{i=1}^n \mu_i^2 e^{\mathbf{z}\cdot\mathbf{m}}\partial_{m_i}\mathbf{p}(t,\mathbf{m}) \\ + \Big(\sum_{i=1}^n m_i - \overline{m}(t)\Big)\,e^{\mathbf{z}\cdot\mathbf{m}}\,\mathbf{p}(t,\mathbf{m}), \end{aligned} \tag{2.57}$$

for all $\mathbf{m} \in (\mathbb{R}_-^*)^n$.

We are now going to integrate (2.57) over $\mathbb{R}_-^n$. To do so, let us first focus on the first two terms of the right-hand side of (2.57). Fix an index $i \in [\![1,n]\!]$ and consider the cubes:

$$B_\varepsilon = \big\{\mathbf{m} = (m_1,\dots,m_n) \in \mathbb{R}_-^n,\ -\varepsilon^{-1} < m_j < -\varepsilon, \text{ for each } j \in [\![1,n]\!]\big\} = (-\varepsilon^{-1},-\varepsilon)^n,$$

with $0 < \varepsilon < 1$. Denote $\widehat{B}_\epsilon = (-\varepsilon^{-1},-\varepsilon)^{n-1}$ and:

$$\widehat{\mathbf{m}} = (m_1,\dots,m_{i-1},m_{i+1},\dots,m_n),\quad \mathbf{m}^\rho = (m_1,\dots,m_{i-1},-\rho,m_{i+1},\dots,m_n),$$

for $\rho \in \mathbb{R}$. By using Fubini's theorem and integrating by parts with respect to the

variable $m_i$, one infers that:

$$\begin{aligned}
&\int_{B_\varepsilon} \Big( m_i e^{\mathbf{z}\cdot\mathbf{m}} \partial_{ii}\mathbf{p}(t,\mathbf{m}) + \frac{3}{2} e^{\mathbf{z}\cdot\mathbf{m}} \partial_{m_i}\mathbf{p}(t,\mathbf{m}) \Big)\, \mathrm{d}\mathbf{m} \\
&= \int_{B_\varepsilon} \Big( \frac{z_i}{2} + m_i z_i^2 \Big) e^{\mathbf{z}\cdot\mathbf{m}}\, \mathbf{p}(t,\mathbf{m})\, \mathrm{d}\mathbf{m} \\
&+ \int_{\widehat{B}_\varepsilon} \Big[ z_i\, \varepsilon\, e^{\mathbf{z}\cdot\mathbf{m}^\varepsilon} \mathbf{p}(t,\mathbf{m}^\varepsilon) - z_i\, \varepsilon^{-1}\, e^{\mathbf{z}\cdot\mathbf{m}^{1/\varepsilon}} \mathbf{p}(t,\mathbf{m}^{1/\varepsilon}) \Big]\, \mathrm{d}\widehat{\mathbf{m}} \\
&+ \int_{\widehat{B}_\varepsilon} \left[ e^{\mathbf{z}\cdot\mathbf{m}^\varepsilon} \left( \frac{\mathbf{p}(t,\mathbf{m}^\varepsilon)}{2} - \varepsilon \partial_{m_i}\mathbf{p}(t,\mathbf{m}^\varepsilon) \right) - e^{\mathbf{z}\cdot\mathbf{m}^{1/\varepsilon}} \left( \frac{\mathbf{p}(t,\mathbf{m}^{1/\varepsilon})}{2} - \varepsilon^{-1} \partial_{m_i}\mathbf{p}(t,\mathbf{m}^{1/\epsilon}) \right) \right] \mathrm{d}\widehat{\mathbf{m}}.
\end{aligned} \tag{2.58}$$

Let us pass to the limit as $\varepsilon \to 0$ in the three integrals of the right-hand side of (2.58). Firstly, since $\mathbf{p}(t,\cdot)$ is nonnegative and the functions $\mathbf{m} \mapsto \mathbf{p}(t,\mathbf{m})$ and $\mathbf{m} \mapsto m_i \mathbf{p}(t,\mathbf{m})$ are in $L^1(\mathbb{R}^n_-)$, it follows from Lebesgue's dominated convergence theorem together with (2.21) and (2.52) that:

$$\int_{B_\varepsilon} \Big( \frac{z_i}{2} + m_i z_i^2 \Big) e^{\mathbf{z}\cdot\mathbf{m}}\, \mathbf{p}(t,\mathbf{m})\, \mathrm{d}\mathbf{m} \to \int_{\mathbb{R}^n_-} \Big( \frac{z_i}{2} + m_i z_i^2 \Big) e^{\mathbf{z}\cdot\mathbf{m}}\, \mathbf{p}(t,\mathbf{m})\, \mathrm{d}\mathbf{m} = \frac{z_i}{2} M_{\mathbf{p}}(t,\mathbf{z}) + z_i^2 \partial_{z_i} M_{\mathbf{p}}(t,\mathbf{z}),$$

as $\varepsilon \to 0$. Secondly, by denoting:

$$\widehat{C}_\varepsilon = \Big( \sqrt{2\varepsilon}, \sqrt{2\varepsilon^{-1}} \Big)^{n-1}, \quad \widehat{\mathbf{z}} = (z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n),$$

and:

$$\widehat{\mathbf{x}} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n), \quad \mathbf{x}^\rho = (x_1, \ldots, x_{i-1}, \sqrt{2\rho}, x_{i+1}, \ldots, x_n),$$

for $\rho \geq 0$, it follows from (2.16) that:

$$\begin{aligned}
\int_{\widehat{B}_\varepsilon} z_i\, \varepsilon\, e^{\mathbf{z}\cdot\mathbf{m}^\varepsilon} \mathbf{p}(t,\mathbf{m}^\varepsilon)\, \mathrm{d}\widehat{\mathbf{m}} &= z_i \sqrt{\varepsilon}\, e^{-z_i\varepsilon} \int_{\widehat{B}_\varepsilon} \frac{2^{n/2}\, e^{\widehat{\mathbf{z}}\cdot\widehat{\mathbf{m}}}\, q^\#(t,\mathbf{x}^{\mathbf{1}}(\mathbf{m}^\varepsilon))}{\sqrt{|m_1 \cdots m_{i-1} m_{i+1} \cdots m_n|}}\, \mathrm{d}\widehat{\mathbf{m}}, \\
&= z_i \sqrt{\varepsilon}\, e^{-z_i\varepsilon}\, 2^{n-1/2} \int_{\widehat{C}_\varepsilon} \exp\Big( -\sum_{j\neq i} \frac{z_j x_j^2}{2} \Big)\, q^\#(t,\mathbf{x}^\varepsilon)\, \mathrm{d}\widehat{\mathbf{x}}.
\end{aligned}$$

Since the continuous function $q^\#(t,\cdot)$ is bounded in $\mathbb{R}^n$ by (2.38) and (2.48), and since $\mathbf{z} \in (\mathbb{R}^*_+)^n$, one then gets that:

$$\int_{\widehat{B}_\varepsilon} z_i\, \varepsilon\, e^{\mathbf{z}\cdot\mathbf{m}^\varepsilon} \mathbf{p}(t,\mathbf{m}^\varepsilon)\, \mathrm{d}\widehat{\mathbf{m}} \to 0, \quad \text{as } \varepsilon \to 0.$$

Similarly, we prove that:

$$\int_{\widehat{B}_\varepsilon} z_i\, \varepsilon^{-1}\, e^{\mathbf{z}\cdot\mathbf{m}^{1/\varepsilon}} \mathbf{p}(t,\mathbf{m}^{1/\varepsilon})\, \mathrm{d}\widehat{\mathbf{m}} = z_i \sqrt{\varepsilon^{-1}}\, e^{-z_i/\varepsilon}\, 2^{n-1/2} \int_{\widehat{C}_\varepsilon} \exp\Big( -\sum_{j\neq i} \frac{z_j x_j^2}{2} \Big)\, q^\#(t,\mathbf{x}^{1/\varepsilon})\, \mathrm{d}\widehat{\mathbf{x}} \to 0,$$

as $\varepsilon \to 0$. Thirdly, from the computations done in the proof of Theorem 15, we

already know that, for all $\mathbf{m} \in (\mathbb{R}_-^*)^n$:

$$\frac{1}{2}\,\mathbf{p}(t,\mathbf{m}) + m_i \partial_{m_i}\mathbf{p}(t,\mathbf{m}) = \frac{2^{n/2}\partial_{x_i} q^{\#}(t,\mathbf{x}^{\mathbf{1}}(\mathbf{m}))}{\sqrt{2|m_1\cdots m_{i-1}m_{i+1}\cdots m_n|}}.$$

Hence, we have:

$$\begin{aligned}\int_{\widehat{B}_\varepsilon} e^{\mathbf{z}\cdot\mathbf{m}^\varepsilon}\left(\frac{\mathbf{p}(t,\mathbf{m}^\varepsilon)}{2} - \varepsilon\partial_{m_i}\mathbf{p}(t,\mathbf{m}^\varepsilon)\right)\mathrm{d}\widehat{\mathbf{m}} &= e^{-z_i\varepsilon}\int_{\widehat{B}_\varepsilon}\frac{2^{n/2-1/2}\,e^{\widehat{\mathbf{z}}\cdot\widehat{\mathbf{m}}}\,\partial_{x_i}q^{\#}(t,\mathbf{x}^{\mathbf{1}}(\mathbf{m}^\varepsilon))}{\sqrt{|m_1\cdots m_{i-1}m_{i+1}\cdots m_n|}}\,\mathrm{d}\widehat{\mathbf{m}},\\ &= e^{-z_i\varepsilon}\,2^{n-1}\int_{\widehat{C}_\varepsilon}\exp\left(-\sum_{j\neq i}\frac{z_j x_j^2}{2}\right)\,\partial_{x_i}q^{\#}(t,\mathbf{x}^\varepsilon)\,\mathrm{d}\widehat{\mathbf{x}}.\end{aligned}$$

Since the function $\partial_{x_i}q^{\#}(t,\cdot)$ is continuous and bounded in $\mathbb{R}^n$ by (2.38) and (2.48), since $\partial_{x_i}q^{\#}(t,\mathbf{x}^0)=0$, by $\#$-symmetry of $q^{\#}(t,\cdot)$, and since $\mathbf{z}\in(\mathbb{R}_+^*)^n$, one then infers from Lebesgue's dominated convergence theorem that:

$$\int_{\widehat{B}_\varepsilon} e^{\mathbf{z}\cdot\mathbf{m}^\varepsilon}\left(\frac{\mathbf{p}(t,\mathbf{m}^\varepsilon)}{2} - \varepsilon\partial_{m_i}\mathbf{p}(t,\mathbf{m}^\varepsilon)\right)\mathrm{d}\widehat{\mathbf{m}} \to 0, \quad \text{as } \varepsilon\to 0.$$

Furthermore, the integral:

$$\begin{aligned}\int_{\widehat{B}_\varepsilon} e^{\mathbf{z}\cdot\mathbf{m}^{1/\varepsilon}}&\left(\frac{\mathbf{p}(t,\mathbf{m}^{1/\varepsilon})}{2} - \varepsilon^{-1}\partial_{m_i}\mathbf{p}(t,\mathbf{m}^{1/\varepsilon})\right)\mathrm{d}\widehat{\mathbf{m}}\\ &= e^{-z_i/\varepsilon}\,2^{n-1}\int_{\widehat{C}_\varepsilon}\exp\left(-\sum_{j\neq i}\frac{z_j x_j^2}{2}\right)\,\partial_{x_i}q^{\#}(t,\mathbf{x}^{1/\varepsilon})\,\mathrm{d}\widehat{\mathbf{x}},\end{aligned}$$

converges to $0$ as $\varepsilon\to 0$. Coming back to (2.58) and passing to the limit as $\varepsilon\to 0$, it follows from the previous estimates that:

$$\int_{B_\varepsilon}\left(m_i e^{\mathbf{z}\cdot\mathbf{m}}\partial_{ii}\mathbf{p}(t,\mathbf{m}) + \frac{3}{2}e^{\mathbf{z}\cdot\mathbf{m}}\partial_{m_i}\mathbf{p}(t,\mathbf{m})\right)\mathrm{d}\mathbf{m} \underset{\varepsilon\to 0}{\longrightarrow} \frac{z_i}{2}\,M_{\mathbf{p}}(t,\mathbf{z}) + z_i^2\partial_{z_i}M_{\mathbf{p}}(t,\mathbf{z}). \quad (2.59)$$

Let us finally remember (2.57) and that the functions:

$$\mathbf{m}\mapsto e^{\mathbf{z}\cdot\mathbf{m}}\partial_t\mathbf{p}(t,\mathbf{m}), \quad \mathbf{m}\mapsto m_i e^{\mathbf{z}\cdot\mathbf{m}}\mathbf{p}(t,\mathbf{m}) \text{ and } \mathbf{m}\mapsto e^{\mathbf{z}\cdot\mathbf{m}}\mathbf{p}(t,\mathbf{m}),$$

are in $L^1(\mathbb{R}_-^n)$ with integrals given by (2.56), (2.52) and (2.21), respectively. Together with (2.58) and (2.59), one concludes that:

$$\partial_t M_{\mathbf{p}}(t,\mathbf{z}) = \sum_{i=1}^n(1-\mu_i^2 z_i^2)\partial_{z_i}M_{\mathbf{p}}(t,\mathbf{z}) - \frac{1}{2}\sum_{i=1}^n \mu_i^2\, z_i\, M_{\mathbf{p}}(t,\mathbf{z}) - \overline{m}(t)M_{\mathbf{p}}(t,\mathbf{z}), \quad (2.60)$$

for every $(t,\mathbf{z})\in\mathbb{R}_+\times(\mathbb{R}_+^*)^n$. Since the right-hand side of the above equation is continuous in $\mathbb{R}_+\times\mathbb{R}_+^n$, one infers that the function $\partial_t M_{\mathbf{p}}$ is extendable by

continuity in $\mathbb{R}_+ \times \mathbb{R}^n_+$ and (2.60) holds in $\mathbb{R}_+ \times \mathbb{R}^n_+$. Owing to the definition $C_{\mathbf{p}} = \log M_{\mathbf{p}}$, one concludes that $\partial_t C_{\mathbf{p}}$ is continuous in $\mathbb{R}_+ \times \mathbb{R}^n_+$ (finally, $C_{\mathbf{p}}$ is of class $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^n_+)$) and:

$$\partial_t C_{\mathbf{p}}(t,\mathbf{z}) = \sum_{i=1}^n (1-\mu_i^2 z_i^2)\partial_{z_i} C_{\mathbf{p}}(t,\mathbf{z}) - \frac{1}{2}\sum_{i=1}^n \mu_i^2\, z_i - \overline{m}(t) = A(\mathbf{z})\cdot\nabla C_{\mathbf{p}}(t,\mathbf{z}) - b(\mathbf{z}) - \overline{m}(t),$$

for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}^n_+$, where $A$ and $b$ are as in (2.24). Therefore, (2.23) holds in $\mathbb{R}_+ \times \mathbb{R}^n_+$ and the proof of Theorem 17 is thereby complete. □

Before going into the proof of the remaining results, let us first observe that, in (2.23)-(2.24), $\overline{m}(t) = \mathbf{1}\cdot\nabla C_{\mathbf{p}}(t,\mathbf{O})$, with:

$$\mathbf{1} = (1,\dots,1) \in \mathbb{R}^n.$$

It turns out that, if $A(\mathbf{z}) = \mathbf{1}$, then equation (2.23) can be solved explicitly by the method of characteristics, as the following lemma shows (this lemma is used later in the proof of Proposition 18 in the general case $A(\mathbf{z})$ given in (2.24)).

**Lemma 2.** *The Cauchy problem:*

$$\left\{\begin{array}{rcl} \partial_t Q(t,\mathbf{z}) &=& \mathbf{1}\cdot(\nabla Q(t,\mathbf{z}) - \nabla Q(t,\mathbf{O})) - \widetilde{b}(\mathbf{z}),\ t\geq 0,\ \mathbf{z}\in\mathbb{R}^n_+, \\ Q(0,\mathbf{z}) &=& Q_0(\mathbf{z}),\ \mathbf{z}\in\mathbb{R}^n_+, \\ Q(t,\mathbf{O}) &=& 0,\ t\geq 0, \end{array}\right. \tag{2.61}$$

*with $\widetilde{b} \in C^1(\mathbb{R}^n_+)$ and $Q_0 \in C^1(\mathbb{R}^n_+)$ such that $\widetilde{b}(\mathbf{O}) = Q_0(\mathbf{O}) = 0$, admits a unique $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^n_+)$ solution, which is given by the expression:*

$$Q(t,\mathbf{z}) = \int_0^t \Big(\widetilde{b}(s\mathbf{1}) - \widetilde{b}(\mathbf{z}+s\mathbf{1})\Big)\,\mathrm{d}s + Q_0(\mathbf{z}+t\mathbf{1}) - Q_0(t\mathbf{1}). \tag{2.62}$$

*Proof.* First of all, it is immediate to check that the function $Q$ given by (2.62) is a $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^n_+)$ solution of (2.61). Let now $Q_1$ and $Q_2$ be two $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^n_+)$ solutions of (2.61) and denote $Q = Q_1 - Q_2$. The function $Q$ is of class $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^n_+)$ and obeys:

$$\partial_t Q(t,\mathbf{z}) = \mathbf{1}\cdot(\nabla Q(t,\mathbf{z}) - \nabla Q(t,\mathbf{O}))\,,$$

for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}^n_+$, together with $Q(0,\mathbf{z}) = 0$ for all $\mathbf{z}\in\mathbb{R}^n_+$ and $Q(t,\mathbf{O}) = 0$ for all $t \geq 0$. It remains to show that $Q = 0$ in $\mathbb{R}_+ \times \mathbb{R}^n_+$. Fix any $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}^n_+$. If $t = 0$, then $Q(0,\mathbf{z}) = 0$, so let us assume that $t > 0$. Consider the $C^1([0,t])$ function $R$ defined by $R(s) = Q(t-s,\mathbf{z}+s\mathbf{1}) - Q(t-s,s\mathbf{1})$ for $s \in [0,t]$ (which is well defined since $\mathbf{z}+s\mathbf{1} \in \mathbb{R}^n_+$). It follows from the equation satisfied by $Q$ that,

for all $s \in [0,t]$, there holds:

$$R'(s) = -\partial_t Q(t-s, \mathbf{z}+s\mathbf{1}) - \partial_t Q(t-s, s\mathbf{1}) + \mathbf{1}\cdot(\nabla Q(t-s, \mathbf{z}+s\mathbf{1}) - \nabla Q(t-s, s\mathbf{1})) = 0.$$

Hence, $Q(t,\mathbf{z}) = Q(t,\mathbf{z}) - Q(t,\mathbf{O}) = R(0) = R(t) = Q(0,\mathbf{z}+t\mathbf{1}) - Q(0,t\mathbf{1}) = 0$, which is the desired conclusion. □

*Proof of Proposition* 18. In order to derive a general formula for the $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^n_+)$ solution $C_{\mathbf{p}}$ of (2.23), we make a substitution of the spatial variable and use the previous special case described in Lemma 2. To do so, we set, for $t \geq 0$ and $z \in \mathbb{R}^n_+$,

$$Q(t,\mathbf{z}) = C_{\mathbf{p}}(t, \mathbf{y}(\mathbf{z})),$$

where $\mathbf{y}(\mathbf{z}) = (y_1(\mathbf{z}), \ldots, y_n(\mathbf{z}))$ and $y_i(\mathbf{z}) = \tanh(\mu_i z_i)/\mu_i$ for every $1 \leq i \leq n$. Notice that $\mathbf{y}(\mathbf{z}) \in \mathbb{R}^n_+$ for every $\mathbf{z} \in \mathbb{R}^n_+$. The function $Q$ is of class $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^n_+)$ and:

$$\mathbf{1}\cdot\nabla Q(t,\mathbf{z}) = \sum_{i=1}^n \partial_{z_i} Q(t,\mathbf{z}) = \sum_{i=1}^n (1-\tanh^2(\mu_i z_i))\partial_{z_i} C_{\mathbf{p}}(t,\mathbf{y}(\mathbf{z})) = A(\mathbf{y}(\mathbf{z}))\cdot\nabla C_{\mathbf{p}}(t,\mathbf{y}(\mathbf{z})),$$

for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}^n_+$, where $A$ is given in (2.24). As $\overline{m}(t) = \mathbf{1}\cdot\nabla C_{\mathbf{p}}(t,\mathbf{O}) = \mathbf{1}\cdot\nabla Q(t,\mathbf{O})$ and $Q(t,\mathbf{O}) = C_{\mathbf{p}}(t,\mathbf{O}) = \log M_{\mathbf{p}}(t,\mathbf{O}) = 0$ by (2.15) and (2.21), it follows from (2.23) that:

$$\left\{ \begin{array}{rcl} \partial_t Q(t,\mathbf{z}) & = & \mathbf{1}\cdot(\nabla Q(t,\mathbf{z}) - \nabla Q(t,\mathbf{O})) - b(\mathbf{y}(\mathbf{z})),\ t \geq 0,\ \mathbf{z} \in \mathbb{R}^n_+, \\ Q(t,\mathbf{O}) & = & 0,\ t \geq 0, \end{array} \right.$$

and $Q(0,\mathbf{z}) = C_{\mathbf{p}_0}(\mathbf{y}(\mathbf{z}))$ for all $\mathbf{z} \in \mathbb{R}^n_+$. The functions $C_{\mathbf{p}_0} \circ \mathbf{y}$ and $\widetilde{b} := b \circ \mathbf{y}$ are of class $C^1(\mathbb{R}^n_+)$ and $C_{\mathbf{p}_0}(\mathbf{y}(\mathbf{O})) = \widetilde{b}(\mathbf{O}) = 0$. Therefore, Lemma 2 implies that:

$$Q(t,\mathbf{z}) = \int_0^t [b(\mathbf{y}(s\mathbf{1})) - b(\mathbf{y}(\mathbf{z}+s\mathbf{1}))]\ \mathrm{d}s + C_{\mathbf{p}_0}(\mathbf{y}(\mathbf{z}+t\mathbf{1})) - C_{\mathbf{p}_0}(\mathbf{y}(t\mathbf{1})), \quad (2.63)$$

for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}^n_+$. Consider now any $t \in \mathbb{R}_+$ and $\mathbf{z} = (z_1, \ldots, z_n) \in [0, 1/\mu_1) \times \cdots \times [0, 1/\mu_n)$. Set:

$$\mathbf{z}' = \left( \frac{\operatorname{atanh}(\mu_1 z_1)}{\mu_1}, \ldots, \frac{\operatorname{atanh}(\mu_1 z_1)}{\mu_1} \right) \in \mathbb{R}^n_+,$$

and observe that $\mathbf{y}(\mathbf{z}') = \mathbf{z}$. Hence, we have:

$$\begin{array}{rcl} C_{\mathbf{p}}(t,\mathbf{z}) & = & C_{\mathbf{p}}(t,\mathbf{y}(\mathbf{z}')), \\ & = & Q(t,\mathbf{z}') = \displaystyle\int_0^t [b(\mathbf{y}(s\mathbf{1})) - b(\mathbf{y}(\mathbf{z}'+s\mathbf{1}))]\ \mathrm{d}s + C_{\mathbf{p}_0}(\mathbf{y}(\mathbf{z}'+t\mathbf{1})) - C_{\mathbf{p}_0}(\mathbf{y}(t\mathbf{1})), \end{array}$$

which leads straightforwardly to the formulae (2.26)-(2.27). Furthermore, for every $t \in \mathbb{R}_+$, the formula $\overline{m}(t) = \mathbf{1}\cdot\nabla Q(t,\mathbf{O})$ together with (2.63) easily

yields (2.28). The proof of Proposition 18 is thereby complete. □

*Proof of Corollary 2.* Let $p$ be the fitness distribution, that is, $p(t,m)\,\mathrm{d}m$ is the pushforward measure of $q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ by the map $\mathbf{x} \mapsto -\|\mathbf{x}\|^2/2$. Let:

$$M_p(t,z) = \int_{-\infty}^{0} e^{zm}\, p(t,m)\,\mathrm{d}m,$$

be the moment generating function of $p$. As the fitness $m \in \mathbb{R}_-$ is the sum of the fitness components $(m_1,\dots,m_n) \in \mathbb{R}_-^n$, we have:

$$M_p(t,z) = \int_{-\infty}^{0} e^{zm}\, p(t,m)\,\mathrm{d}m = \int_{\mathbb{R}_-^n} e^{z(m_1+\cdots+m_n)}\, \mathbf{p}(t,\mathbf{m})\,\mathrm{d}\mathbf{m} = M_{\mathbf{p}}(t,z\mathbf{1}),$$

for all $(t,z) \in \mathbb{R}_+ \times \mathbb{R}_+$. This implies that $C_p(t,z) = C_{\mathbf{p}}(t,z\mathbf{1})$ for all $(t,z) \in \mathbb{R}_+ \times \mathbb{R}_+$ and that $C_p$ is of class $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+)$, with initial condition $C_p(0,\cdot) = C_{\mathbf{p}_0}(\cdot\mathbf{1})$. Thanks to Eqs. (2.23)-(2.24) satisfied by $C_{\mathbf{p}}$, it follows that:

$$\begin{aligned}\partial_t C_p(t,z) &= \sum_{i=1}^{n}\left[(1-\mu^2 z^2)\partial_{z_i} C_{\mathbf{p}}(t,z\mathbf{1}) - \partial_{z_i} C_{\mathbf{p}}(t,\mathbf{O}) - \frac{\mu^2}{2}\,z\right],\\ &= (1-\mu^2 z^2)\partial_z C_p(t,z) - \partial_z C_p(t,0) - \frac{n}{2}\mu^2 z,\end{aligned}$$

for all $(t,z) \in \mathbb{R}_+ \times \mathbb{R}_+$. This is the desired result and the proof of Corollary 2 is thereby complete. □

*Proof of Corollary 3.* We have seen in the proof of Corollary 2 that $C_p(t,z) = C_{\mathbf{p}}(t,z\mathbf{1})$ for all $(t,z) \in \mathbb{R}_+ \times \mathbb{R}_+$. Thus, formulae (2.26)-(2.28) straightforwardly yield (2.29)-(2.30) for $t \geq 0$ and $z \in [0, 1/\mu)$. □

### 2.4.4. Stationary states

*Proof of Theorem 19.* We use the notations of Proposition 18. Let $\mathbf{z} \in [0,1/\mu_1) \times \cdots \times [0,1/\mu_n)$. As $\psi(t,\mathbf{z}) \to (\mu_1^{-1},\dots,\mu_n^{-1})$ as $t \to +\infty$, the continuity of $C_{\mathbf{p}_0}$ and the formulae (2.26)-(2.27) yield that:

$$C_{\mathbf{p}}(t,\mathbf{z}) \to \frac{1}{2}\sum_{i=1}^{n} \log\left[\exp\left(-\operatorname{atanh}\left(\mu_i z_i\right)\right)\,\cosh\left(\operatorname{atanh}\left(\mu_i z_i\right)\right)\right] = -\frac{1}{2}\sum_{i=1}^{n}\log(1+\mu_i z_i), \tag{2.64}$$

as $t \to +\infty$. It then follows from the generalization of the Curtiss theorem (Yakymiv 2011) that, if the limit as $t \to +\infty$ of the cumulant generating functions $C_{\mathbf{p}}(t,\cdot)$ is the cumulant generating function of $\mathbf{p}_\infty$ given by (2.31) in some subset of $\mathbb{R}_-^n$ with non-empty interior, then the distributions $\mathbf{p}(t,\cdot)$ weakly converge to $\mathbf{p}_\infty$ as $t \to +\infty$. So let us compute the CGF of $\mathbf{p}_\infty$. For all $\mathbf{z} \in \mathbb{R}_+^n$, Fubini's theorem

yields that:

$$\begin{aligned}\int_{\mathbb{R}^n_-} e^{\mathbf{z}\cdot\mathbf{m}}\mathbf{p}_\infty(\mathbf{m})\,\mathrm{d}\mathbf{m} &= \frac{1}{\pi^{n/2}}\left(\prod_{i=1}^n \mu_i^{-1/2}\right)\int_{\mathbb{R}^n_-}\left(\prod_{i=1}^n |m_i|^{-1/2}\right)\exp\left(\sum_{i=1}^n m_i/\mu_i\right)e^{\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{m},\\ &= \frac{1}{\pi^{n/2}}\prod_{i=1}^n\left[\mu_i^{-1/2}\int_{-\infty}^0 |m_i|^{-1/2}\exp\left((1+\mu_i z_i)\,m_i/\mu_i\right)\mathrm{d}m_i\right],\\ &= \frac{1}{\pi^{n/2}}\prod_{i=1}^n\left[\frac{1}{\sqrt{1+\mu_i z_i}}\int_0^{+\infty}|x_i|^{-1/2}\,e^{-x_i}\,\mathrm{d}x_i\right] = \prod_{i=1}^n (1+\mu_i z_i)^{-1/2}.\end{aligned}$$

Hence, the CGF of $\mathbf{p}_\infty$ is equal to the function:

$$\mathbf{z} = (z_1,\ldots,z_n) \mapsto -(1/2)\sum_{i=1}^n \log\left(1+\mu_i z_i\right),$$

that is, the limit in (2.64). As a consequence, the distributions $\mathbf{p}(t,\cdot)$ weakly converge to $\mathbf{p}_\infty$ in $\mathbb{R}^n_-$ as $t\to+\infty$.

On the other hand, thanks to Proposition 18, we also know that:

$$\overline{m}(t) = A(\psi(t,\mathbf{O}))\cdot\nabla C_{\mathbf{p}_0}(\psi(t,\mathbf{O})) - b(\psi(t,\mathbf{O})),$$

for every $t\geq 0$, with $A(\mathbf{z}) = (1-\mu_1^2 z_1^2,\ldots,1-\mu_n^2 z_n^2)$ and $b(\mathbf{z}) = \sum_{i=1}^n \mu_i^2 z_i/2$. Notice that $\psi(t,\mathbf{O})\to(\mu_1^{-1},\ldots,\mu_n^{-1})$, $A(\psi(t,\mathbf{O}))\to\mathbf{O}$ and $b(\psi(t,\mathbf{O}))\to\sum_{i=1}^n\mu_i/2$, as $t\to+\infty$. Hence, $\overline{m}(t)\to-\sum_{i=1}^n\mu_i/2 =: \overline{m}_\infty$, as $t\to+\infty$. It is also straightforward to check that:

$$\sum_{i=1}^n\int_{\mathbb{R}^n_-} m_i\,\mathbf{p}_\infty(\mathbf{m})\,\mathrm{d}\mathbf{m} = \overline{m}_\infty,$$

and that $\mathbf{p}_\infty$ is a classical $C^2((\mathbb{R}^*_-)^n)$ solution of (2.32) (this property is also a consequence of the fact that $\mathbf{p}$ satisfies (2.17) and the distributions $\mathbf{p}(t,\cdot)$ weakly converge to $\mathbf{p}_\infty$ as $t\to+\infty$). The proof of Theorem 19 is thereby complete. □

*Proof of Corollary 4.* By the same arguments as in the proof of Theorem 19, thanks to (2.29), we see that, for all $z\in[0,1/\mu)$, $C_p(t,z)\to-(1/2)\log(1+\mu\,z)$ as $t\to+\infty$. This limiting function corresponds to the cumulant generating function of a random variable distributed according to $-\Gamma(n/2,\mu)$. Since a distribution is uniquely determined by its cumulant generating function, this implies that $p_\infty$ is the probability density function of this random variable, *i.e.*, for all $m<0$,

$$p_\infty(m) = \frac{1}{\Gamma(n/2)\,\mu^{n/2}}\,|m|^{\frac{n}{2}-1}\exp\left(\frac{m}{\mu}\right),$$

where $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}\mathrm{d}t$ is the standard Gamma function. □

*Proof of Corollary 5.* We assume here that $q_0$ is #-symmetric, that is, $q_0 = q_0^{\#}$.

As already emphasized in Section 2.2.1, the uniqueness for problem (2.10) implies that $q(t,\cdot)$ is also $\#$-symmetric at each time $t \geq 0$. Proposition 14 (or formula (2.16)) yields:

$$q(t,\mathbf{x}) = 2^{-n}|x_1\cdots x_n|\,p\Big(t,-\frac{x_1^2}{2},\dots,-\frac{x_n^2}{2}\Big),$$

for all $(t,\mathbf{x}) \in \mathbb{R}_+ \times (\mathbb{R}^*)^n$ and, for each function $\phi \in C_c^\infty((\mathbb{R}^*)^n)$:

$$\begin{aligned}\int_{\mathbb{R}^n} q(t,\mathbf{x})\,\phi(\mathbf{x})\,\mathrm{d}\mathbf{x} &= 2^{-n}\int_{\mathbb{R}^n} |x_1\cdots x_n|\,p\Big(t,-\frac{x_1^2}{2},\dots,-\frac{x_n^2}{2}\Big)\,\phi(\mathbf{x})\,\mathrm{d}\mathbf{x},\\ &= \int_{\mathbb{R}^n_+} |x_1\cdots x_n|\,p\Big(t,-\frac{x_1^2}{2},\dots,-\frac{x_n^2}{2}\Big)\,\phi^{\#}(\mathbf{x})\,\mathrm{d}\mathbf{x},\\ &= \int_{\mathbb{R}^n_-} \mathbf{p}(t,\mathbf{m})\,\phi^{\#}(\mathbf{x}^{\mathbf{1}}(\mathbf{m}))\,\mathrm{d}\mathbf{m},\\ &\to \int_{\mathbb{R}^n_-} \mathbf{p}_\infty(\mathbf{m})\,\phi^{\#}(\mathbf{x}^{\mathbf{1}}(\mathbf{m}))\,\mathrm{d}\mathbf{m} = \int_{\mathbb{R}^n} q_\infty(\mathbf{x})\,\phi(\mathbf{x})\,\mathrm{d}\mathbf{x},\end{aligned}$$

as $t \to +\infty$, with $\mathbf{x}^{\mathbf{1}}(\mathbf{m}) = (\sqrt{-2m_1},\dots,\sqrt{-2m_n})$ and:

$$q_\infty(\mathbf{x}) = 2^{-n}\,|x_1\cdots x_n|\,\mathbf{p}_\infty\Big(-\frac{x_1^2}{2},\dots,-\frac{x_n^2}{2}\Big).$$

The above formula corresponds to (2.33). Furthermore, since $\int_{\mathbb{R}^n} q_\infty(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$ and since $\int_{\mathbb{R}^n} q(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$, for every $t \geq 0$, it then easily follows from the previous estimates that:

$$\int_{\mathbb{R}^n} q(t,\mathbf{x})\,\phi(\mathbf{x})\,\mathrm{d}\mathbf{x} \underset{t\to+\infty}{\longrightarrow} \int_{\mathbb{R}^n} q_\infty(\mathbf{x})\,\phi(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

for every $\phi \in C_c^\infty(\mathbb{R}^n)$. In other words, the distributions $q(t,\cdot)$ weakly converge in $\mathbb{R}^n$ to $q_\infty$ as $t \to +\infty$. Lastly, the formula $\lim_{t\to+\infty}\overline{m}(t) = \overline{m}_\infty = \int_{\mathbb{R}^n} m(\mathbf{x})\,q_\infty(\mathbf{x})\,\mathrm{d}\mathbf{x}$ is a consequence of the previous arguments and Theorem 19. The proof of Corollary 5 is thereby complete. □

### 2.4.5. Plateaus: proofs of Proposition 20 and Remark 1

*Proof of Proposition 20.* We show in this section that, given an initial phenotype $\mathbf{x}_0 = (x_{0,1},\dots,x_{0,n}) \in \mathbb{R}^n$, a value $\mu_1 > 0$, and a duration $T > 0$, we can choose some positive real numbers $\mu_2,\dots,\mu_n$ (here, $n \geq 2$) such that the mean fitness:

$$t \mapsto \overline{m}(t) := \sum_{i=1}^{n}\Big(\frac{x_{0,i}^2}{2}\Big(\tanh^2(\mu_i t) - 1\Big) - \frac{\mu_i}{2}\tanh(\mu_i t)\Big),$$

is close to each of the plateaus:

$$\overline{m}_{\infty,k} = -\sum_{i=k+1}^{n} \frac{x_{0,i}^2}{2} - \sum_{i=1}^{k} \frac{\mu_i}{2}, \quad \text{for } k = 1, \dots, n-1,$$

at least during a time interval of duration $T$. We recall that the above formula for $\overline{m}(t)$ corresponds to the limit of formula (2.28), when the initial conditions $q_0$ approach the Dirac distribution $\delta_{\mathbf{x}_0}$.

More precisely, we are given $\mathbf{x}_0 = (x_{0,1}, \dots, x_{0,n}) \in \mathbb{R}^n$, $T > 0$, $\varepsilon > 0$ and $\mu_1 > 0$, and we shall prove the existence of $(\mu_2, \dots, \mu_n) \in (\mathbb{R}_+^*)^{n-1}$ such that, for each $k \in [\![1, n-1]\!]$, the set:

$$\Big\{ t \geq 0, \ |\overline{m}(t) - \overline{m}_{\infty,k}| \leq \varepsilon \Big\}, \tag{2.65}$$

contains an interval of length at least equal to $T$. In that respect, we firstly define some functions:

$$s_i : (\mu, t) \in \mathbb{R}_+^* \times \mathbb{R}_+ \mapsto \left| \frac{x_{0,i}^2}{2} \tanh(\mu\, t) - \frac{\mu}{2} \right| \ \text{ for } i = 1, \dots, n-1.$$

Secondly, by iteration for $k = 1, \dots, n-1$, we can then define:

- ◇ a function $S_k : t \in \mathbb{R}_+ \mapsto \displaystyle\sum_{i=1}^{k} \left| \frac{x_{0,i}^2}{2} \Big( \tanh(\mu_i t) + 1 \Big) - \frac{\mu_i}{2} \right|$,
- ◇ a time $\tau_k > \tau_{k-1} + T$, (with $\tau_0 = -T$) such that:

$$\Big( \max_{t \in [\tau_k, \tau_k + T]} S_k(t) \Big) \times (1 - \tanh(\mu_k \tau_k)) \leq \varepsilon \frac{n+k}{n}, \tag{2.66}$$

- ◇ a real number $\mu_{k+1} \in (0, \mu_k)$ such that:

$$\forall\, i \in [\![1, k]\!], \ \forall\, t \in [\tau_i, \tau_i + T], \quad s_{k+1}(\mu_{k+1}, t) \tanh(\mu_{k+1}(\tau_i + T)) \leq \frac{\varepsilon}{2n}.$$

Note that the last property implies that:

$$\forall\, k \in [\![1, n-1]\!], \ \forall\, i \in [\![k+1, n]\!], \ \forall\, t \in [\tau_k, \tau_k + T], \ s_i(\mu_i, t) \tanh(\mu_i(\tau_k + T)) \leq \frac{\varepsilon}{2n}. \tag{2.67}$$

Fix now an index $k \in [\![1, n-1]\!]$ and a time $t \in [\tau_k, \tau_k + T]$ $(\subset \mathbb{R}_+^*)$. There holds:

$$\begin{aligned} |\overline{m}(t) - \overline{m}_{\infty,k}| &= \Big| \sum_{i=1}^{k} \Big( \frac{x_{0,i}^2}{2} \left( \tanh^2(\mu_i t) - 1 \right) - \frac{\mu_i}{2} \left( \tanh(\mu_i t) - 1 \right) \Big) \\ &\qquad + \sum_{i=k+1}^{n} \Big( \frac{x_{0,i}^2}{2} \tanh^2(\mu_i t) - \frac{\mu_i}{2} \tanh(\mu_i t) \Big) \Big|, \\ &\leq \sum_{i=1}^{k} \Big| \frac{x_{0,i}^2}{2} \left( \tanh(\mu_i t) + 1 \right) - \frac{\mu_i}{2} \Big| \left( 1 - \tanh(\mu_i t) \right) \\ &\qquad + \sum_{i=k+1}^{n} \Big| \frac{x_{0,i}^2}{2} \tanh(\mu_i t) - \frac{\mu_i}{2} \Big| \tanh(\mu_i t). \end{aligned} \tag{2.68}$$

As $t \in [\tau_k, \tau_k + T]$, we have $0 < 1 - \tanh(\mu_i t) \leq 1 - \tanh(\mu_k \tau_k)$ for every $i \in [\![1, k]\!]$ (remember that $0 < \mu_k < \mu_{k-1} < \cdots < \mu_1$), whereas $\tanh(\mu_i\, t) \leq \tanh(\mu_i\, (\tau_k + T))$ for every $i \in [\![k+1, n]\!]$. It then follows from (2.66)-(2.68) that, for every $k \in [\![1, n-1]\!]$ and $t \in [\tau_k, \tau_k + T]$:

$$|\overline{m}(t) - \overline{m}_{\infty,k}| \leq S_k(t) \left( 1 - \tanh(\mu_k \tau_k) \right) + \sum_{i=k+1}^{n} s_i(\mu_i, t) \tanh(\mu_i\, (\tau_k + T)) \leq \varepsilon \frac{n+k}{2n} + \sum_{i=k+1}^{n} \frac{\varepsilon}{2n} = \varepsilon.$$

Thus, with this choice of $(\mu_1, \ldots, \mu_n) \in (\mathbb{R}_+^*)^n$, each set defined in (2.65) contains an interval of length at least equal to $T$. This proves Proposition 20. □

*Proof of results in Remark 1.* Take $n = 2$, $\mu_1 = 1$, $\mu_2 = 10^{-k}$ for some $k \geq 1$ and $x_{0,1} = x_{0,2} = 1$. Differentiating the expression (2.34), we observe that:

$$\overline{m}'(t) = \left( \tanh(t) - \frac{1}{2} \right) \left( 1 - \tanh^2(t) \right) + \mu_2 \left( \tanh(\mu_2\, t) - \frac{\mu_2}{2} \right) \left( 1 - \tanh^2(\mu_2\, t) \right).$$

We observe that, as $t \to 0$, $\overline{m}'(t) \to -1/2 - 10^{-2\,k}/2$. Then $\overline{m}'(t)$ rapidly increases, to reach significantly positive values, *e.g.*, at $t_0 = \ln(3)$, $\overline{m}'(t_0) > 27/250 - \mu_2/2 > 0.1 - 10^{-k}/2$.

Then, consider the interval:

$$I := \left( \ln\left( 2 \cdot 10^k \right), 10^{k/2} \right).$$

For each $t$ in $I$, $1 - \tanh^2(t) < 4\, \exp(-2\, t) < 10^{-2\,k}$ and $\tanh(\mu_2\, t) - \mu_2/2 < \mu_2\, t < 10^{-k/2}$. Thus,

$$|\overline{m}'(t)| \leq 10^{-2\,k} + 10^{-3\,k/2},$$

which means that $\overline{m}(t)$ remains stable within this interval, corresponding to a part of the plateau. Latter on, $\overline{m}'(t)$ reaches again significantly positive values which are significantly higher than in this interval, *e.g.*, at $t_1 = 10^k\, \ln[(\sqrt{2} + \sqrt{6})/2] >$

$10^{k/2}$, straightforward computations show that:

$$\begin{aligned}\overline{m}'(t_1) &\geq \mu_2 \left(\tanh(\mu_2\, t_1) - \tfrac{\mu_2}{2}\right)\, (1 - \tanh^2(\mu_2\, t_1)),\\ &= 2 \cdot 10^{-k} \left[\tfrac{(2+\sqrt{3})(2+2\sqrt{3})}{(3+\sqrt{3})^3} - 10^{-k}\, \tfrac{(2+\sqrt{3})}{(3+\sqrt{3})^2}\right],\end{aligned}$$

which show that $\overline{m}'(t)$ is of order $10^{-k}$ at $t_1$, versus $10^{-3k/2}$ in the interval I. □

## 2.5. A formal derivation of the diffusive approximation of the mutation effects

The goal of this appendix is to give a formal justification of the diffusion term in (2.4). The case $n = 1$ is classical and can be found *e.g.* in Roques 2013; Turchin 1998. The anisotropic case $n \geq 2$ is less standard, but it will easily follow from the same arguments.

Namely, we assume that the mutation effects on phenotypes follow a normal distribution $\mathcal{N}(0, \Lambda)$, with $\Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$ and $\lambda_i > 0$ for each $i \in [\![1, n]\!]$, and that these mutations occur with a rate $U > 0$. In other words, the dynamics of the phenotype distribution under the mutation effects only (*i.e.*, without selection) can be described by an integro-differential equation:

$$\partial_t q(t, \mathbf{x}) = U\, (J \star q - q)(t, \mathbf{x}),\ t \geq 0,\ \mathbf{x} \in \mathbb{R}^n,$$

where $\star$ is the standard convolution product in $\mathbb{R}^n$ defined by:

$$(J \star q)(t, \mathbf{x}) = \int_{\mathbb{R}^n} q(t, x_1 - y_1, \ldots, x_n - y_n)\, J(y_1, \ldots, y_n)\, \mathrm{d}y_1 \ldots \mathrm{d}y_n,$$

and $J$ the (Gaussian) probability density function associated with the normal distribution $\mathcal{N}(0, \Lambda)$.

Formally, by writing a Taylor expansion of $q(t, \mathbf{x} - \mathbf{y})$ at $\mathbf{x}$:

$$q(t, \mathbf{x} - \mathbf{y}) = \sum_{k_1, \ldots, k_n = 0}^{\infty} (-1)^{k_1 + \cdots + k_n}\, \frac{y_1^{k_1} \cdots y_n^{k_n}}{k_1! \cdots k_n!}\, \frac{\partial^{k_1 + \cdots + k_n} q}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(t, \mathbf{x}),$$

and by defining the central moments of the normal distribution:

$$\omega_{k_1, \ldots, k_n} = \int_{\mathbb{R}^n} y_1^{k_1} \cdots y_n^{k_n}\, J(y_1, \ldots, y_n)\, \mathrm{d}y_1 \ldots \mathrm{d}y_n,$$

we then get that:

$$(J \star q)(t, \mathbf{x}) = \sum_{k_1, \ldots, k_n = 0}^{\infty} (-1)^{k_1 + \cdots + k_n}\, \frac{\omega_{k_1, \ldots, k_n}}{k_1! \cdots k_n!}\, \frac{\partial^{k_1 + \cdots + k_n} q}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(t, \mathbf{x}).$$

Since $\omega_{k_1,\ldots,k_n} = 0$ if at least one of the $k_i$'s is odd, and since:

$$\omega_{k_1,\ldots,k_n} = \Big( \prod_{i=1}^{n} \frac{k_i!}{2^{k_i/2}(k_i/2)!} \Big) \times \Big( \prod_{i=1}^{n} \lambda_i^{k_i/2} \Big),$$

otherwise, one infers in particular that the second-order moments with even indexes are such that $\omega_{0,\ldots,0,k_i=2,0,\ldots,0} = \lambda_i$. Assuming that $\max\limits_{1 \leq i \leq n} \lambda_i \ll 1$, we may formally neglect the moments of order $k_1 + \cdots + k_n \geq 4$, leading to:

$$(J \star q)(t, \mathbf{x}) \approx q(t, \mathbf{x}) + \sum_{i=1}^{n} \frac{\lambda_i}{2} \partial_{ii} q(t, \mathbf{x}).$$

Finally, setting $\mu_i^2 = U \, \lambda_i$, we obtain:

$$U \, (J \star q - q)(t, \mathbf{x}) \approx \sum_{i=1}^{n} \frac{\mu_i^2}{2} \partial_{ii} q(t, \mathbf{x}).$$

3

# Adaptation in an anisotropic mobile phenotype-fitness landscape

**F. Lavigne**[a,b,c]

[a] BioSP, INRA, 84914, Avignon, France
[b] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
[c] ISEM (UMR 5554), CNRS, 34095, Montpellier, France

**Abstract**

Evolution of an asexual population, as viruses, bacteria and cancer cells, depends on some external factors, as antibiotic concentrations. Because of adaptation under mutations and selection, this pathogen can manage to resist to drugs, which yields an important issue to the understanding of the impact of the environmental changes.

This paper is devoted to model this problem with a nonlocal PDE, describing the dynamics of such a phenotypically structured population, in a changing environment. The large-time behaviour of this model, with particular forms of environmental changes (linear or periodically fluctuations), has been previously studied. A recent paper (Roques; Patout; Bonnefon; Martin 2020) has developed a new mathematical approach for this problem, considering a very general form of environmental variations, and giving an analytic description of the full trajectories of adaptation. This study has limited to isotropic mutations.

However, recent studies have shown that considering anisotropic mutation kernel can change the evolutionary dynamics of the population (Hamel; Lavigne; Martin; Roques 2020). For example, when the mutation variances have different scaling, some evolutionary plateaus can appear.

Thus the aim of this paper is to mix the two previous studies, with an anisotropic mutation kernel, and a changing environment. The main idea is to study a multivariate distribution of $(n+1)$ "fitness components", solving a degenerate parabolic equation. Its generating function solves so a transport equation, and describes the distribution of fitness at any time. To check the curracy of this approach, we use stochastic individual-based simulations.

# Sommaire

## 3.1. Introduction

External factors can have different impact on the evolution of an asexual population, as viruses, bacteria and cancer cells. Because of genetic adaptation, the growth rate of pathogens can become positive, and so this pathogen can manage to resist to drugs. Understand the impact of the environmental changes on the evolution is so an important issue to prevent drug resistance, which can be approximated by mathematical model.

Different papers have already studied the dynamics of phenotypic distribution in a steady environment (Desai; Fisher 2011; Gerrish; Colato; Perelson; Sniegowski 2007; Gil; Hamel; Martin; Roques 2017; Gil; Hamel; Martin; Roques 2019; Hamel; Lavigne; Martin; Roques 2020; Sniegowski; Gerrish 2010; Tsimring; Levine; Kessler 1996). A phenotype is modeled by a set of $n$ biological traits, *i.e.*, by a vector $\mathbf{x} \in \mathbb{R}^n$. Many recent studies have modeled the distribution of the biological traits $q(t, \mathbf{x})$ as a solution of an integro-differential equation (IDE) (Desai; Fisher 2011; Gerrish; Colato; Perelson; Sniegowski 2007; Gil; Hamel; Martin; Roques 2017; Gil; Hamel; Martin; Roques 2019; Sniegowski; Gerrish 2010). This model is based on the following IDE:

$$\partial_t q(t, \mathbf{x}) = U\left[J * q(t, \mathbf{x}) - q(t, \mathbf{x})\right] + (m(\mathbf{x}) - \overline{m}(t))q(t, \mathbf{x}), \quad \forall t > 0, \forall \mathbf{x} \in \mathbb{R}^n, \tag{3.1}$$

where $J * q$ is the classic convolution of $q$ with a given mutational kernel $J$, $U > 0$ is the mutation rate, $m(\mathbf{x})$ is the Malthusian fitness, and the mean fitness $\overline{m}(t)$ is defined by:

$$\overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})\, q(t, \mathbf{x})\, \mathrm{d}\mathbf{x}.$$

The Malthusian fitness is connected with the Malthusian growth rate $r$ thanks to a constant $r_{\max} > 0$ (growth rate of an optimum phenotype) and the relation $r(\mathbf{x}) = r_{\max} + m(\mathbf{x})$. Unfortunately, we do not find any analytic results on the solution of (3.1), except for asymptotic properties (see for example Gil; Hamel; Martin; Roques 2017).

The mutational term $U(J * q - q)$ can be approached by a diffusion term (for more details see for example Appendix, Hamel; Lavigne; Martin; Roques 2020), which modified the IDE (3.1) into a parabolic partial differential equation:

$$\partial_t q(t, \mathbf{x}) = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\, \partial_{x_i x_i} q(t, \mathbf{x}) + (m(\mathbf{x}) - \overline{m}(t))\, q(t, \mathbf{x}), \ t > 0, \ \mathbf{x} \in \mathbb{R}^n, \tag{3.2}$$

Making assumption on the fitness function $m$, Alfaro; Carles 2017; Alfaro; Veruete 2018 (in the one dimension case) and Hamel; Lavigne; Martin; Roques 2020 (in the general $n$ dimension case) found general formula of (3.2), for any initial condition $q_0$. The most important case for us is the Fisher's geometrical model (FGM; a single peak phenotype-fitness landscape), which has shown promising potential when compared to empirical measures of fitness epistasis.

According to this model, the fitness $m$ of a phenotype $\mathbf{x} \in \mathbb{R}^n$ decreases away from an optimum $\mathbf{O} \in \mathbb{R}^n$, and the decrease is assumed to be quadratic, *i.e.*:

$$\forall \mathbf{x} \in \mathbb{R}^n,\ m(\mathbf{x}) = -\frac{\|\mathbf{x} - \mathbf{O}\|^2}{2},$$

with $\|\cdot\|$ the standard Euclidian norm in $\mathbb{R}^n$.

With an isotropic dependence between the phenotype $\mathbf{x} \in \mathbb{R}^n$ and the fitness $m(\mathbf{x})$, and with an isotropic effect of mutations on phenotypes (*i.e.*, $\mu_1 = \cdots = \mu_n$), Martin; Roques 2016 have proven some analytic results, thanks to an analysis of a PDE satisfied by the generating function of the fitness distribution. However, fitting the analytic formula of $\overline{m}(t)$ to experimental data based on *E. Coli* evolution (Lenski; Rose; Simpson; Tadler 1991; Lenski; Travisano 1994; Wiser; Ribeck; Lenski 2013), we can see that this model is not realistic: a second growth of the mean fitness can be observed in the data (see Hamel; Lavigne; Martin; Roques 2020). However, one hypothesis can explain this second growth. The authors of Hamel; Lavigne; Martin; Roques 2020 have studied the same model as Martin; Roques 2016, without the assumption of isotropic effect of mutations. Fitting now the new formula of the mean fitness $\overline{m}(t)$ given by Hamel; Lavigne; Martin; Roques 2020 gives a better approximation than the isotropic model: the adjusted $R^2$ is equal to 0.89 for the anisotropic model, whereas it is 0.57 for the isotropic one. Moreover, the anisotropic model can imply some evolutive plateaus, as Hamel; Lavigne; Martin; Roques 2020 have shown, which cannot be seen with an isotropic effect of mutations. This implication leads the need to a better understanding of the anisotropic model.

Taking now into account the environmental changes, the optimum $\mathbf{O}$ is considered as a point $\mathbf{O}(t) \in \mathbb{R}^n$, which can move over time. Thus the fitness $m(\mathbf{x})$ in (3.2) has to be transformed into:

$$m(t, \mathbf{x}) = -\frac{\|\mathbf{x} - \mathbf{O}(t)\|^2}{2}, \tag{3.3}$$

with $\mathbf{O}(t) \in C(\mathbb{R}_+, \mathbb{R}^n)$. For the sake of simplicity, we will make the hypothesis that the optimum moves just in the direction of one axis. Without loss of generality, we assume in the sequel that $\mathbf{O}(0) = (0, \ldots, 0)$ and that $\mathbf{O}(t) = \Delta(t)\mathbf{e}_1$, with $\Delta \in \mathcal{C}(\mathbb{R}_+)$ and $\mathbf{e}_1 = (1, 0, \ldots, 0) \in \mathbb{R}^n$. Thus the equation modeling the evolution of the phenotypic distribution becomes:

$$\partial_t q(t, \mathbf{x}) = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\, \partial_{x_i x_i} q(t, \mathbf{x}) + (m(t, \mathbf{x}) - \overline{m}(t))\, q(t, \mathbf{x}),\ t > 0,\ \mathbf{x} \in \mathbb{R}^n, \tag{3.4}$$

with:

$$\overline{m}(t) = \int_{\mathbb{R}^n} m(t, \mathbf{x})\, q(t, \mathbf{x})\, \mathrm{d}\mathbf{x}.$$

While Hamel; Lavigne; Martin; Roques 2020 has focused on the distribution

of the fitness $\mathbf{m}$, we cannot make the same computations here. In fact, we can explicit a relation between this distribution and $q(t,\mathbf{x})$ as in Hamel; Lavigne; Martin; Roques 2020, but we cannot find a closed differential equation satisfied by this. To understand the adaptation dynamics, we define a transformed fitness components vector $\mathbf{v} = (v_0, v_1, \dots, v_n) \in \mathbb{R}^{n+1}$ with:

$$v_0(\mathbf{x}) = x_1 \in \mathbb{R}, \quad \text{and} \quad \forall 1 \leq i \leq n, v_i(\mathbf{x}) = -\frac{x_i^2}{2} \in \mathbb{R}_-,$$

and so its joint distribution $\mathbf{p}(t,\mathbf{v})$ is the pushforward measure of the measure $q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ by the map $\mathbf{x} \mapsto (x_1, -x_1^2/2, \dots, -x_n^2/2)$. In the isotropic case (Roques; Patout; Bonnefon; Martin 2020), we can directly focus on the equation satisfied by the distribution of a fitness vector $\mathbf{m}$ and neither by $q(t,\mathbf{x})$ nor $\mathbf{p}(t,\mathbf{v})$. We can check that the mean fitness $\overline{m}(t)$ can be easily connected with $\mathbf{p}$ by:

$$\overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x}) q(t,\mathbf{x}) \mathrm{d}\mathbf{x} = -\frac{\Delta(t)^2}{2} + \int_{\Gamma_{n+1}} \boldsymbol{\alpha}(t) \cdot \mathbf{v}\, \mathbf{p}(t,\mathbf{v}) \mathrm{d}\mathbf{v},$$

with $\boldsymbol{\alpha}(t) = (\Delta(t), 1, \dots, 1) \in \mathbb{R}^{n+1}$ and:

$$\Gamma_{n+1} = \{(v_0, \dots, v_n) \in \mathbb{R} \times \mathbb{R}_-^n,\ v_1 = -v_0^2/2\}.$$

In the general case that $\mathbf{O}(t) = (O_1(t), \dots, O_n(t))$ is not only moving in one axis direction, we assume that this study can be generalized taking the new fitness vector $\mathbf{v}' = (v'_1, , \dots, v'_{2n}) \in \mathbb{R}^{2n}$ with:

$$\forall 1 \leq i \leq n,\ v'_i(\mathbf{x}) = x_i \in \mathbb{R}, \quad \text{and} \quad v'_{n+i} = -\frac{x_i^2}{2} \in \mathbb{R}_-.$$

## 3.2. Main results

### 3.2.1. The time-dependent problem

#### Solution of the Cauchy problem associated with equation (3.4) for $q(t,\mathbf{x})$

First, we show that the Cauchy problem admits a unique solution with the following assumptions on the initial distribution $q_0$:

- ◇ For some $\alpha \in (0,1)$, $\|q_0\|_{C^{2+\alpha}(\mathbb{R}^n)} < +\infty$, *i.e.*:

$$q_0 \in C^{2+\alpha}(\mathbb{R}^n); \tag{3.5}$$

- ◇ The function $q_0$ is a probability distribution:

$$\int_{\mathbb{R}^n} q_0(\mathbf{x}) \mathrm{d}\mathbf{x} = 1; \tag{3.6}$$

◇ There exists a function $g : \mathbb{R}_+ \to \mathbb{R}_+$ (with $\mathbb{R}_+ = [0, +\infty)$) such that:

$$g \text{ is non-increasing}, \quad 0 \leq q_0 \leq g(\|\cdot\|) \text{ in } \mathbb{R}^n,$$
$$\forall b > 0,\ \mathbf{x} \mapsto \exp(b\|\mathbf{x}\|) g(\|\mathbf{x}\|) \text{ is bounded in } \mathbb{R}^n$$
$$\text{and } \int_{\mathbb{R}^n} \exp(b\|\mathbf{x}\|)\, g(\|\mathbf{x}\|)\, \mathrm{d}\mathbf{x} < +\infty. \quad (3.7)$$

*These assumptions are made throughout the paper, and are therefore not repeated in the statements of the results below.*

By the same arguments as in Hamel; Lavigne; Martin; Roques 2020, the distribution of phenotypes is well-defined: the Cauchy problem admits a unique solution.

**Theorem 22. (Aronson; Besala 1967; Chabrowski 1970)**
*There exists a unique nonnegative solution $q \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ of:*

$$\begin{cases} \partial_t q(t,\mathbf{x}) &= \displaystyle\sum_{i=1}^n \frac{\mu_i^2}{2} \partial_{x_i x_i} q(t,\mathbf{x}) + (m(t,\mathbf{x}) - \overline{m}(t))\, q(t,\mathbf{x}), \quad t \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \\ q(0,\mathbf{x}) &= q_0(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n, \end{cases} \quad (3.8)$$

*such that $q \in L^\infty((0,T) \times \mathbb{R}^n)$ for all $T > 0$, and the function:*

$$t \mapsto \overline{m}(t) = \int_{\mathbb{R}^n} m(t,\mathbf{x})\, q(t,\mathbf{x})\, \mathrm{d}\mathbf{x},$$

*is real-valued and continuous in $\mathbb{R}_+$. Moreover, we have for all $t \geq 0$, $\int_{\mathbb{R}^n} q(t,\mathbf{x})\, \mathrm{d}\mathbf{x} = 1$.*

As in Hamel; Lavigne; Martin; Roques 2020, it is possible to find an explicit solution of (3.8) in the particular case where the phenotypes are initially Gaussian-distributed.

**Corollary 6.**
*Assume that the initial distribution of phenotype frequencies is Gaussian, that is,*

$$\forall\, \mathbf{x} \in \mathbb{R}^n,\ q_0(\mathbf{x}) = (2\pi)^{-n/2}\Big(\prod_{i=1}^{n}(s_i^0)^{-1/2}\Big)\ \exp\left(-\sum_{i=1}^{n}\frac{(x_i-\overline{q}_i^0)^2}{2s_i^0}\right), \tag{3.9}$$

*for some parameters $\overline{q}_i^0 \in \mathbb{R}$ and $s_i^0 > 0$. Then the solution $q(t,\mathbf{x})$ of the Cauchy problem (3.8) is Gaussian at all time:*

$$\forall\, t \geq 0,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\ q(t,\mathbf{x}) = (2\pi)^{-n/2}\Big(\prod_{i=1}^{n}(s_i(t))^{-1/2}\Big)\ \exp\left(-\sum_{i=1}^{n}\frac{(x_i-\overline{q}_i(t))^2}{2s_i(t)}\right), \tag{3.10}$$

*with:*

$$\begin{aligned}
\overline{q}_1(t) &= \frac{\mu_1\overline{q}_1^0}{\mu_1\cosh(\mu_1 t)+s_1^0\sinh(\mu_1 t)} + \mu_1\int_0^t \Delta(\tau)\,\frac{\mu_1\sinh(\mu_1\tau)+s_1^0\cosh(\mu_1\tau)}{\mu_1\cosh(\mu_1 t)+s_1^0\sinh(\mu_1 t)}\mathrm{d}\tau,\\
\forall 2\leq i\leq n,\ \overline{q}_i(t) &= \frac{\mu_i\overline{q}_i^0}{\mu_i\cosh(\mu_i t)+s_i^0\sinh(\mu_i t)},\\
\text{and } \forall 1\leq i\leq n,\ s_i(t) &= \mu_i\,\frac{\mu_i\sinh(\mu_i t)+s_i^0\cosh(\mu_i t)}{\mu_i\cosh(\mu_i t)+s_i^0\sinh(\mu_i t)}.
\end{aligned} \tag{3.11}$$

*Moreover, we have:*

$$\overline{m}(t) = -\sum_{i=1}^{n}\frac{\overline{q}_i^2(t)+s_i(t)}{2} + \overline{q}_1(t)\Delta(t) - \frac{\Delta(t)^2}{2}.$$

### A degenerate parabolic PDE satisfied by $\mathbf{p}(t,\mathbf{v})$

Our objective here is to derive an equation for $\mathbf{p}(t,\mathbf{v})$ that only involves linear dependencies with respect to the coefficients $v_i$, and holds for general initial phenotype distributions $q_0$.

First, we generalize the formula given in Hamel; Lavigne; Martin; Roques 2020, which expresses the distribution of the fitness components $\mathbf{p}(t,\mathbf{v})$ in terms of the distribution of phenotypes $q(t,\mathbf{x})$.

**Proposition 23.** *For all $t \geq 0$ and $\mathbf{v} = (v_0, \dots, v_n) \in \Gamma_{n+1}$, there holds:*

$$\mathbf{p}(t, \mathbf{v}) = \frac{2^{-(n-1)/2}}{\sqrt{|v_2 \cdots v_n|}} \sum_{\varepsilon \in \{\pm 1\}^{n-1}} q(t, \mathbf{x}^\varepsilon(t, \mathbf{v})), \tag{3.12}$$

*where the variable change $\mathbf{x}^\varepsilon(\mathbf{v})$ is defined by $\mathbf{x}^\varepsilon(\mathbf{v}) = (v_0, x_2^\varepsilon(\mathbf{v}), \dots, x_n^\varepsilon(\mathbf{v}))$, with $x_i^\varepsilon(\mathbf{v}) = \varepsilon_i \sqrt{-2v_i}$ and $\varepsilon = (\varepsilon_2, \dots, \varepsilon_n) \in \{\pm 1\}^{n-1}$. Furthermore, we have:*

$$\forall\, t \geq 0, \ \int_{\Gamma_{n+1}} \mathbf{p}(t, \mathbf{v})\, \mathrm{d}\mathbf{v} = 1 \ \text{ and } \ \overline{m}(t) = -\frac{\Delta(t)^2}{2} + \int_{\Gamma_{n+1}} \boldsymbol{\alpha}(t) \cdot \mathbf{v}\, \mathbf{p}(t, \mathbf{v}) \mathrm{d}\mathbf{v}, \tag{3.13}$$

*with $\boldsymbol{\alpha}(t) = (\Delta(t), 1, \dots, 1) \in \mathbb{R}^{n+1}$, where all above integrals are convergent.*

We remark that, when $q$ satisfies some symmetry properties, the expression (3.12) becomes simpler. Hereafter, for a given function $f \in C(\mathbb{R}^n)$, we define its $\$-$symmetrization $f^\$ \in C(\mathbb{R}^n)$ by:

$$\forall\, \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n, \ f^\$(\mathbf{x}) = 2^{-(n-1)} \sum_{\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{\pm 1\}^{n-1}} f(x_1, \varepsilon_2 x_2, \dots, \varepsilon_n x_n).$$

It is clearly true that if $q$ is the solution of (3.4) with initial condition $q_0$ satisfying the conditions of Theorem 22, then $q^\$$ is the solution of (3.4) with initial condition $q_0^\$$. Moreover, the two distributions $q(t, \cdot)$ and $q^\$(t, \cdot)$ have the same mean fitness $\overline{m}(t)$ at every time $t \geq 0$. By the expression (3.12), this yields that $\mathbf{p}(t, \mathbf{v})$ can be described in terms of the $\$-$symmetrization of $q(t, \cdot)$:

$$\mathbf{p}(t, \mathbf{v}) = \frac{2^{(n-1)/2}}{\sqrt{|v_2 \cdots v_n|}}\, q^\$(t, \mathbf{x}^{\mathbf{1}}(\mathbf{v})), \tag{3.14}$$

for every $t \geq 0$ and $\mathbf{v} \in \Gamma_{n+1}$, with:

$$\mathbf{x}^{\mathbf{1}}(\mathbf{v}) = (v_0, \sqrt{-2v_2}, \dots, \sqrt{-2v_n}) \in \mathbb{R} \times (\mathbb{R}_+^*)^{n-1}.$$

This function $\mathbf{p}(t, \mathbf{v})$ satisfies a nonlocal degenerate parabolic equation, as the following result shows:

**Theorem 24.**
*The distribution function of the fitness components* $\mathbf{p}$ *is a classical* $C^{1,2}(\Gamma_{n+1})$ *solution of:*

$$\partial_t \mathbf{p}(t,\mathbf{v}) = \frac{\mu_1^2}{2}\partial_{v_0v_0}\mathbf{p}(t,\mathbf{v}) - \mu_1^2 v_0 \partial_{v_0v_1}\mathbf{p}(t,\mathbf{v}) + \frac{\mu_1^2 v_0^2}{2}\partial_{v_1v_1}\mathbf{p}(t,\mathbf{v}) - \frac{\mu_1^2}{2}\partial_{v_1}\mathbf{p}(t,\mathbf{v})$$
$$- \sum_{i=2}^{n} \mu_i^2 \left[ v_i \partial_{v_iv_i}\mathbf{p}(t,\mathbf{v}) + \frac{3}{2}\partial_{v_i}\mathbf{p}(t,\mathbf{v}) \right] + (\boldsymbol{\alpha}(t)\cdot\mathbf{v} - \overline{v}(t))\ \mathbf{p}(t,\mathbf{v}), \quad (3.15)$$

*for* $t \geq 0$ *and* $\mathbf{v} \in \Gamma_{n+1}$*, with initial condition:*

$$\mathbf{p}_0(\mathbf{v}) = \frac{2^{(n-1)/2}}{\sqrt{|v_2 \cdots v_n|}}\, q_0^{\$}(\mathbf{x}^{\mathbf{1}}(\mathbf{v})), \quad (3.16)$$

*and:*

$$\overline{v}(t) = \overline{m}(t) + \frac{\Delta(t)^2}{2} = \int_{\Gamma_{n+1}} \boldsymbol{\alpha}(t)\cdot\mathbf{v}\mathbf{p}(t,\mathbf{v})\mathrm{d}\mathbf{v}. \quad (3.17)$$

Note that the set $\Gamma_{n+1}$ is a submanifold. Thus the partial derivatives of $\mathbf{p}$, appearing in (3.15), are defined thanks to the Lie derivative. More precisely, the fourth first terms of the right hand side term are the second derivative with respect to the tangent direction of the function $\mathbf{p}$, where $(v_2,\ldots,v_n)$ are fixed.

### Generating functions

We define the *moment generating functions* (MGFs) $M_{\mathbf{p}}$ of $\mathbf{p}$ and its logarithm – the *cumulant generating function* (CGF) – $C_{\mathbf{p}}$ by:

$$M_{\mathbf{p}}(t,\mathbf{z}) = \int_{\Gamma_{n+1}} e^{\mathbf{z}\cdot\mathbf{v}}\mathbf{p}(t,\mathbf{v})\,\mathrm{d}\mathbf{v}, \quad \text{and} \quad C_{\mathbf{p}}(t,\mathbf{z}) = \log M_{\mathbf{p}}(t,\mathbf{z}), \quad (3.18)$$

for $t \geq 0$, $\mathbf{z} \in \mathbb{R}_+^n$. The following result gives the equation satisfied by $C_{\mathbf{p}}$.

**Theorem 25.**
*The cumulant generating function $C_{\mathbf{p}}$ of $\mathbf{p}$ is well-defined and is of class $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^{n+1})$ [a] and it solves:*

$$\begin{cases} \partial_t C_{\mathbf{p}}(t,\mathbf{z}) &= \boldsymbol{\alpha}(t)\cdot\nabla C_{\mathbf{p}}(t,\mathbf{z}) - \overline{v}(t) + V(\mathbf{z})\cdot\nabla C_{\mathbf{p}}(t,\mathbf{z}) + \gamma(\mathbf{z}),\ t\geq 0,\ \mathbf{z}\in\mathbb{R}_+^n, \\ C_{\mathbf{p}}(0,\mathbf{z}) &= C_{\mathbf{p}_0}(\mathbf{z}),\ \mathbf{z}\in\mathbb{R}_+^n, \end{cases} \tag{3.19}$$

*where:*

$$\boldsymbol{\alpha}(t) = (\Delta(t),1,\dots,1),\ V(\mathbf{z}) = \left(-\mu_1^2 z_0 z_1, -\mu_1^2 z_1^2, \dots, -\mu_n^2 z_n^2\right),$$
$$\gamma(\mathbf{z}) = \frac{\mu_1^2}{2} z_0^2 - \sum_{i=1}^n \frac{\mu_i^2}{2} z_i, \text{ and } \overline{v}(t) = \boldsymbol{\alpha}(t)\cdot\nabla C_{\mathbf{p}}(t,0), \tag{3.20}$$

*and $\nabla C_{\mathbf{p}}(t,\mathbf{z})$ denotes the gradient of $C_{\mathbf{p}}$ with respect to the variable $\mathbf{z}$.*

*a.* This means that the partial derivatives of $C_{\mathbf{p}}$ with respect to the variables $t$ and $\mathbf{z}$ exist and are continuous in $\mathbb{R}_+ \times \mathbb{R}_+^{n+1}$.

The last results of this section provides some explicit expression of $C_{\mathbf{p}}(t,\mathbf{z})$, when $\mathbf{z}$ is close enough to $(0,\dots,0)$.

**Proposition 26.** *The mean fitness is given by, for all $t \geq 0$:*

$$\begin{aligned} \overline{m}(t) = &\left[\frac{\Delta(t)}{\cosh[\mu_1 t]} - \mu_1 \int_0^t \Delta(s)\frac{\sinh[\mu_1 s]}{\cosh^2[\mu_1 t]}\mathrm{d}s\right]\partial_{z_0} C_{\mathbf{p}}(0,\boldsymbol{\psi}(t)) \\ &- \frac{1}{2}\left[\Delta(t) - \mu_1\int_0^t \Delta(s)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s\right]^2 \\ &+ \sum_{i=1}^n \left[\left[1-\tanh^2(\mu_i t)\right]\partial_{z_i} C_{\mathbf{p}}\left(0,\boldsymbol{\psi}(t)\right) - \frac{\mu_i}{2}\tanh(\mu_i t)\right], \end{aligned} \tag{3.21}$$

*with:*

$$\boldsymbol{\psi}(t) = \left(\int_0^t \Delta(s)\frac{\cosh[\mu_1(t-s)]}{\cosh[\mu_1 t]}\mathrm{d}s, \frac{\tanh(\mu_1 t)}{\mu_1}, \dots, \frac{\tanh(\mu_n t)}{\mu_n}\right) \in \mathbb{R}^{n+1}.$$

### 3.2.2. Explicit expressions for $\overline{m}(t)$: some examples

The formula (3.21) for $\overline{m}$ can be simplified in some different cases, which are shown in this section, assuming smooth enough optimum function $\mathbf{O}(t)$. None proof of the following results are developed, as they are direct consequences of

the formula (3.21).

**Optimum shifting with a constant speed.** We assume that the optimum $\mathbf{O}(t)$ moves at a constant speed $c > 0$, as in Alfaro; Berestycki; Raoul 2017; Iglesias; Mirrahimi 2019; Roques; Patout; Bonnefon; Martin 2020.

**Proposition 27.** *Assume that $\Delta(t) = ct$, for some $c \in \mathbb{R}$. Then the mean fitness is given by:*

$$\overline{m}(t) = \frac{c\tanh(\mu_1 t)}{\mu_1 \cosh(\mu_1 t)}\,\partial_{z_0} C_{\mathbf{p}}(0, \boldsymbol{\psi}(t)) - \frac{c^2}{2\mu_1^2}\tanh^2[\mu_1 t] + \sum_{i=1}^{n}\left[[1-\tanh^2(\mu_i t)]\,\partial_{z_i} C_{\mathbf{p}}\,(0, \boldsymbol{\psi}(t)) - \frac{\mu_i}{2}\tanh(\mu_i t)\right], \quad (3.22)$$

*with:*

$$\boldsymbol{\psi}(t) = \left(\frac{c}{\mu_1^2}\left(1 - \frac{1}{\cosh[\mu_1 t]}\right), \frac{\tanh(\mu_1 t)}{\mu_1}, \ldots, \frac{\tanh(\mu_n t)}{\mu_n}\right).$$

Passing to the limit in (3.22) as $t \to +\infty$ leads that:

$$\overline{m}(t) \to \overline{m}(\infty) = -\sum_{i=1}^{n}\frac{\mu_i}{2} - \frac{c^2}{2\mu_1^2}. \quad (3.23)$$

Taking $c = 0$ (steady optimum), we retrieve the formula of the mutational load in the anisotropic case (Hamel; Lavigne; Martin; Roques 2020). When $c \neq 0$, an additional negative term $-c^2/(2\mu_1^2)$ appears, because of the decrease in fitness due to the shifting in the optimum: this term is called "lag load". This formula is consistent with the isotropic case studied in Roques; Patout; Bonnefon; Martin 2020: this decrease in fitness only depends on the mutation rate $\mu_1$.

**Optimum convergence.** We assume that the optimum $\mathbf{O}(t)$ converges to $\mathbf{O}(\infty) = \Delta(\infty)\mathbf{e}_1$, with $\Delta(\infty) \neq 0$. Note that:

$$\Delta(t) - \mu_1 \int_0^t \Delta(s)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s \underset{t\to\infty}{\sim} \frac{\Delta(\infty)}{\cosh[\mu_1 t]},$$

which converges to 0, as $t \to +\infty$. This remark yields the following proposition:

**Proposition 28.** *Assume that $\Delta(t) \to \Delta(\infty) \neq 0$ as $t \to +\infty$.*
*Then $\overline{m}(t) \to -\sum_{i=1}^{n}\frac{\mu_i}{2}$, as $t \to +\infty$.*

When the moving optimum goes to converge as $t \to +\infty$, then the adaptation criterion does not depend on the variation: the lag load is null. We illustrate this proposition with the special case $\Delta(t) = 1 - \exp(-\nu t)$, with some parameter $\nu > 0$. In this special case, the mean fitness is given by:

$$\begin{aligned}\overline{m}(t) = &\left[\frac{1}{\cosh[\mu_1 t]} - e^{-\nu t} + \frac{\mu_1 e^{-\nu t}}{\mu_1^2 - \nu^2}\left(\mu_1 + \nu \tanh[\mu_1 t]\right) - \frac{\mu_1^2}{\mu_1^2 - \nu^2}\frac{1}{\cosh[\mu_1 t]}\right]\frac{\partial_{z_0} C_{\mathbf{p}}(0, \boldsymbol{\psi}(t))}{\cosh[\mu_1 t]} \\ &- \frac{1}{2}\left[\frac{1}{\cosh[\mu_1 t]} - e^{-\nu t} + \frac{\mu_1 e^{-\nu t}}{\mu_1^2 - \nu^2}\left(\mu_1 + \nu \tanh[\mu_1 t]\right) - \frac{\mu_1^2}{\mu_1^2 - \nu^2}\frac{1}{\cosh[\mu_1 t]}\right]^2 \\ &+ \sum_{i=1}^{n}\left[\left[1 - \tanh^2(\mu_i t)\right]\partial_{z_i} C_{\mathbf{p}}\left(0, \boldsymbol{\psi}(t)\right) - \frac{\mu_i}{2}\tanh(\mu_i t)\right]. \quad (3.24)\end{aligned}$$

**Concave case.** We make the assumption now that $\Delta(t)$ is an increasing concave function of $t$. For example, $\Delta(t) = t^\alpha$ for $\alpha \in (0, 1)$. Note that for all $t \geq 0$:

$$\Delta(t) - \mu_1 \int_0^t \Delta(s)\frac{\sin[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s \geq \Delta(t)\left[1 - \mu_1 \int_0^t \frac{\sinh[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s\right] = \frac{\Delta(t)}{\cosh[\mu_1 t]} \geq 0,$$

and:

$$\begin{aligned}\Delta(t) - \mu_1 \int_0^t \Delta(s)\frac{\sin[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s &= \frac{\Delta(t)}{\cosh[\mu_1 t]} + \mu_1 \int_0^t \left(\Delta(t) - \Delta(s)\right)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s, \\ &\leq \frac{\Delta(t)}{\cosh[\mu_1 t]} + \mu_1 \frac{\Delta(t)}{t}\int_0^t (t - s)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s, \\ &\leq \frac{\Delta(t)}{\mu_1 t}\tanh[\mu_1 t].\end{aligned}$$

These inequalities imply that the population adaptation tends to be the same as in the case of a steady optimum, as in the isotropic case (see Roques; Patout; Bonnefon; Martin 2020, for the special case of sublinear functions):

**Proposition 29.** *Assume that $\Delta$ is an increasing concave function. Thus the mean fitness $\overline{m}(t)$ tends to $-\sum_{i=1}^{n}\frac{\mu_i}{2}$, as $t \to +\infty$.*

**Convex case.** Let us turn to the case that $\Delta$ is an increasing convex function. By the same remark as in the convex case, we can check that:

$$\Delta(t) - \mu_1 \int_0^t \Delta(s)\frac{\sin[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s \to +\infty, \text{ as } t \to +\infty,$$

which implies that the adaptation can never occur, when the optimum variation is convex.

**Proposition 30.** *Assume that $\Delta$ is an increasing convex function. Thus the mean fitness $\overline{m}(t)$ diverges to $-\infty$ as $t \to +\infty$.*

We retrieve the case studied by Roques; Patout; Bonnefon; Martin 2020, where the mutation is isotropic and $\Delta(t) = t^\alpha$, for $\alpha > 1$.

**Periodically varying optimum.** The phenotype optimum $\mathbf{O}(t)$ can oscillate periodically, because of external factors which can be periodic (concentration of an antibiotic, temperature, light, *etc.*). Thus if the period of $\Delta$ is $p > 0$, remark that for all $t = pN + \tau$, with $N \in \mathbb{N}$ and $0 \leq \tau < p$:

$$\begin{aligned}\int_0^t \Delta(s)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1 t]}\mathrm{d}s &= \int_0^x \Delta(s)\frac{\sinh[\mu_1(s+pN)]}{\cosh[\mu_1 t]}\mathrm{d}s + \sum_{j=0}^{N-1}\int_0^p \Delta(s)\frac{\sinh[\mu_1(s+pj)]}{\cosh[\mu_1 t]}\mathrm{d}s,\\ &= \int_0^x \Delta(s)\frac{\sinh[\mu_1(s+pN)]}{\cosh[\mu_1 t]}\mathrm{d}s\\ &\quad + \int_0^p \Delta(s)\frac{\sinh[\mu_1 pN/2]\sinh[\mu_1(s+p(N-1)/2)]}{\sinh[\mu_1 p/2]\cosh[\mu_1 t]}\mathrm{d}s,\end{aligned}$$

which converges as $N \to +\infty$. Thus the large time behaviour of the mean fitness can be known:

**Proposition 31.** *Assume that $\Delta$ is $p$-periodic, for some $p \in \mathbb{R}_+^*$. The mean fitness $\overline{m}(t)$ "tends to" the $p$-periodic function:*

$$\overline{m}_\infty(t) = -\frac{1}{2}\left(\Delta(t) - \mu_1\int_0^t \Delta(s)e^{\mu_1(s-t)}\,\mathrm{d}s - \frac{\mu_1\exp[\mu_1 p/2]}{2\sinh[\mu_1 p/2]}\int_0^p \Delta(s)e^{\mu_1(s-t)}\,\mathrm{d}s\right)^2 - \sum_{i=1}^n \frac{\mu_i}{2},$$

*as $t \to +\infty$, in the sense that $\overline{m}(t) - \overline{m}_\infty(t) \to 0$, as $t \to +\infty$.*

This formula is a generalization of the case studied by Roques; Patout; Bonnefon; Martin 2020. First, the mutation are not assumed to be isotropic. Secondly, we make no assumptions on the form of periodicity of $\Delta$. In the particular case

$\Delta(t) = \delta_{\max} \sin(\omega t)$, with $\omega \in \mathbb{R}_+^*$ and $\delta_{\max} > 0$, the mean fitness is given by:

$$\begin{aligned}\overline{m}(t) = \frac{\delta_{\max}\omega}{\omega^2+\mu_1^2} \frac{\omega \sin[\omega t] + \mu_1 \cos[\mu_1 t] \tanh[\mu_1 t]}{\cosh[\mu_1 t]} \partial_{z_0} C_{\mathbf{p}}(0, \boldsymbol{\psi}(t)) \\ - \frac{1}{2}\left(\frac{\delta_{\max}\omega}{\omega^2+\mu_1^2}\right)^2 (\omega \sin[\omega t] + \mu_1 \cos[\mu_1 t] \tanh[\mu_1 t])^2 \\ + \sum_{i=1}^{n} \left[ [1 - \tanh^2(\mu_i t)] \, \partial_{z_i} C_{\mathbf{p}}(0, \boldsymbol{\psi}(t)) - \frac{\mu_i}{2} \tanh(\mu_i t) \right], \quad (3.25)\end{aligned}$$

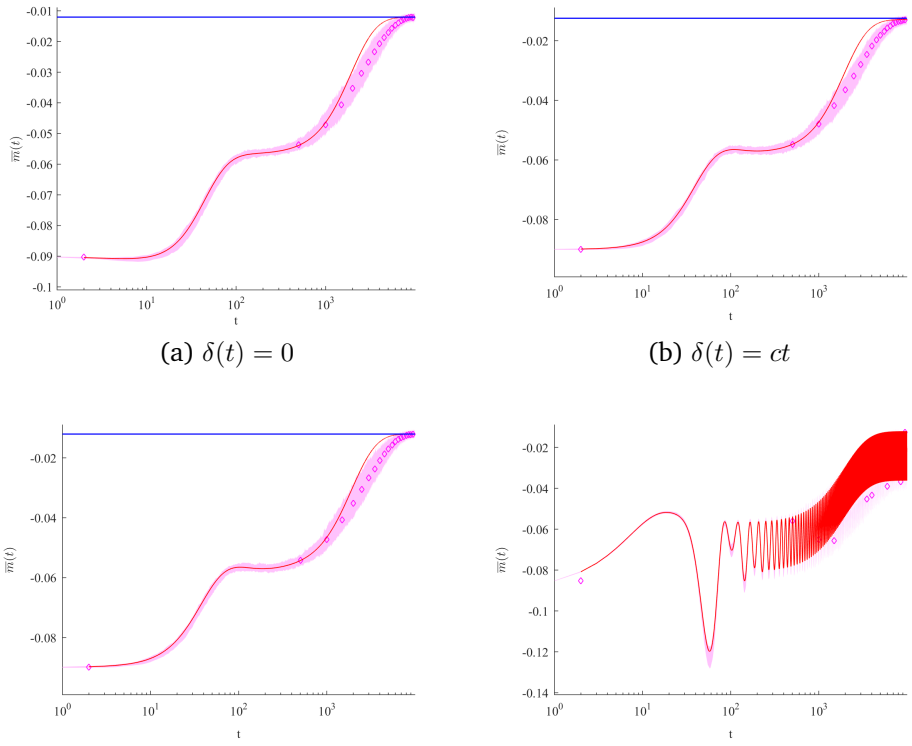
which "tends to" the periodic function:

$$\overline{m}_\infty(t) = -\sum_{i=1}^{n} \frac{\mu_i}{2} - \frac{1}{2}\left(\frac{\delta_{\max}\omega}{\omega^2+\mu_1^2}\right)^2 (\omega \sin[\omega t] + \mu_1 \cos[\mu_1 t])^2 .$$

## 3.3. Numerical computations

This section is devoted to the validation of the diffusion approximation (3.4), thanks to a comparison between our analytic formula and a stochastic individual-based simulations of a standard model of genetic adaptation. This stochastic model is the same model described in Roques; Patout; Bonnefon; Martin 2020, taking into account mutational anisotropy:

**An anisotropic Wright-Fisher individual-based model (IBM) with moving optimum and parameter values.** The population is assumed to have a constant size $N = 10^4$ individuals, which each one is caracterized by a phenotype $\mathbf{x}_i \in \mathbb{R}^n$, by the Fisher's geometrical model (in this section, we assume $n = 2$). Initially, all individuals are assumed to have the same phenotype $\mathbf{x}_{init} = (0.3, 0.3)$. Let us remind that, at time $t$, the fitness of an individual is given by (3.3), *i.e.*, $m_i = -\|\mathbf{x}_i - \mathbf{O}(t)\|^2/2$. Its discrete counterpart is the geometric growth rate, named Darwinian fitness, given by $W_i = \exp(m_i)$. We assume non overlapping generations of duration $\Delta t = 1$. During one generation, because of the joint effect of selection and genetic drift, $N$ individuals are sampled from the previous generation, according to the multinomial sampling, with weight given by the Darwinian fitnesses $W_i$. Then, each individual suffers mutations. Their number is randomly drawn according to a Poisson law with rate $U = 50/300 \simeq 0.167$. Using an anisotropic Gaussian FGM, the phenotype $\mathbf{x}_i$ mutates to an other phenotype $\mathbf{x}_i + \mathrm{d}\mathbf{x}$, with a random phenotypic effect $\mathrm{d}\mathbf{x}$ drawn into a multivariate Gaussian distribution $\mathrm{d}\mathbf{x} \sim \mathcal{N}(0, \Lambda)$, where the variance-covariance matrix $\Lambda \in \mathcal{M}_n(\mathbb{R})$ is the diagonal matrix $\Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & 5\lambda 10^{-4} \end{pmatrix}$ with $\lambda = 1/300$.



(a) $\delta(t) = 0$ (b) $\delta(t) = ct$

(c) $\delta(t) = 1 - \exp(-\nu t)$ (d) $\delta(t) = \delta_{\max} \sin(\omega t)$

Figure 3.1. – **Trajectory of mean fitness: immobile optimum, linearly, converging and periodically varying optimum, with mutational anisoptropy.** Plots associated of simulations of an asexual population evolving according to the IBM. The magenta diamond is the mean value of the mean fitness, averaged over 100 replicates. The pink shading corresponds to the interval between the 0.025 and 0.975 quantiles of the distribution of $\overline{m}(t)$. The red line is the theoretical value of $\overline{m}(t)$ (given by (3.22) and (3.25)). The blue line corresponds to the limit of the mean fitness $\overline{m}(\infty)$ (given by (3.23)). Parameter values: (b) $c = 10^{-3}$; (c) $\nu = 1.2 \cdot 10^{-3}$; (d) $\delta_{\max} = 4 \times \sqrt{1/300}$ and $\omega = \mu_1 \pi$.

**Numerical results.** We first check the accuracy of the model with an immobile optimum in Fig. 3.1(a). This is equivalent to a "linear" moving optimum with constant velocity $c = 0$. The analytical result, given in Hamel; Lavigne; Martin; Roques 2020 and in (3.22), gives a good fit of the mean fitness $\overline{m}(t)$, compared to the average value over 100 replicates of the IBM. As in Hamel; Lavigne; Martin; Roques 2020, we can see some plateaus for some values of $\mu_1$ and $\mu_2$. We choose these parameter $\mu_1 = \sqrt{U\lambda}$ and $\mu_2 = \sqrt{U\lambda/100}$ to see this phenomenon. When the optimum moves linearly according to the time $t$, plateaus can also appear,

as we can see in Fig. 3.1(b). In this figure, Eq. (3.22), for $c = 10^{-3} > 0$, yields a good approximation of the mean value $\overline{m}(t)$, averaged over replicates, and so of the lag load.

When the optimum converges to a given point $\mathbf{x}_\infty$ (see Figure 3.1(c)), the mean fitness $\overline{m}(t)$ has the same kind of evolution than in the static and the mobile optimum, with some plateaus.

Figure 3.1(d) depicts the trajectories of mean fitness with a periodically varying optimum. The analytic description given by (3.25) captures the oscillations and the short temporary plateau of the average dynamics of the IBM.

## 3.4. Discussion

In this paper, the study of the generating functions, based on several articles (Gil; Hamel; Martin; Roques 2017; Hamel; Lavigne; Martin; Roques 2020; Martin; Roques 2016; Roques; Patout; Bonnefon; Martin 2020), yields to an explicit formula of the mean fitness associated to (3.4), without assumption on the initial phenotype distribution. The only assumption that we need on the movement on the optimum is that it moves on one direction which is given by one of the direction axis. The results presented here are a direct generalization of the static optimum case with anisotropy (Hamel; Lavigne; Martin; Roques 2020), and the moving optimum case with isotropy (Roques; Patout; Bonnefon; Martin 2020). Anisotropy and optimum move have additive effect on the formula of the mean fitness, as we can see in Fig. 3.1: the four different examples show different plateaus for a good choice of parameters $\mu_1$ and $\mu_2$, as in the static optimum, but the optimum move changes the variation similarly as in the isotropic case (decreasing of the limit value of the mean fitness, see Roques; Patout; Bonnefon; Martin 2020).

This recent framework, based on generating function of the fitness distribution, leads us to a total description of the evolution of an asexual population across time, while traditional approaches – "traveling waves", Hamilton-Jacobi, "moment closure" (as "Gaussian solution") or spectral analysis – implies results on the asymptotic behaviour. However, Hamel; Lavigne; Martin; Roques 2020 has already shown that the dynamics of the mean fitness can be more complex with the anisotropic assumption, than with the isotropic one (appearing of plateaus, when the mutation parameters $\mu_i$ have different scales).

The issue of this paper is to be a first step to the description of the transient dynamics of the mean fitness, with anisotropy and without assumption on the optimum movement $\mathbf{O}(t)$. As said in the Introduction, dropping this hypothesis would yield to study the distribution of a fitness vector $\mathbf{v}' \in \mathbb{R}^n \times \mathbb{R}^n_-$.

## 3.5. Proofs

This section is devoted to the proofs of the results announced in Section 3.2. Section 3.5.1 is concerned with the PDE satisfied by the distribution of the fitness vector $\mathbf{v}$. This study yields a description of the generating functions of this distribution, developed in Section 3.5.2.

### 3.5.1. A degenerate parabolic PDE satisfied by $\mathbf{p}(t, \mathbf{m})$

*Proof of Theorem 24.* Formula (3.14) yields that the function $\mathbf{p}$ is of class $C^{1,2}(\Gamma_{n+1})$ with initial condition $\mathbf{p}_0$ given by (3.16). Furthermore, it is easily to check that, for all $2 \leq i \leq n$ for all $t \geq 0$ and $\mathbf{v} \in \Gamma_{n+1}$:

$$2^{-(n-1)/2}\, \partial_{v_i}\mathbf{p}(t,\mathbf{v}) = -\frac{q^{\$}(t,\mathbf{x^1}(\mathbf{v}))}{2v_i\sqrt{|v_2\cdots v_n|}} - \frac{1}{\sqrt{2|v_i|}}\,\frac{\partial_{x_i}q^{\$}(t,\mathbf{x^1}(\mathbf{v}))}{\sqrt{|v_2\cdots v_n|}},$$

$$2^{-(n-1)/2}\, \partial_{v_iv_i}\mathbf{p}(t,\mathbf{v}) = \frac{3}{4}\,\frac{q^{\$}(t,\mathbf{x^1}(\mathbf{v}))}{v_i^2\sqrt{|v_2\cdots v_n|}} + \frac{3}{2}\,\frac{\partial_{x_i}q^{\$}(t,\mathbf{x^1}(\mathbf{v}))}{v_i\sqrt{2|v_i|}\sqrt{|v_2\cdots v_n|}} - \frac{1}{2v_i}\,\frac{\partial_{x_ix_i}q^{\$}(t,\mathbf{x^1}(\mathbf{v}))}{\sqrt{|v_2\cdots v_n|}},$$

with $\mathbf{x^1}(\mathbf{v}) = (v_0, \sqrt{-2v_2}, \ldots, \sqrt{-2v_n}) \in \mathbb{R}\times(\mathbb{R}_+^*)^{n-1}$. Moreover, fixing the coordinates $(v_2,\ldots,v_n)\in\mathbb{R}^n$, the Lie derivative of $\mathbf{p}$ with respect to $(v_0,v_1)$ at $t>0$ is given by:

$$\mathcal{L}\mathbf{p}(t,\mathbf{v}) = (1,-v_0)\cdot(\partial_{v_0}\mathbf{p}(t,\mathbf{v}),\partial_{v_1}\mathbf{p}(t,\mathbf{v})) = \partial_{v_0}\mathbf{p}(t,\mathbf{v}) - v_0\,\partial_{v_1}\mathbf{p}(t,\mathbf{v}),$$

and the second Lie derivative by:

$$\begin{aligned}\mathcal{L}^2\mathbf{p}(t,\mathbf{v}) &= (1,-v_0)\cdot(\partial_{v_0}[\mathcal{L}\mathbf{p}](t,\mathbf{v}),\partial_{v_1}[\mathcal{L}\mathbf{p}](t,\mathbf{v})),\\ &= \partial_{v_0v_0}\mathbf{p}(t,\mathbf{v}) - \partial_{v_1}\mathbf{p}(t,\mathbf{v}) - 2v_0\partial_{v_0v_1}\mathbf{p}(t,\mathbf{v}) + v_0^2\partial_{v_1v_1}\mathbf{p}(t,\mathbf{v}).\end{aligned}$$

This last relation, with Proposition 23, yields that for all $t>0$ and $\mathbf{v}\in\Gamma_{n+1}$:

$$\begin{aligned}\partial_{v_0v_0}\mathbf{p}(t,\mathbf{v}) - \partial_{v_1}\mathbf{p}(t,\mathbf{v}) - 2v_0\partial_{v_0v_1}\mathbf{p}(t,\mathbf{v}) + v_0^2\partial_{v_1v_1}\mathbf{p}(t,\mathbf{v}) &= \mathcal{L}^2\mathbf{p}(t,\mathbf{v}),\\ &= \frac{2^{-(n-1)/2}}{\sqrt{|v_2\cdots v_n|}}\,\partial_{x_1x_1}q^{\$}(t,\mathbf{x^1}(\mathbf{v})).\end{aligned}$$

Hence, we have for all $t > 0$ and $\mathbf{v} \in \Gamma_{n+1}$:

$$\begin{aligned}
&\sum_{i=2}^{n} \mu_i^2 \left( v_i \partial_{v_i v_i} \mathbf{p}(t,\mathbf{v}) + \frac{3}{2} \partial_{v_i} \mathbf{p}(t,\mathbf{v}) \right) \\
&\qquad = -\sum_{i=2}^{n} \frac{2^{(n-1)/2}}{\sqrt{|v_2 \cdots v_n|}} \frac{\mu_i^2}{2} \partial_{x_i x_i} q^{\$}(t, \mathbf{x}^{\mathbf{1}}(\mathbf{v})), \\
&\qquad = -\frac{2^{(n-1)/2}}{\sqrt{|v_2 \cdots v_n|}} \left[ \partial_t q^{\$}(t, \mathbf{x}^{\mathbf{1}}(\mathbf{v})) - \frac{\mu_1^2}{2} \partial_{x_1 x_1} q^{\$}(t, \mathbf{x}^{\mathbf{1}}(\mathbf{v})) \right. \\
&\qquad\qquad \left. - (m(t, \mathbf{x}^{\mathbf{1}}(\mathbf{v})) - \overline{m}(t))\, q^{\$}(t, \mathbf{x}^{\mathbf{1}}(\mathbf{v})) \right], \\
&\qquad = -\partial_t \mathbf{p}(t,\mathbf{v}) + \frac{\mu_1^2}{2} \left[ \partial_{v_0 v_0} \mathbf{p}(t,\mathbf{v}) - 2 v_0 \partial_{v_0 v_1} \mathbf{p}(t,\mathbf{v}) + v_0^2 \partial_{v_1 v_1} \mathbf{p}(t,\mathbf{v}) - \partial_{v_1} \mathbf{p}(t,\mathbf{v}) \right] \\
&\qquad\qquad + \left( -\frac{\Delta(t)^2}{2} + \Delta(t) v_0 + \sum_{i=1}^{n} v_i - \overline{m}(t) \right) \mathbf{p}(t,\mathbf{v}).
\end{aligned}$$

Theorem 24 is thereby proven. □

### 3.5.2. Generating functions

*Proof of Theorem 25.* The proof of the well-possedness and the regularity of $M_{\mathbf{p}}$ are similar as in Hamel; Lavigne; Martin; Roques 2020.

The proof of Theorem 2.6 in Hamel; Lavigne; Martin; Roques 2020 can be generalized in our case, to justify the following formal computations.

Using the fact that $M_{\mathbf{p}}(t,\mathbf{z}) = 2^{n-1} \int_{\mathbb{R}_+^n} \exp\left[ z_0 x_1 - \sum_{i=1}^{n} \frac{z_i x_i^2}{2} \right] q^{\$}(t,\mathbf{x}) \mathrm{d}\mathbf{x}$, we have:

$$\begin{aligned}
&2^{n-1} \int_{\mathbb{R}_+^n} \exp\left[ z_0 x_1 - \sum_{i=1}^{n} \frac{z_i x_i^2}{2} \right] \partial_{x_1 x_1} q^{\$}(t,\mathbf{x}) \mathrm{d}\mathbf{x} \\
&\qquad = 2^{n-1} \int_{\mathbb{R}_+^n} \left[ z_0^2 - 2 z_0 z_1 x_1 + z_1^2 x_1^2 - z_1 \right] \exp\left[ z_0 x_1 - \sum_{i=1}^{n} \frac{z_i x_i^2}{2} \right] q^{\$}(t,\mathbf{x}) \mathrm{d}\mathbf{x}, \\
&\qquad = (z_0^2 - z_1) M_{\mathbf{p}}(t,\mathbf{z}) - 2 z_0 z_1 \partial_{z_0} M_{\mathbf{p}}(t,\mathbf{z}) - 2 z_1^2 \partial_{z_1} M_{\mathbf{p}}(t,\mathbf{z}).
\end{aligned}$$

By the computations developed in Hamel; Lavigne; Martin; Roques 2020, one concludes that:

$$\begin{aligned}
\partial_t M_{\mathbf{p}}(t,\mathbf{z}) = -\mu_1^2 z_0 z_1 \partial_{z_0} M_{\mathbf{p}}(t,\mathbf{z}) - \sum_{i=1}^{n} \mu_i^2 z_i^2 \partial_{z_i} M_{\mathbf{p}}(t,\mathbf{z}) + \left[ \frac{\mu_1^2}{2} z_0^2 - \sum_{i=1}^{n} \frac{\mu_i^2}{2} z_i \right] M_{\mathbf{p}}(t,\mathbf{z}) \\
+ \boldsymbol{\alpha}(t) \cdot \nabla M_{\mathbf{p}}(t,\mathbf{z}) - \overline{v}(t) M_{\mathbf{p}}(t,\mathbf{z}), \quad (3.26)
\end{aligned}$$

for every $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}_+^{n+1}$, where $\boldsymbol{\alpha}$ is as in (3.20). Since the right-hand side of the above equation is continuous in $\mathbb{R}_+ \times \mathbb{R}_+^{n+1}$, one infers that the function $\partial_t M_{\mathbf{p}}$ is extendable by continuity in $\mathbb{R}_+ \times \mathbb{R}_+^{n+1}$ and (3.26) holds in $\mathbb{R}_+ \times \mathbb{R}_+^{n+1}$. Owing to the definition $C_{\mathbf{p}} = \log M_{\mathbf{p}}$, one concludes that $\partial_t C_{\mathbf{p}}$ is continuous in

$\mathbb{R}_+ \times \mathbb{R}_+^{n+1}$ (finally, $C_{\mathbf{p}}$ is of class $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^{n+1})$) and:

$$\partial_t C_{\mathbf{p}}(t,\mathbf{z}) = -\mu_1^2 z_0 z_1 \partial_{z_0} C_{\mathbf{p}}(t,\mathbf{z}) - \sum_{i=1}^{n} \mu_i^2 z_i^2 \partial_{z_i} C_{\mathbf{p}}(t,\mathbf{z}) + \left[\frac{\mu_1^2}{2} z_0^2 - \sum_{i=1}^{n} \frac{\mu_i^2}{2} z_i\right] + \boldsymbol{\alpha}(t) \cdot \nabla C_{\mathbf{p}}(t,\mathbf{z}) - \overline{v}(t),$$

for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}_+^n$. Therefore, (3.19) holds in $\mathbb{R}_+ \times \mathbb{R}_+^{n+1}$ and the proof of Theorem 25 is thereby complete. □

Before going into the proof of the remaining results, let us first enounce the following lemma, which is a generalization of Lemma 4.3 in Hamel; Lavigne; Martin; Roques 2020:

**Lemma 3.** *The Cauchy problem:*

$$\begin{cases} \partial_t Q(t,\mathbf{z}) &= \mathbf{1} \cdot (\nabla Q(t,\mathbf{z}) - \nabla Q(t,0)) - \widetilde{b}(t,\mathbf{z}),\ t \geq 0,\ \mathbf{z} \in \mathbb{R}_+^{n+1}, \\ Q(0,\mathbf{z}) &= Q_0(\mathbf{z}),\ \mathbf{z} \in \mathbb{R}_+^{n+1}, \\ Q(t,0) &= 0,\ t \geq 0, \end{cases} \tag{3.27}$$

*with $\widetilde{b} \in \mathbb{R}_+^{n+1}$ and $Q_0 \in C^1(\mathbb{R}_+ \times \mathbb{R}_+^{n+1})$ such that $\widetilde{b}(t,0) = Q_0(0) = 0$, admits a unique $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^{n+1})$ solution, which is given by the expression:*

$$Q(t,\mathbf{z}) = \int_0^t [\widetilde{b}(t-s, s\mathbf{1}) - \widetilde{b}(t-s, \mathbf{z}+s\mathbf{1})]\, ds + Q_0(\mathbf{z}+t\mathbf{1}) - Q_0(t\mathbf{1}), \tag{3.28}$$

*where $\mathbf{1} = (1,\ldots,1) \in \mathbb{R}^{n+1}$.*

*Proof of Proposition 26.* In order to derive a general formula for the $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^{n+1})$ solution $C_{\mathbf{p}}$ of (3.19), a substitution of the spatial variable let us use the previous special case described in Lemma 3. To do so, we set, for $t \geq 0$ and $z \in \mathbb{R}_+^{n+1}$,

$$Q(t,\mathbf{z}) = C_{\mathbf{p}}(t, \mathbf{y}(t,\mathbf{z})),$$

where $\mathbf{y}(t,\mathbf{z}) = (y_0(t,\mathbf{z}),\ldots,y_n(t,\mathbf{z}))$ and:

$$\begin{cases} y_0(t,\mathbf{z}) = -\frac{\cosh[\mu_1(z_0+t)]}{\cosh[\mu_1 z_0]} \left[z_1 - z_0 - \int_t^{z_0+t} \Delta(s) \frac{\cosh[\mu_1(z_0+t-s)]}{\cosh[\mu_1(z_0+t)]}\, \mathrm{d}s\right], \\ y_1(t,\mathbf{z}) = \tanh(\mu_1 z_0)/\mu_1, \\ y_i(t,\mathbf{z}) = \tanh(\mu_i z_i)/\mu_i \text{ for every } 1 \leq i \leq n. \end{cases}$$

Notice that $\mathbf{y}(\mathbf{z}) \in \mathbb{R}_+^{n+1}$ for every $\mathbf{z} \in \mathbb{R}_+^{n+1}$. The function $Q$ is of class $C^{1,1}(\mathbb{R}_+ \times$

$\mathbb{R}_+^{n+1}$) and:

$$\begin{aligned}
\mathbf{1}\cdot\nabla Q(t,\mathbf{z}) &= \sum_{i=0}^{n}\partial_{z_i}Q(t,\mathbf{z}),\\
&= -\mu_1\frac{\sinh[\mu_1 t]}{\cosh^2[\mu_1 z_0]}\left[z_1 - z_0 - \int_t^{z_0+t}\Delta(s)\frac{\cosh[\mu_1(z_0+t-s)]}{\cosh[\mu_1(z_0+t)]}\,\mathrm{d}s\right]\partial_{z_0}C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))\\
&\quad+\left[\Delta(z_0+t)-\mu_1\int_t^{z_0+t}\Delta(s)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1(z_0+t)]}\mathrm{d}s\right]\frac{\partial_{z_0}C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))}{\cosh[\mu_1 z_0]}\\
&\quad+(1-\tanh^2(\mu_1 z_0))\partial_{z_1}C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))+\sum_{i=2}^{n}\left(1-\tanh^2(\mu_i z_i)\right)\partial_{z_i}C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z})),
\end{aligned}$$

and:

$$\begin{aligned}
\partial_t Q(t,\mathbf{z}) &= \partial_t C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))+\partial_t y_0(t,\mathbf{z})\partial_{z_0}C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z})),\\
&= \partial_t C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))\\
&\quad-\mu_1\frac{\sinh[\mu_1(z_0+t)]}{\cosh[\mu_1 z_0]}\left[z_1 - z_0 - \int_t^{z_0+t}\Delta(s)\frac{\cosh[\mu_1(z_0+t-s)]}{\cosh[\mu_1(z_0+t)]}\,\mathrm{d}s\right]\partial_{z_0}C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))\\
&\quad+\left[\Delta(z_0+t)-\Delta(t)\cosh[\mu_1 z_0]-\mu_1\int_t^{z_0+t}\Delta(s)\frac{\sinh[\mu_1 s]}{\cosh[\mu_1(z_0+t)]}\,\mathrm{d}s\right]\frac{\partial_{z_0}C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))}{\cosh[\mu_1 z_0]}.
\end{aligned}$$

This two formulae yield that:

$$\partial_t Q(t,\mathbf{z})-\mathbf{1}\cdot\nabla Q(t,\mathbf{z}) = \partial_t C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z}))-[\boldsymbol{\alpha}(t)+V(\mathbf{y}(t,\mathbf{z}))]\cdot\nabla C_{\mathbf{p}}(t,\mathbf{y}(t,\mathbf{z})),$$

for all $(t,\mathbf{z})\in\mathbb{R}_+\times\mathbb{R}_+^{n+1}$, where $\boldsymbol{\alpha}$ and $V$ are given in (3.20). As $\overline{v}(t)=\boldsymbol{\alpha}(t)\cdot\nabla C_{\mathbf{p}}(t,0)=\mathbf{1}\cdot\nabla Q(t,0)$ and $Q(t,0)=C_{\mathbf{p}}(t,0)=\log M_{\mathbf{p}}(t,0)=0$ by (3.13), it follows from (3.19) that:

$$\begin{cases}\partial_t Q(t,\mathbf{z}) &= \mathbf{1}\cdot(\nabla Q(t,\mathbf{z})-\nabla Q(t,0))+\gamma(\mathbf{y}(t,\mathbf{z})),\ t\geq 0,\ \mathbf{z}\in\mathbb{R}_+^{n+1},\\ Q(t,0) &= 0,\ t\geq 0,\end{cases}$$

and $Q(0,\mathbf{z})=C_{\mathbf{p}}(0,\mathbf{y}(0,\mathbf{z}))$ for all $\mathbf{z}\in\mathbb{R}_+^{n+1}$. The functions $C_{\mathbf{p}}(0,\mathbf{y}(0,\cdot))$ and $\widetilde{b}:=b\circ\mathbf{y}$ are of class $C^1(\mathbb{R}_+\times\mathbb{R}_+^{n+1})$ and $C_{\mathbf{p}}(0,\mathbf{y}(0,0))=\widetilde{b}(t,0)=0$. Therefore, Lemma 3 implies that:

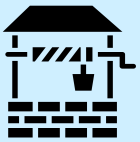
$$\begin{aligned}
Q(t,\mathbf{z}) = \int_0^t[\gamma(\mathbf{y}(t-s,\mathbf{z}+s\mathbf{1}))-\gamma(\mathbf{y}(t-s,s\mathbf{1}))]\,\mathrm{d}s\\
+C_{\mathbf{p}}(0,\mathbf{y}(0,\mathbf{z}+t\mathbf{1}))-C_{\mathbf{p}}(0,\mathbf{y}(0,t\mathbf{1})),
\end{aligned}\tag{3.29}$$

for all $(t,\mathbf{z})\in\mathbb{R}_+\times\mathbb{R}_+^{n+1}$. Using the relations $\overline{m}(t)=\overline{v}(t)-\frac{\Delta(t)^2}{2}$ and $\overline{v}(t)=\boldsymbol{\alpha}(t)\cdot\nabla C_{\mathbf{p}}(t,0)=\mathbf{1}\cdot\nabla Q(t,0)$, we can check (3.21). The proof of Proposition 26 is thereby complete. □

# II

# Adaptation with continuous migration

4



# When sinks become sources: adaptive colonization in asexuals

**F. Lavigne[a,b,c], G. Martin[c,*], Y. Anciaux[c,d], J. Papaïx[a] and L. Roques[a,*]**

[a] BioSP, INRA, 84914, Avignon, France
[b] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
[c] ISEM (UMR 5554), CNRS, 34095, Montpellier, France
[d] BIRC, Aarhus University, C.F. Møllers Allé 8, DK-8000 Aarhus C, Denmark
*Corresponding authors: guillaume.martin@umontpellier.fr and lionel.roques@inra.fr

**Abstract**

The successful establishment of a population into a new empty habitat outside of its initial niche is a phenomenon akin to evolutionary rescue in the presence of immigration. It underlies a wide range of processes, such as biological invasions by alien organisms, host shifts in pathogens or the emergence of resistance to pesticides or antibiotics from untreated areas.

In this study, we derive an analytically tractable framework to describe the coupled evolutionary and demographic dynamics of asexual populations in a source-sink system. In particular, we analyze the influence of several factors — immigration rate, mutational parameters, and harshness of the stress induced by the change of environment — on the establishment success in the sink (i.e. the formation of a self-sufficient population in the sink), and on the time until establishment. To this aim, we use a classic phenotype-fitness landscape (Fisher's geometrical model in $n$ dimensions) where source and sink habitats determine distinct phenotypic optima. The harshness of stress, in the sink, is determined by the distance between the fitness optimum in the sink and that of the source. The dynamics of the full distribution of fitness and of population size in the sink are analytically predicted under a strong mutation strong immigration limit where the population is always polymorphic.

The resulting eco-evolutionary dynamics depends on mutation and immigration rates in a non straightforward way. Below some mutation rate threshold, establishment always occurs in the sink, following a typical four-phases trajectory of the mean fitness. The waiting time to this establishment is independent of the immigration rate and decreases with the mutation rate. Beyond the mutation rate threshold, lethal mutagenesis impedes establishment and the sink population remains so, albeit with an equilibrium state that depends on the details of the fitness landscape. We use these results to get some insight into possible effects of several management strategies.

# Sommaire

## 4.1. Introduction

Most natural populations are spread over a heterogeneous set of environments, to which local subpopulations may be more or less adapted. When these local populations exchange migrants we can define “source" and “sink" populations. Source populations, where the local genotypes have positive growth rate, are self-sustained and can send migrants to the rest of the system. They may be connected to sink populations, where local genotypes are so maladapted that they have negative growth rates (Pulliam 1988). A recent review (Furrer; Pasinelli 2016) showed that empirical examples of sources and sinks exist throughout the whole animal kingdom. In the absence of any plastic or evolutionary change, source-sink systems are stable, with the sources being close to their carrying capacity and the sinks being only maintained by incoming maladapted migrants from source environments. In the literature, different source-sink systems have been categorized by their pattern of immigration and emigration (for more detail on these different categories see Fig. 1 in Sokurenko; Gomulkiewicz; Dykhuizen 2006 and Table 1 in Loreau; Daufresne; Gravel; Guichard, et al. 2013). One particular system, defined as “black-hole sink" (Gomulkiewicz; Holt; Barfield 1999), corresponds to a demographic dead-end, from which emigration is negligible. These black-hole sinks, and their demographic and evolutionary dynamics, are the canonical model for studying the invasion of a new environment, outside of the initial species “niche", and thus initially almost empty (Holt; Gomulkiewicz; Barfield 2003; Holt; Barfield; Gomulkiewicz 2004). In this article, we will only consider black-hole sinks: for compactness, we hereafter simply use the term ’sink’, when in fact referring to a black-hole sink population. The demographic and evolutionary process leading, or not, to the invasion of a sink is akin to evolutionary rescue in the presence of immigration. It underlies a wide range of biological processes: invasion of new habitats by alien organisms (Colautti; Alexander; Dlugosch; Keller, et al. 2017), host shifts in pathogens or the emergence of resistance to pesticides or antibiotics, within treated areas or patients (discussed e.g. in Jansen; Coors; Stoks; De Meester 2011 and Sokurenko; Gomulkiewicz; Dykhuizen 2006). The issues under study in these situations are the likelihood and timescale of successful invasions (or establishment) of sinks from neighboring source populations. “Establishment" in a sink is generally considered successful when the population is self-sustaining in this new environment, even if immigration was to stop (e.g., Blackburn; Pyšek; Bacher; Carlton, et al. 2011, for a definition of this concept in the framework of biological invasions).

A rich theoretical literature has considered the effects of demography and/or evolution in populations facing a heterogeneous environment connected by migration, both in sexuals (e.g., Kirkpatrick; Barton 1997) and asexuals (e.g., Débarre; Ronce; Gandon 2013). The source-sink model is a sub-case of this general problem, that has received particular attention (for a review, see Holt; Barfield; Gomulkiewicz 2005): below, we quickly summarize the relevance and key proper-

ties of source-sink models. The asymmetric migration (from source to sink alone), characteristic of black-hole sinks, provides a key simplification, while remaining fairly realistic over the early phase of invasion, where success or failure is decided. For the same reason, some models further ignore density-dependent effects in the sink, although both high (logistic growth) and/or low (Allee effect) densities could further impact the results, when relevant (discussed in Holt 2009).

Some source-sink models (e.g., Drury; Drake; Lodge; Dwyer 2007; Garnier; Roques; Hamel 2012), focus on detailed demographic dynamics, in the absence of any evolutionary forces. Evolutionary forces (selection, mutation, migration, drift and possibly recombination/segregation) can greatly alter the outcome. These forces may yield both local adaptation or maladaptation, favoring or hindering (respectively) the ultimate invasion of the sink ("adaptive colonization", Gomulkiewicz; Holt; Barfield; Nuismer 2010), however harsh. In this context, mutation and migration are double edged swords, both increasing the local variance available for selection but generating mutation and migration loads, due to the adverse effects of deleterious mutations and maladapted migrant inflow (resp.). For a review of the ambivalent effects of mutation and migration see e.g., (Lenormand 2002) and (Débarre; Ronce; Gandon 2013). Disentangling the complex interplay of these forces with demographic dynamics is challenging, and modelling approaches have used various ecological simplifications: e.g. no age or stage structure, constant stress, constant migration rate.

The associated evolutionary processes are also simplified. As for evolutionary rescue models (discussed in Alexander; Martin; Martin; Bonhoeffer 2014), evolutionary source-sink models may be divided into two classes, based on the presence or absence of context-dependence in the genotype-fitness map they assume (Gomulkiewicz; Holt; Barfield; Nuismer 2010). In context-independent models, fitness in the sink is additively determined by a single or a set of freely recombining loci, and adaptation occurs by directional selection on fitness itself (Gomulkiewicz; Holt; Barfield; Nuismer 2010; Barton; Etheridge 2017). In context-dependent models, which arguably forms the vast majority of source-sink models, fitness is assumed to be a concave function (typically quadratic or Gaussian) of an underlying phenotype, with the source and sink environments corresponding to alternative optima for this phenotype (e.g., Holt; Gomulkiewicz; Barfield 2003; Holt; Barfield; Gomulkiewicz 2004). Such nonlinear phenotype-fitness maps, with environment dependent optima, generate context-dependent interactions for fitness (epistasis and genotype x environment or "G x E" interactions): the effect of a given allele depends on the genetic and environmental background in which it is found. These models reproduce observed empirical patterns of mutation fitness effects across backgrounds (Martin; Elena; Lenormand 2007; MacLean; Hall; Perron; Buckling 2010; Trindade; Sousa; Gordo 2012), reviewed in (Tenaillon 2014). However, their analysis is more involved. Most analytical treatments have thus relied on stationarity assumptions: e.g. describing the ultimate (mutation-selection-migration) equilibrium in asexuals (Débarre;

Ronce; Gandon 2013), or assuming a constant genetic variance and Gaussian distribution for the underlying trait in sexuals (e.g., Gomulkiewicz; Holt; Barfield 1999; Holt; Barfield; Gomulkiewicz 2004). While numerical explorations (by individual-based simulations) often relax these stationarity assumptions, they are necessarily bound to study a limited set of parameter value combinations.

In this paper, we explore a complementary scenario: a source-sink system, out of equilibrium, in an asexual population. The focus on asexuals is intended to better capture pathogenic microorganisms or microbial evolution experiments. We ignore density-dependence by assuming that it is negligible (no Allee effects) before and during the critical early phase of the sink invasion (far below the population reaches the carrying capacity). Considering asexuals and density-independent populations implies that several complex effects of migration (both genetic and demographic) can be ignored. Because migrants do not hybridize/recombine with locally adapted genotypes or use up limiting resources, the maladaptive effects of migration are limited. Migration meltdown and gene swamping (see Lenormand 2002) are thus expected to be absent. This simplification allows to analytically track out-of-equilibrium dynamics, in a context-dependent model (with epistasis and G x E), without requiring stable variance or Gaussian and moment-closure approximations for the phenotypic distribution.

More precisely, we study the transient dynamics of a sink under constant immigration from a source population at mutation-selection balance and a sink initially empty (invasion process). We use the classic quadratic phenotype-fitness map with an isotropic version of Fisher's geometrical model (FGM) with mutation pleiotropically affecting $n$ phenotypic traits. To make analytical progress, we use a deterministic approximation (as in Martin; Roques 2016) that neglects stochastic aspects of migration, mutation and genetic drift, but tracks the full distribution of fitness and phenotypes. Under a weak selection strong mutation (WSSM) regime, when mutation rates are large compared to mutation effects, we further obtain an analytically tractable coupled partial-ordinary differential equation (PDE-ODE) model describing the evolutionary and demographic dynamics in the sink. This framework allows us to derive analytic formulae for the demographic dynamics and the distribution of fitness, at all times, which we test by exact stochastic simulations. We investigate the effect of demographic and evolutionary parameters on the establishment success, on the establishment time, and on the equilibrium mean fitness in the sink. In particular, we focus on the effects of the immigration rate, the harshness of stress (distance between source and sink optima), and mutational parameters (rate, phenotypic effects and dimension $n$).

## 4.2. Methods

Throughout this paper, we follow the dynamics of the fitness distribution of the individuals in the sink environment, under the joint action of mutation, selection

and immigration from the source. The latter remains stable at mutation-selection balance, as migration is asymmetric in this black-hole sink. We consider an asexual population evolving in continuous time. Consistently, we focus on Malthusian fitness $m$ (hereafter 'fitness'): the expected growth rate (over stochastic demographic events) of a given genotypic class, per arbitrary time units. Absolute Malthusian fitnesses $r$ are therefore (expected) growth rates, and without loss of generality, $m$ is measured relative to that of the phenotype optimal in the sink, with growth rate $r_{\max}$. We thus have $m = r - r_{\max}$, and the mean absolute fitness $\overline{r}(t)$ and mean relative fitness $\overline{m}(t)$, at time $t$, satisfy:

$$\overline{r}(t) = r_{\max} + \overline{m}(t).$$

We use a *deterministic approximation* which neglects variations among replicate populations. Under this approximation, $\overline{r}(t)$ (respectively $\overline{m}(t)$), the mean absolute (resp. relative) fitness within each population can be equated to their expected values (across stochastic events). In general, the bar $\bar{}$ denotes averages taken over the sink population. The main notations are summarized in Table 4.1.

### 4.2.1. Demographic model and establishment time $t_0$

In our simple scenario without density-dependence, evolutionary and demographic dynamics are entirely coupled by the mean absolute Malthusian fitness (mean growth rate). We consider a sink population with mean growth rate $\overline{r}(t)$ at time $t$, receiving on average $d$ individuals per unit time by immigration. Under the deterministic approximation, the population size dynamics in the sink environment are therefore given by:

$$N'(t) = \overline{r}(t)\, N(t) + d, \tag{4.1}$$

with $N'(t)$ the derivative of $N$ with respect to $t$ at time $t$.

In the absence of adaptation, $\overline{r}$ is constant, leading to an equilibrium population size $N = d/(-\overline{r})$ when $\overline{r} < 0$, as mentioned in the Introduction. When genetic adaptation is taken into account, we need further assumptions to describe the dynamics of $\overline{r}(t)$ in the sink.

We always assume that the new environment is initially empty ($N(0) = 0$) and that the individuals from the source are, on average, maladapted in the sink ($\overline{r}(0) < 0$). Following a classic definition (Blackburn; Pyšek; Bacher; Carlton, et al. 2011), we define the establishment time $t_0$ as the first time when the growth rate of the sink becomes positive in the absence of immigration:

$$t_0 := \inf\{t > 0 \text{ s.t. } \overline{r}(t) > 0\}.$$

This means that, from time $t_0$, the sink population is self-sustaining in the absence of immigration and further adaptation. By definition (assuming that $\overline{r}$ is continu-

| Notation | Description |
|---|---|
| $n$ | number of pleiotropic phenotypes |
| $\mathbf{x}$ | (breeding value for) phenotype of a given genotype |
| $\mathbf{x}^*$ | Optimal phenotype (source) |
| $d$ | Immigration rate |
| $U$ | Genomic mutation rate |
| $\lambda$ | Mutational variance per trait |
| $\mu$ | $\sqrt{U\,\lambda}$ |
| $m$ | Malthusian fitness in the sink, relative to a genotype optimal in the sink |
| $m_D$ | Harshness of stress (fitness distance between source and sink optima) |
| $r_D$ | Decay rate, in the sink, of a genotype optimal in the source $r_D = m_D - r_{\max}$ |
| $m_{source}$ | Fitness of the migrants in the source |
| $m_{migr}$ | Fitness of the migrants in the sink |
| $p_{migr}$ | Probability density of $m_{migr}$ |
| $r_{\max}$ | Maximum absolute fitness (sink) |
| $r$ | Absolute Malthusian fitness: genotypic growth rate $r = r_{\max} + m$ |
| $N(t)$ | Population size at time $t$ |
| $\overline{m}(t)$ | Mean relative fitness |
| $\overline{r}(t)$ | Mean absolute fitness: mean growth rate of the population $\overline{r}(t) = r_{\max} + \overline{m}(t)$ |
| $t_0$ | Establishment time |
| $C_t(z)$ | Cumulant generating function of the relative fitness distribution in the sink |

Table 4.1. – **Main notations**

ous), $t_0$ satisfies $\overline{r}(t_0) = 0$. Depending on the behavior of $\overline{r}(t)$, $t_0$ may therefore be finite (successful establishment) or infinite (establishment failure).

### 4.2.2. Fisher's geometric model

We use Fisher's geometric model (FGM) to describe the relationships between genotypes, phenotypes and fitnesses in each environment. This phenotype-fitness landscape model has the advantage of yielding realistic distributions of mutation effects on fitnesses (Trindade; Sousa; Gordo 2012; Hietpas; Bank; Jensen; Bolon 2013; Tenaillon 2014) and of generating a coupling between stress levels, the distribution of fitnesses among migrants from the source and that among *de novo* random mutants arising in the sink (Anciaux; Chevin; Ronce; Martin 2018).

*Phenotype-fitness relationships in the two environments*. The FGM assumes that each genotype is characterized by a given breeding value for phenotype at $n$ traits (hereafter simply denoted 'phenotype'), namely a vector $\mathbf{x} \in \mathbb{R}^n$. Each environment (the source and the sink) is characterized by a distinct phenotypic optimum. The distance between these optima determines the stress induced by a change of the environment. An optimal phenotype in the sink has maximal absolute fitness $r_{\max}$ (relative fitness $m = 0$) and sets the origin of phenotype space ($\mathbf{x} = 0$). Fitness decreases away from this optimum. Following the classic version of the FGM, Malthusian fitness is a quadratic function of the breeding value $r(\mathbf{x}) = r_{\max} - \|\mathbf{x}\|^2/2$ and $m(\mathbf{x}) = -\|\mathbf{x}\|^2/2$.

In the source, due to a different phenotype optimum $\mathbf{x}^* \in \mathbb{R}^n$, the relative fitness is $m^*(\mathbf{x}) = -\|\mathbf{x} - \mathbf{x}^*\|^2/2$. As the population size is kept constant in the source (see below), only relative fitness matters in this environment. The harshness of stress $m_D > 0$ is the fitness distance between source and sink optima:

$$m_D = -m(\mathbf{x}^*) = \|\mathbf{x}^*\|^2/2. \tag{4.2}$$

The decay rate, in the sink, of a population composed of individuals with the optimal phenotype from the source, is thus $r_D = m_D - r_{\max}$.

*Mutations*. In the two environments, mutations occur at rate $U$ and create independent and identically distributed (iid) random variations $\mathrm{d}\mathbf{x}$ around the phenotype of the parent, for each trait. We assume here a standard Gaussian distribution of the mutation phenotypic effects (Kimura 1965; Lande 1980): $\mathrm{d}\mathbf{x} \sim \mathcal{N}(0, \lambda I_n)$, where $\lambda$ is the mutational variance at each trait, and $I_n$ is the identity matrix in $n$ dimensions. These assumptions induce a distribution of the mutation effects on fitness, given the relative fitness $m_p \leq 0$ of the parent. This distribution has stochastic representation (Martin 2014) $s \sim -m_p - \frac{\lambda}{2}\chi^2_n(-2\, m_p/\lambda)$, where $\chi^2_n(-2\, m_p/\lambda)$ denotes the noncentral chi-square distribution with $n$ degrees of freedom and noncentrality $-2\, m_p/\lambda$. This distribution is detailed elsewhere

(reviewed in Tenaillon 2014), its mean is $\mathbb{E}[s] = -n\,\lambda/2$. Alternatively, it can be characterized by its moment generating function:

$$\mathbb{E}[e^{s\,z}|m_p] = M_*(z)\,e^{\omega(z)\,m_p}, \tag{4.3}$$

with:

$$M_*(z) = \frac{1}{(1+\lambda\,z)^{n/2}} \text{ and } \omega(z) = \frac{-\lambda\,z^2}{1+\lambda\,z}. \tag{4.4}$$

*Migration events.* Migration sends randomly sampled individuals from the source into the sink, at rate $d > 0$ per unit time. Their relative fitness in the sink is $m_{migr}(\mathbf{x}) = -\|\mathbf{x}\|^2/2$, with $\mathbf{x}$ randomly sampled from the source's standing phenotype distribution.

### 4.2.3. Fitness distribution of the migrants

We assume that the distribution of phenotypes in the source is at mutation-selection balance. The resulting equilibrium distribution of phenotypes yields an equilibrium fitness distribution in the source. Under a weak selection strong mutation (WSSM) regime, a simple expression for this equilibrium fitness distribution is (Martin; Roques 2016, equation (10)): $m_{source} \sim -\Gamma(n/2, \mu)$, with $\mu := \sqrt{U\,\lambda}$, where $\Gamma(a, b)$ denotes a gamma deviate with shape $a$ and scale $b$. This WSSM regime can be quantitatively defined by the inequality $U > U_c := n^2\,\lambda/4$ (Martin; Roques 2016, Appendix E).

To understand the dynamics of the fitness distribution in the sink, we need to compute the distribution of the relative fitness of the migrants $m_{migr}$ when they arrive into the sink. In our case, a handy way to describe this distribution is to compute its moment generating function: $e^{\phi(z)} := \mathbb{E}[e^{m_{migr}\,z}]$, for any $z \geq 0$. Computations in Appendix 4.5.1 show that for any $z \geq 0$:

$$\phi(z) = -\frac{n}{2}\ln(1+\mu z) - m_D\,z + \frac{m_D\,\mu\,z^2}{1+\mu\,z}. \tag{4.5}$$

The corresponding distribution of $m_{migr}$ (see Appendix 4.5.1) is:

$$p_{migr}(m) = \begin{cases} \frac{1}{\mu}\,\left(\frac{|m|}{m_D}\right)^{\frac{1}{2}\left(\frac{n}{2}-1\right)}\,e^{\frac{m-m_D}{\mu}}\,I_{\frac{n}{2}-1}\left[\frac{2\sqrt{m_D|m|}}{\mu}\right], \text{ if } m < 0, \\ 0, \text{ if } m \geq 0, \end{cases} \tag{4.6}$$

where $I_\nu$ is the modified Bessel function of the first kind. The accuracy of this formula is illustrated in Fig. 4.1. We observe that the mean absolute fitness of the migrants, which coincides with $\overline{r}(0) = \lim \overline{r}(t)$ as $t \to 0$, is given by:

$$\overline{r}(0) = r_{\max} + \phi'(0) = r_{\max} - m_D - \mu\,n/2 = -r_D - \mu\,n/2, \tag{4.7}$$

with $\phi$ defined by (4.5). This initial growth rate is negative and corresponds to the decay rate ($r_D$) of the mean phenotype from the source (which is optimal there) plus a variance load ($\mu\, n/2$) due to the equilibrium variation around this mean.

The assumption that the individuals from the source are initially decaying ($\overline{r}(0) < 0$) can therefore be expressed by the inequality $r_{\max} - \mu\, n/2 < m_D$.

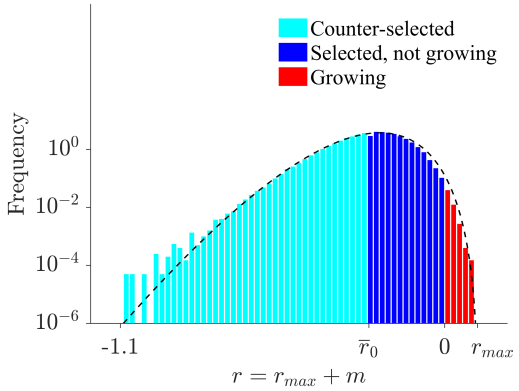



Figure 4.1. – **Distribution of absolute fitness of the migrants in the sink.** The dashed line corresponds to the theoretical expected values of this distribution $p_{migr}(\cdot - r_{\max})$ given by formula (4.6). The histogram corresponds to the distribution of migrants obtained in exact stochastic simulations after reaching the mutation-selection balance in the source (see Section 4.2.5). When the sink is empty, individuals are 'counter-selected' if their fitness is below the mean fitness $\overline{r}(0)$ given by (4.7), 'selected' if their fitness is above $\overline{r}(0)$, and 'growing' if their fitness is positive. The parameter values are $r_{\max} = 0.1$, $U = 0.1$, $m_D = 0.3$, $\lambda = 1/300$, $n = 6$ and $N = 10^6$.

## 4.2.4. Trajectories of fitness in the sink: a PDE approach

Assume that at time $t$, the population in the sink consists of the phenotypes $\{\mathbf{x}_i(t)\}_{i=1,\dots,N(t)}$ (with $N(t) \in \mathbb{N}$), with the corresponding values of relative fitnesses $\{m_i(t)\}_{i=1,\dots,N(t)}$. In the absence of demography and immigration, the dynamics of the fitness distribution is traditionally investigated by a moment closure approximation (Burger 1991; Gerrish; Sniegowski 2012): the variations of the moment of order $k$ depend on the moments of order larger than $(k+1)$ through a linear ordinary differential equation, and the resulting system is solved by assuming that the moments vanish for $k$ larger than some value. A way around this issue is the use of cumulant generating functions (CGFs), which handle all moments in a single function. In a relatively wide class of evolutionary models of mutation and selection, the CGF of the fitness distribution satisfies a partial differential equation (PDE) that can be solved without requiring a moment closure approximation (Martin; Roques 2016, Appendix B). We follow this approach here.

The empirical CGF of the relative fitness in a population of $N(t)$ individuals with fitnesses $m_1(t), \ldots, m_{N(t)}(t)$ is defined by:

$$C_t(z) = \ln\left(\frac{1}{N(t)}\sum_{i=1}^{N(t)} e^{m_i(t)\,z}\right), \tag{4.8}$$

for all $z \geq 0$. The mean fitness and the variance in fitness in the sink can readily be derived from derivatives, with respect to $z$, of the CGF, taken at $z = 0$: $\overline{m}(t) = \partial_z C_t(0)$ (and $\overline{r}(t) = r_{\max} + \partial_z C_t(0)$), and $V(t) = \partial_{zz} C_t(0)$ (the variance in fitness). In the absence of demography and immigration, and under a weak selection strong mutation (WSSM) regime, Martin; Roques 2016, Appendix A derived a deterministic nonlocal PDE for the dynamics of $C_t$. We extend this approach to take into account immigration effects and varying population sizes. This leads to the following PDE (derived in Appendix 4.5.2):

$$\begin{aligned}\partial_t C_t(z) = \underbrace{\partial_z C_t(z) - \partial_z C_t(0)}_{selection} \; & \underbrace{- \mu^2 \left(z^2\, \partial_z C_t(z) + \frac{n}{2} z\right)}_{mutation} \\ & \underbrace{+ \frac{d}{N(t)}\left(e^{\phi(z) - C_t(z)} - 1\right)}_{migration,\ demography}, \; z \geq 0,\end{aligned} \tag{4.9}$$

where we recall that $\mu := \sqrt{U\,\lambda}$. The immigration term depends on $\phi(z)$, which is given by (4.5), and on $N(t)$, which satisfies the ODE (4.1), i.e. $N'(t) = (\partial_z C_t(0) + r_{\max})\, N(t) + d$. This leads to a well-posed coupled system (4.1) & (4.9) which can be solved explicitly, as shown in Appendix 4.5.3.

The selection term in eq. (4.9) stems from the increase in frequency of each lineage proportionally to its Malthusian fitness (frequency-independent selection). The second term is the WSSM approximation ($U > U_c$) to a more complex term (Martin; Roques 2016, Appendix A) describing the effect of mutation: it depends on the current background distribution (on $C_t(z)$) because of the fitness epistasis inherent in the FGM. The last term describes the effect of the inflow of migrants on lineage frequencies. It tends to equate $C_t(z)$ with $\phi(z)$, the CGF of fitnesses among migrants, proportionally to $d/N(t)$, the dilution factor of migrants into the current sink population.

### 4.2.5. Individual-based stochastic simulations

To check the validity of our approach, we used as a benchmark an individual-based, discrete time model of genetic drift, selection, mutation, reproduction and migration with non-overlapping generations.

*Source population.* A standard Wright-Fisher model with constant population

size was used to compute the equilibrium distribution of phenotypes in the source. Our computations were carried out with $N^* = 10^6$ individuals in the source. Each individual $i = 1, \ldots, N^*$ has phenotype $\mathbf{x}_i \in \mathbb{R}^n$ and relative Malthusian fitness $m_i = -\|\mathbf{x}_i - \mathbf{x}^*\|^2/2$, with corresponding Darwinian fitness $e^{m_i}$ (discrete time counterpart of the Malthusian fitness). At each generation, $N^*$ individuals are sampled with replacement proportionally to their Darwinian fitness. Mutations are simulated by randomly drawing, every generation and for each individual, a Poisson number of mutations, with rate $U$. Mutation acts additively on phenotype, with individual effects $d\mathbf{x}$ drawn into an isotropic multivariate Gaussian distribution with variance $\lambda$ per trait (see Section 4.2.2). Simulations were started with a homogeneous population ($\mathbf{x}_i = \mathbf{x}^*$ for all $i$ at initial time) and ran for $20/\sqrt{\mu}$ generations (the predicted time taken to reach a proportion $q$ of the final equilibrium mean fitness is $\text{atanh}(q)/\sqrt{\mu}$, see Appendix E, Section “Characteristic time" in Martin; Roques 2016; with $\text{atanh}(q) = 20$, one can consider that the equilibrium has been reached). An example of the distribution of absolute fitness in the resulting (equilibrium) source population, after migrating into the sink (distribution of $r_{\max} - \|\mathbf{x}_i\|^2/2$) is presented in Fig. 4.1.

*Sink population.* We started with $N(0) = 0$ individuals in the sink. Then, the process to go from generation $t$ to generation $(t+1)$ is divided into three steps: (i) migration: a Poisson number of migrants, with rate $d$, was randomly sampled from the equilibrium source population, and added to the population in the sink; (ii) reproduction, selection and drift: each individual produced a Poisson number of offspring with rate $\exp(r_i) = \exp(r_{\max} + m_i)$ (absolute Darwinian fitness in the sink); (iii) mutation followed the same process as in the source population. The stopping criterion was reached when $N(t) > 1.5 \cdot 10^6$ individuals or $t > 5 \cdot 10^3$ to limit computation times.

All the Matlab® codes to generate individual-based simulations are provided in Supplementary File 1.

# 4.3. Results

## 4.3.1. Trajectories of mean fitness

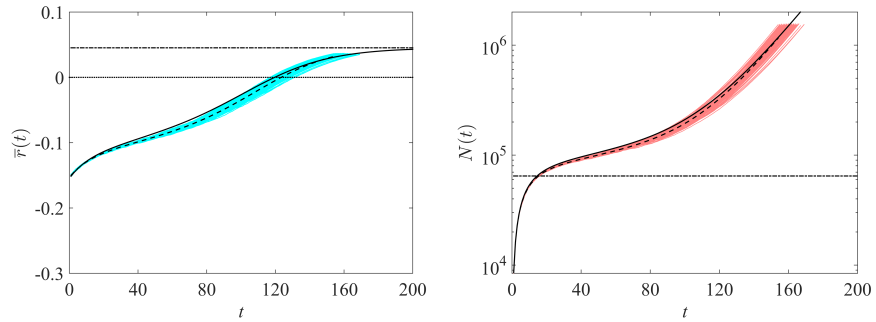
**Dynamics of $\overline{r}(t)$ and $N(t)$.** The system (4.1) & (4.9) leads to an expression for the mean absolute fitness (Appendix 4.5.3):

$$\overline{r}(t) = \frac{f(t) - 1}{\int_0^t f(\tau)\,\mathrm{d}\tau}, \text{ with } f(t) = \exp\left[\left(r_{\max} - \mu\,\frac{n}{2}\right)t + \frac{m_D}{2\,\mu}(e^{-2\,\mu\,t} - 1)\right]. \quad (4.10)$$

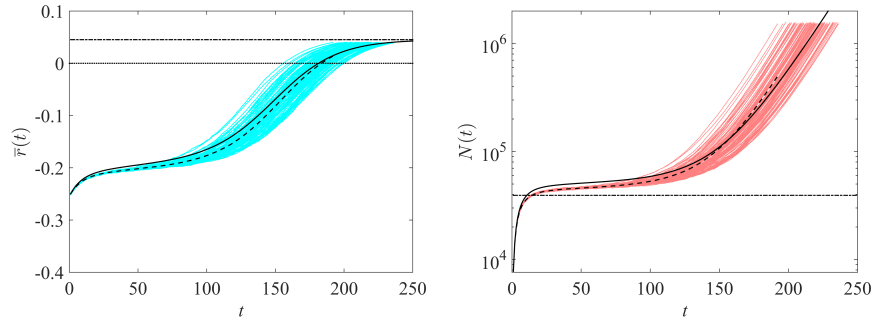
It also leads to an expression for the population size thanks to $N'(t) = \overline{r}(t)\,N(t) + d$. (see eq. (4.16) in Appendix 4.5.3).

The good accuracy of eq. (4.10) is illustrated in Figs. 4.2-4.4, by comparing it with the results of individual-based stochastic simulations, under the WSSM assumption ($U > U_c := n^2\,\lambda/4$). Both the individual-based simulations and the analytic expressions show that sink invasion tends to follow four different phases, which are all the more pronounced as the harshness of stress $m_D$ increases. *Phase 1:* During the first generations, the mean fitness slightly increases; *Phase 2:* The mean fitness remains stable. *Phase 3:* Rapid increase in mean fitness. *Phase 4:* The mean fitness stabilizes at some asymptotic value. In the case of establishment failure (Fig. 4.4), the adaptation process remains in Phase $2$.

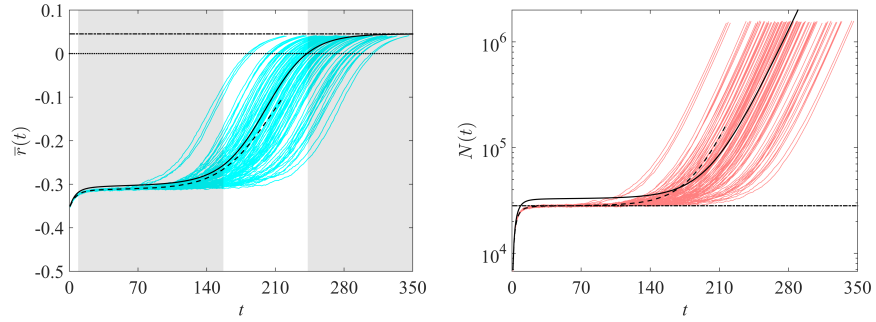
In all cases, formula (4.7) gives an accurate prediction of the mean fitness of the migrants, as shown by the agreement between theoretical and simulated values of $\overline{r}(0)$. Other trajectories, outside of the WSSM regime ($U < U_c$) are presented in Appendix 4.5.4 (and discussed in Section 4.3.3).




(a) $m_D = 0.2$ (b) $m_D = 0.2$



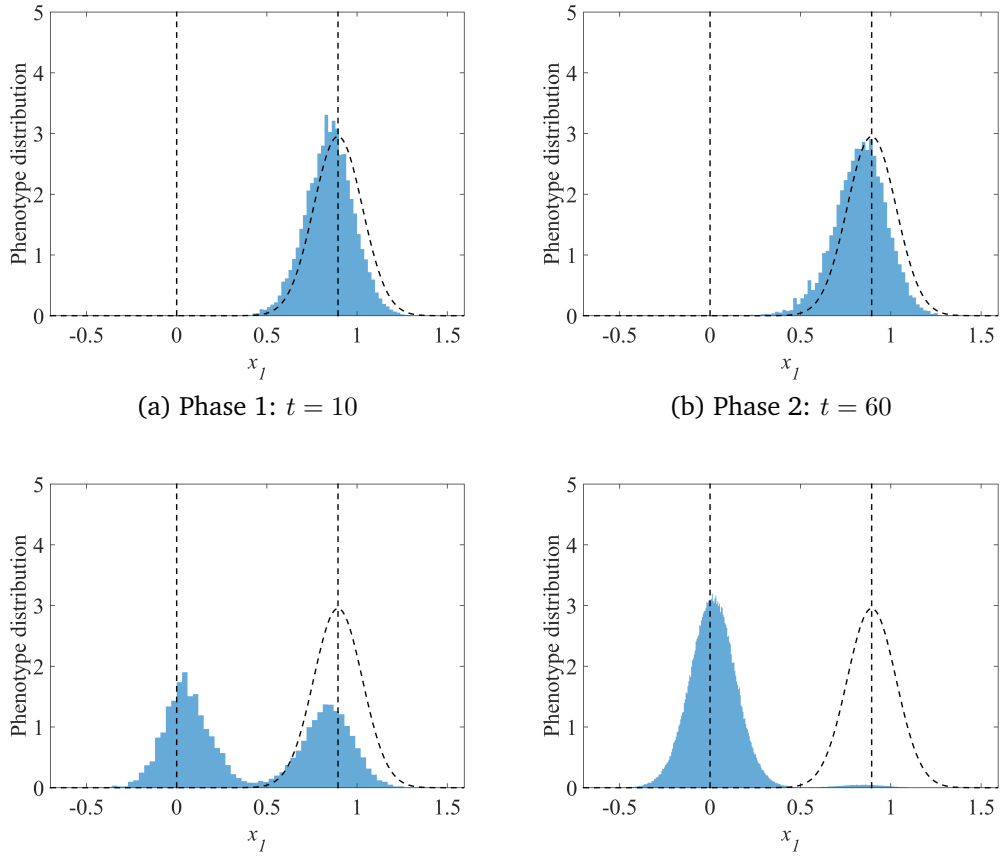

(c) $m_D = 0.3$ (d) $m_D = 0.3$




(e) $m_D = 0.4$ (f) $m_D = 0.4$

Figure 4.2. – **Trajectories of mean fitnesses and population sizes in a WSSM regime, depending on the harshness of stress.** Solid lines: analytical predictions given by formulae (4.1) and (4.10) vs 100 trajectories obtained by individual-based simulations (blue curves for $\overline{r}(t)$ and red curves for $N(t)$; dashed lines: mean values averaged over the 100 populations). Horizontal dashed-dotted lines: theoretical value of $\overline{r}(\infty) = r_{\max} - \mu\, n/2$ (left panels) and equilibrium population size $-d/\overline{r}(0)$ in the absence of adaptation (right panels). The four phases of invasion (Phases 1-4, see main text) are illustrated by distinct shaded areas on panel (e). The parameter values are $U = 0.1$ (thus, $U > U_c = 0.03$, which is consistent with the WSSM regime), $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$. Due to the stopping criterion $N(t) = 1.5 \cdot 10^6$ was reached, the mean values could not be computed over the full time span.

**Phenotypic dynamics over the different phases of invasion.** Obviously the dichotomy into four phases could be deemed somewhat arbitrary, and it is clearly less marked with milder stress (top panels of Fig. 4.2). However, it does convey the qualitative chronology of the whole process in all cases. This can be further understood by exploring the dynamics of the phenotypic distribution over time: a typical example for a single simulation is given in Fig. 4.3, at four times corresponding to each of the four phases. We show here the phenotypic distribution along the one meaningful dimension, that for which the optimum is shifted between source and sink (the optimum in the sink is $0$, and the optimum in the source $\mathbf{x}^* = (\sqrt{2\,m_D}, 0, \ldots, 0)$). The corresponding trajectories of fitness and population size are available in Appendix 4.5.5 (Fig. 4.9). A video file of the phenotype distribution is also available as Supplementary File 2.




(a) Phase 1: $t = 10$

(b) Phase 2: $t = 60$

(c) Phase 3: $t = 220$

(d) Phase 4: $t = 300$

Figure 4.3. – **Phenotype distribution in the sink, along the direction $x_1$.** The vertical dotted lines correspond to the sink ($x_1 = 0$) and source ($x_1 = \sqrt{2\,m_D}$) optima. The black dotted curve corresponds to the theoretical distribution of migrant's phenotypes in the sink (Gaussian distribution, centered at $x_1 = \sqrt{2\,m_D}$, and with variance $\mu = \sqrt{U\,\lambda}$). In all cases, the parameter values are $m_D = 0.4$, $U = 0.1$, $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$.

During Phases 1 and 2, the phenotypic distribution is fairly stable and slightly shifted from the source distribution towards the sink optimum. The short Phase 1 merely witnesses an increase in population size from zero to the semi-stable Phase 2. We suggest that this semi-stable state approximately corresponds to a macroscopic "equilibrium" between migration and selection on the bulk of phenotypes. Here, we conjecture a negligible impact of mutation on this bulk because simulations in the absence of mutation in the sink yield a very similar phenotypic distribution during Phase 2 (Appendix 4.5.10, Fig. 4.12). However, over the course of Phase 2, a second mode slowly appears closer to the sink optimum, due to the invasion of rare, better adapted, phenotypes (generated by the combined effects of rare adapted migrants and *de novo* mutation in the

sink). When this second mode becomes significant in frequency, Phase 3 starts with a rapid increase of the second mode (and of mean fitness), because phenotypic and fitness variance are then maximized. The last Phase 4 corresponds to the new equilibrium dominated by a mutation selection balance around the sink optimum. In the present model without density limitations, migration becomes ultimately negligible as the sink population explodes, and its phenotypic distribution ultimately reaches exactly a new mutation-selection balance.

**Effect of the immigration rate.** Unexpectedly, the value of $\overline{r}(t)$ in formula (4.10) does not depend on the immigration rate $d$. Thus, only the population size dynamics are influenced by the immigration rate, but not the evolutionary dynamics. To understand this phenomenon, we may divide the equation $N'(t) = \overline{r}(t)\,N(t) + d$ by $d$, leading to $P'(t) = \overline{r}(t)P(t) + 1$ with $P(t) = N(t)/d$. Then, we observe that the main system (4.1) & (4.9) can be written in terms of $P(t)$, independently of $N$ and $d$. This means that the ratio $N(t)/d$ is not influenced by $d$. This yields the independence of the evolutionary dynamics of $d$, because the effect of migration on mean fitness in (4.9) only depends on $d/N(t)$.

A simple mathematical argument (Appendix 4.5.6) shows that this property will apply beyond the present model. The result arises for any model where (i) the evolutionary and demographic dynamics in the sink are density-independent (apart from the impact of migration) and (ii) the sink is initially empty (or at least $d \gg N(0)$). This means that it should apply for a broad class of models of asexual evolution in black-hole sinks. Note however, that sex and recombination, for example, necessarily create density-dependent evolution as recombination with migrants affects the genotype frequencies beyond the pure demographic impact of migration.

An intuition for the independence of $\overline{r}(t)$ on $d$ might be framed as follows: if $d$ is increased (resp. decreased), the sink fills in more (resp. less) rapidly, from $N(0) = 0$, proportionally to the increase (resp. decrease) in $d$, at all times. Therefore things cancel out in the migration contribution on frequencies ($d/N(t)$ is unaffected), and this contribution is the only one where $d$ enters the dynamics. Overall increasing or decreasing $d$ thus has no effect on genotype frequency dynamics, although it does affect population sizes. This balanced effect likely exists qualitatively in even more general conditions, but the exact cancelling out only happens with exponential (density-independent) growth/decay, density independent mutation and selection, and an initially empty sink.

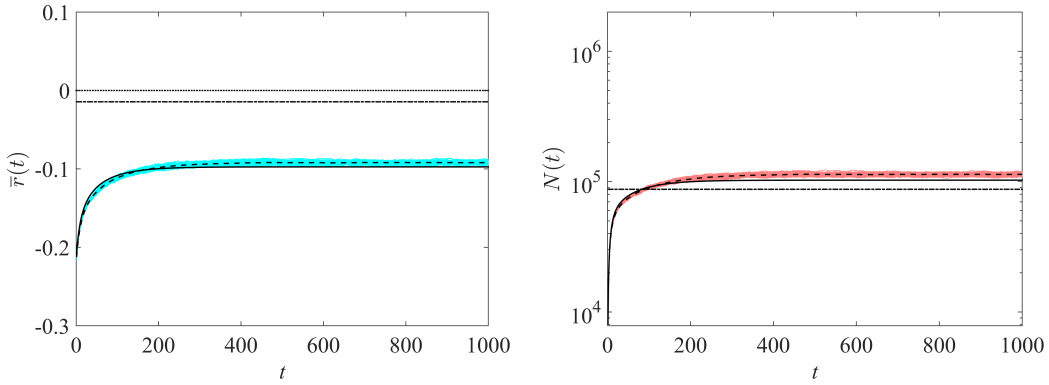
**Large time behavior.** As seen in Fig. 4.2, $\overline{r}(t)$ converges towards an asymptotic value $\overline{r}(\infty)$ at large times. The expression (4.10) shows that this value depends on $r_{\max}$, $\mu$ and $n$. Interestingly, it becomes dependent on the harshness of stress $m_D$, only in the case of establishment failure. More precisely, we get:

$$\begin{array}{l} \text{if } r_{\max} - \mu\, n/2 \geq 0 \text{ then } \overline{r}(\infty) = r_{\max} - \mu\, n/2, \text{ and } N(\infty) = \infty \\ \text{if } r_{\max} - \mu\, n/2 < 0 \text{ then } \overline{r}(\infty) = r_{\max} - \mu\, n/2 - \delta(m_D), \text{ and } N(\infty) = -d/\overline{r}(\infty), \end{array} \tag{4.11}$$

for some function $\delta(m_D)$ such that $m_D > \delta(m_D) > m_D/8$ for $\mu$ large enough (the inequality $\delta(m_D) > m_D/8$ is true whatever the phenotype dimension $n$). When $n$ is large enough, sharper lower bounds can be obtained, e.g. $\delta(m_D) > 3\,m_D/8$ for $n \geq 6$), see Appendix 4.5.7.

These asymptotic results can be interpreted as follows. Below some threshold ($U < U_{lethal} := 4r_{\max}^2/(\lambda\, n^2)$, or equivalently $\mu < \mu_{lethal} := 2r_{\max}/n$), establishment is always successful and the sink population ultimately explodes (as we ignore density-dependence in the sink). As $d/N(\infty) = 0$, the demographic and evolutionary effects of migrants thus become negligible (being diluted in an effectively infinite population). The sink population thus reaches mutation-selection balance, with a mutation load $\mu\, n/2$, as if it was isolated. It ultimately grows exponentially at rate $r_{\max} - \mu\, n/2$ as illustrated in Fig. 4.2.

On the contrary, large mutation rates ($U \geq U_{lethal}$ or equivalently $\mu \geq \mu_{lethal}$) lead to establishment failure, which is a form of lethal mutagenesis (see Bull; Sanjuan; Wilke 2007 for viruses and Bull; Wilke 2008 for bacteria) illustrated in Fig. 4.4. In this regime, the mutation load $\mu\, n/2$ is larger than the absolute maximal fitness $r_{\max}$ in the sink. Therefore, at mutation-selection balance and even in the absence of any migration, the population could never show positive growth: establishment is impossible because the fitness peak is too low, given the mutation rate and effect. We further identify a “jump” of amplitude $\delta(m_D)$ in the equilibrium mean fitness, as $\mu$ increases beyond the lethal mutagenesis threshold (illustrated in Fig. 4.5). Then, the population ultimately reaches a stable size determined by an immigration - decay equilibrium: a migration load can build up at equilibrium ($\delta(m_D)$) together with the mutation load ($\mu\, n/2$). This migration load is produced by the constant inflow of maladapted genotypes from the source and does depend on the harshness of stress $m_D$. It is this migration load that creates the “phase transition" in equilibrium fitness as $\mu$ crosses beyond $\mu_{lethal}$, the lethal mutagenesis threshold (Fig. 4.5). Note, however, that contrary to what happens with sexuals, migrants entering an asexual population do not interbreed with locally adapted genotypes, which simplifies the effect of migration. Note also that, in this lethal mutagenesis regime, the sink population does establish to a stable size, that may be higher than that expected in the absence of mutation and adaptation. However, this is not an establishment in that the population would still get extinct if migration was to be stopped.




(a) $U = 0.44 > U_{lethal}$ (b) $U = 0.44 > U_{lethal}$

Figure 4.4. – **Trajectories of mean fitnesses and population sizes, lethal mutagenesis regime.** Same legend as in Fig. 4.2. Other parameter values are $m_D = 0.2$, $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$, leading to a theoretical threshold value for lethal mutagenesis $U_{lethal} = 4r_{\max}^2/(\lambda\, n^2) = 0.33$. The panel (a) illustrates the bifurcation in the behavior of the equilibrium mean fitness as $r_{\max} - \mu\, n/2$ becomes negative.

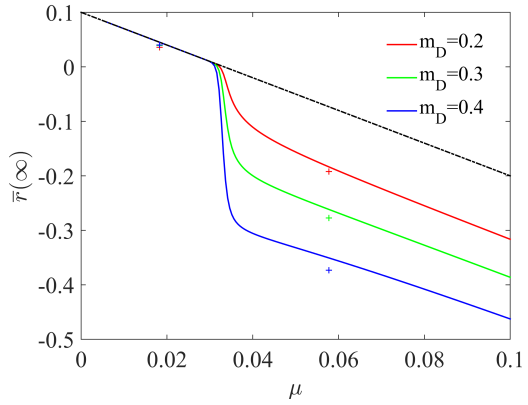



Figure 4.5. – **Mean fitness at large times, dependence with $\mu$ and $m_D$.** The solid lines are the values given by formula (4.11). The crosses correspond to the result of individual-based simulations. The dashed-dot line corresponds to $r_{\max} - \mu\, n/2$; the gap between the dashed-dot line and the solid lines represents the amplitude of the jump $\delta(m_D)$. Parameter values: $r_{\max} = 0.1$, $n = 6$.

### 4.3.2. Establishment time $t_0$

Of critical importance is the waiting time until the sink becomes a source, when this happens, namely the time $t_0$ at which $\overline{r}(t)$ becomes positive. This section is devoted to the analysis of this time.

**Derivation of an analytical expression.** Using the expression (4.10), we can solve the equation $\bar{r}(t_0) = 0$. We recall that, due to our assumptions, $t_0 > 0$, i.e. $\bar{r}(0) = r_{\max} - \mu\, n/2 - m_D < 0$.

The result in (4.11) shows that $t_0 = \infty$ if $r_{\max} - \mu\, n/2 \leq 0$ (establishment failure). In the case of successful establishment ($m_D > \bar{r}(\infty) = r_{\max} - \mu\, n/2 > 0$), the waiting time to this establishment is:

$$t_0 = \frac{1}{2\mu}\left[c + W_0\left(-c\, e^{-c}\right)\right],\ c = \frac{m_D}{r_{\max} - \mu\, n/2}, \tag{4.12}$$

with $W_0$ the principal branch of the Lambert-W function (see Appendix 4.5.8).

First of all, eq. (4.12) shows that the waiting time is independent of the dispersal rate $d$. This was further supported by individual-based simulations (Fig. 4.6a) as $t_0$ was found to drop rapidly to its predicted value as $d$ increases (as the deterministic approximation becomes accurate), to then become independent of $d$. The waiting time shows a transition (around $c = 1$) from $t_0 \approx c/2\mu$ for small $c$ to $t_0 \approx c/\mu$ for large $c$, so the establishment time always increases close to linearly with the harshness of stress $m_D$. This was also the case in individual-based simulations (Fig. 4.6c), at least until stress becomes too strong, compared to mutation and migration. In that case, the sink population remains fairly small for a long time and our deterministic approximation no longer applies, at least in the early phases (1 and 2) of invasion (see Section 4.3.3). Eq. (4.12) also implies that the establishment time $t_0$ decreases with $r_{\max}$ and increases with $n$. The dependence with respect to the mutational parameter $\mu$ is more subtle: as $\mu$ is increased, $t_0(\mu)$ first decreases until $\mu$ reaches an 'optimal value' (minimizing invasion time), then $t_0(\mu)$ increases until $\mu$ reaches the lethal mutagenesis threshold ($\mu_{lethal} = 2\, r_{\max}/n$). This behaviour always holds, as proven analytically in Appendix 4.5.8. This non-monotonous variation of $t_0$ with mutation rate (here with $\mu = \sqrt{U\lambda}$) was also found in individual-based simulations (Fig. 4.6b).

Most of these effects are fairly intuitive: it takes more time to establish from a more maladapted source ($m_D$), with a smaller mutational variance ($U\lambda$), although their particularly simple quantitative effect on $t_0$ was somewhat unexpected. The effect of $r_{\max}$, although quantitatively simple, has multiple aspects. Indeed, $r_{\max}$ affects various parameters of the establishment process, all else being equal: it decreases the initial rate of decay ($\bar{r}(0) = r_{\max} - m_D - \mu\, n/2$) and increases the proportion of migrants that are resistant to the sink environment (fitness peak height) which both speed adaptation. It also increases the ultimate exponential growth rate of the population ($\bar{r}(\infty) = r_{\max} - \mu\, n/2$). The latter effect is likely irrelevant to $t_0$, however, as this growth phase occurs after the establishment time.

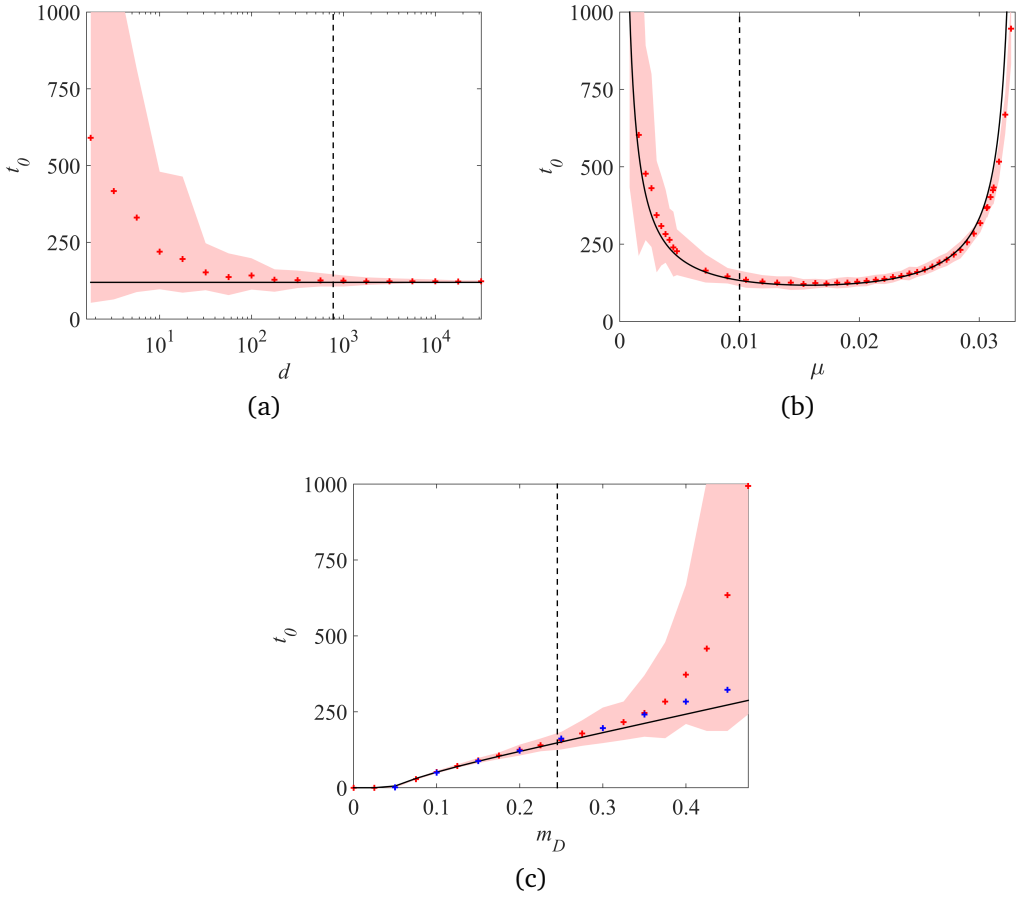



Figure 4.6. – **Establishment time $t_0$, dependence with the immigration rate $d$, the mutational parameter $\mu$ and the harshness of stress $m_D$.** Theoretical value of $t_0$ (black curve) vs value obtained with individual-based simulations (red crosses) and 95% confidence intervals, with fixed $m_D = 0.2$, $U = 0.1$ (panel a), $m_D = 0.2$, $d = 10^3$ (panel b) and fixed $d = 10^3$, $U = 0.1$ (panel c). The vertical dotted lines correspond to the values of $d$, $\mu$ and $m_D$ such that $-dU/\overline{r}(0) = 500$ (panels a and c; this corresponds to a mutant input $N(t)\,U \approx 500$ at small times, ensuring that criterion (ii) in Section 4.3.3 is fulfilled) and $U = U_c$ (panel b). The blue crosses in panel (c) correspond to the establishment time $t_0^I(m_D)$, obtained by individual-based simulations, in the presence of an intermediate habitat with phenotype optimum $\mathbf{x}^I$ such that $\|\mathbf{x}^* - \mathbf{x}^I\|^2/2 = \|\mathbf{x}^I\|^2/2 = m_D/2$. In all cases, the parameter values are $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$.

**Effect of an intermediate sink.** The simulations identify a sharp transition, in the harshness of stress, beyond which establishment does not occur (or occurs at very large times), see Appendix 4.5.9. We see in Fig. 4.6 that as $m_D$ gets close to this threshold, the dependence between $t_0$ and $m_D$ shifts from linear to

superlinear (convex). Based on previous results on evolutionary rescue in the FGM (Anciaux; Chevin; Ronce; Martin 2018), we conjecture that this pattern is inherent to the phenotype fitness landscape model. In the FGM, increased stress (higher $m_D$) is caused by a larger shift in optimum from source to sink. This has two effects, (i) a demographic effect (faster decay of new migrants, on average) and (ii) an evolutionary effect. This latter effect is simply due to the geometry of the landscape. Indeed, when the shift in optimum from source to sink is larger, there are fewer genotypes, in the migrant pool, that can grow in the sink and they tend to grow more slowly. This effect is highly non-linear with stress, showing a sharp transition in the proportion of resistant genotypes beyond some threshold stress (for more details see Anciaux; Chevin; Ronce; Martin 2018).

We argue that this type of dependence has important implications for the potential effect of an intermediate milder sink, with phenotype optimum $\mathbf{x}^I$ in between $\mathbf{x}^*$ (optimum in the source) and $0$ (optimum in the sink), connected by a stepping-stone model of migration. A natural question is then whether the presence of this intermediate sink affects the waiting time to establish in the harsher sink. In that respect, assume that the overall harshness of stress (fitness distance between optima) is the same with and without the intermediate habitat $I$: schematically, $m_D = m_D(\mathbf{x}^* \to 0) = m_D(\mathbf{x}^* \to \mathbf{x}^I) + m_D(\mathbf{x}^I \to 0)$. When $m_D$ is low, $t_0$ is roughly linear with $m_D$ so that it may take a similar time to establish in two step and in one (the sum of intermediate establishment times would be the same as that to establish in a single jump). However, for harsher stress levels where $t_0$ is superlinear with $m_D$, the intermediate habitat could provide a springboard to invade the final sink, if both intermediate jumps are much faster than the leap from source to final sink.

To check this theory, we considered a new individual-based model with an intermediate habitat with phenotype optimum $\mathbf{x}^I$ such that $\|\mathbf{x}^* - \mathbf{x}^I\|^2/2 = \|\mathbf{x}^I\|^2/2 = m_D/2$. The dynamics between the source and the sink are the same as those described in Section 4.2.5. In addition, we assume that (1) the source also sends migrants to the intermediate habitat at a rate $d$; (2) reproduction, selection and drift occur in the intermediate habitat following the same rules as in the sink, until the population $N_I(t)$ in the intermediate habitat reaches the carrying capacity $K = N^*$ (same population size as in the source); (3) the intermediate habitat sends migrants to the ultimate sink, at rate $d\,N_I(t)/N^*$. Then, we computed the time $t_0^I(m_D)$ needed to establish in the final sink, in the presence of the intermediate habitat (value averaged over 100 replicate simulations).

The results presented in Fig. 4.6c (blue crosses) confirm that for small $m_D$, the presence of an intermediate habitat has almost no effect ($t_0^I(m_D) \approx t_0(m_D)$). However, when $m_D$ becomes larger and $t_0(m_D)$ becomes superlinear, the establishment time in the sink is dramatically reduced by the presence of the intermediate sink ($t_0^I(m_D) \ll t_0(m_D)$; e.g., for $m_D = 0.5$, $5 \cdot 10^3 \approx t_0(m_D) \gg t_0^I(m_D) \approx 364$).

**Effect of mutation in the sink on the establishment time.** We have seen in Fig 4.6 b that mutation has a non-monotonous impact on establishment time. However, a higher mutation rate affects both the source equilibrium state and the sink dynamics. A natural question to ask is thus whether local mutation *in the sink* helps or hinders invasion. Indeed, mutation in the FGM (and other models with both deleterious and beneficial mutations) can have antagonistic effects: it generates fitness variance to fuel adaptation but lowers the mean fitness by creating a mutation load. This is of course also true for mutation in the source, but the interaction with migration in the sink makes the outcome less straightforward to grasp.

To tell apart the influences of local mutation on invasion speed, we analyzed (Appendix 4.5.10) a scenario where mutation is absent in the sink, but still active in the source, so that the latter is unchanged. An expression equivalent to eq. (4.10) is obtained in this case for the mean fitness trajectory. We compared the corresponding time to establishment, noted $t_0^0$, with the establishment time $t_0$ to check whether local mutation (in the sink) speeds or slows invasion.

The results in Fig. 4.7 show that local mutation can either slow down or accelerate invasion, depending on the mutational variance ($\mu$) and stress level ($m_D$). For a given level of stress ($m_D$), local mutation tends to speed invasion as long as mutational variance ($\mu$) is limited (left part of the graph) but hinders it when it becomes larger (right part of the graph). The transition from helping to hindering invasion happens at larger $\mu$ values when the stress is harsher (higher $m_D$). It thus appears that the beneficial effect of local mutation in producing variance dominates when mutation is limited while its negative effect in load buildup takes over as $\mu$ is increased. The transition occurs at higher $\mu$ under harsher stress because the former effect is more critical then, while the latter is roughly independent of stress. This pattern illustrates quite strikingly the complex implications, for adaptation dynamics, of the ambivalent nature of mutation in the FGM.

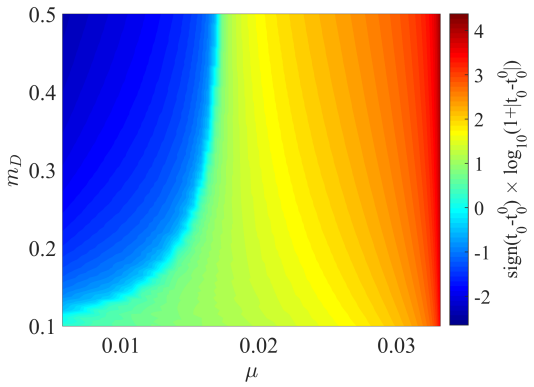



Figure 4.7. – **Comparison between the establishment times $t_0$ (with mutation in the sink) and $t_0^0$ (without mutation in the sink).** The heat map corresponds to $\text{sign}(t_0 - t_0^0)\ \log_{10}(1 + |t_0 - t_0^0|)$: negative values indicate that $t_0^0 > t_0$ (faster establishment with mutation in the sink) and positive values indicate that $t_0 > t_0^0$ (faster establishment without mutation in the sink).

### 4.3.3. Range of validity of the model

We explored the range of validity of the analytical model by comparing theory and simulations over a wide range of parameter values. The raw results are given in Appendix 4.5.9. Overall, the model is more accurate as $U$ and $d$ increase and $m_D$ (equivalently, $r_D = m_D - r_{\max}$), $n$ and $\lambda$ decrease. More precisely, theoretical and numerical analysis yield two (*a priori* conservative) conditions that should lead to the model being accurate: (i) $U \geq U_c = n^2\lambda/4$, for the WSSM to apply; (ii) $dU/r_D \gg 1$, for the large $d$ approximation to apply.
Below we detail each criterion, their robustness and possible empirical insight on their realism.

**Criterion (i):** it is formally derived in Appendix E of (Martin; Roques 2016) and guarantees that the mutation term associated with the FGM linearizes to produce an analytically tractable PDE. While the model is indeed accurate whenever $U > U_c$, it remains reasonably so even at fairly lower mutation rates. Even for mutation rates $U = U_c/30$ (but keeping a large $d$), $\overline{r}(t)$ and $N(t)$ from eq. (4.10) still accurately capture the average trajectories (Fig. 4.8), although the length of Phase 2 in the numerical simulations becomes more variable, around this average, as $U$ is decreased. Consistently, Fig. 4.6b shows that the invasion time in eq. (4.12) accurately captures the average of simulations far below $U = U_c$, with larger variability around this mean as $U$ decreases.

As an example, empirical estimates in *E. coli*, based on a recent mutation accumulation experiment (Trindade; Perfeito; Gordo 2010) suggest $U \in [0.004, 0.006]$ and $\mathbb{E}[s] = n\lambda/2 \in [0.02, 0.04]$ (mean effect of mutations on fitness), which yields $U/U_c \in [0.2, 0.6]$ for $n = 1$ and $U/U_c \in [0.033, 0.1]$ for $n = 6$. This suggests that *E. coli* may lie somewhere below the critical mutation rate, at a similar order.

Note however that estimates of these quantities are fairly scarce (even in this well studied biological model) and seem to vary substantially across experiments (medium, strain, growth conditions). We suspect that viruses (especially RNA viruses) may lie well within $U \geq U_c$, while bacteria may vary widely around $U = U_c$. Obviously any proper statement on this issue would require a full review of empirical estimates (appropriately scaled in consistent time units), wherever available.

**Criterion (ii):** this criterion, which is confirmed by the simulations in Appendix 4.5.9 and Fig. 4.6 panels (a) and (c), stems from the following argument: the early population size in the sink is of order $N(t) \approx d/|\overline{r}(0)|$ (no evolution), with $|\overline{r}(0)| = |r_D + \mu\, n/2| \approx r_D$ (when $\mu \ll r_D$). Thus whenever $d \gg r_D/U$, the mutant input $N(t)U$ in the sink population quickly reaches a large value $N(t)\, U \approx d\, U/r_D \gg 1$ and only increases later on. Adaptive evolution can then take place within the sink, in a way that is accurately captured by a deterministic approximation (see the dotted lines in Fig. 4.6). Conversely, when $d$ is smaller and/or $r_D$ is larger, the early population size in the sink is small, so that the deterministic approximation does not apply anymore. In this case, we see that the time $t_0$ is much more variable, and increases on average with smaller $d$ and larger $r_D$ (or equivalently $m_D$), see Fig. 4.6.

Empirically evaluating the criterion (ii) requires estimates of $d, U, r_D$ on the same timescale (hours, days, generations) in a well defined sink. Such estimates should be possible from dedicated experiments controlling the immigration rate, in strains with known mutational parameters, and environmental stresses with well characterized demographic effect. They would greatly help our understanding of source-sink dynamics. However, to the best of our knowledge, they are not available to date.

## 4.4. Discussion

We derived an analytically tractable PDE-ODE framework describing evolutionary and demographic dynamics of asexuals in a source-sink system. Comparison with individual-based stochastic simulations shows that the approach is accurate in the WSSM regime (large mutation rates compared to mutation effects) and with a large migration rate, and seems robust to mild deviations from this regime. This approach reveals the typical shape of the trajectories of mean fitness and population sizes in a sink: (1) in the case of establishment failure, after a brief increase, the mean fitness remains stable at some negative level which depends on the harshness of stress; (2) in the case of successful establishment, this “plateau" is followed by a sudden increase in mean fitness up to the point where it becomes positive and the sink becomes a source. Note that here, we ignored density dependent effects in the sink, so that mean fitness ultimately converges towards an equilibrium that is independent of any migration effect, the latter being diluted

into an exploding population.

The three first phases predicted by the model, for the case of successful establishment, are qualitatively observed in (Dennehy; Friedenberg; McBride; Holt, et al. 2010), an experimental study of invasion of a black-hole sink (an asexual bacteriophage shifting to a new bacterial host). The “host shift" scenario in their Fig. 3 corresponds roughly to our scenario with a population evolved on the native host sending migrants to a new host. The conditions may differ however as the population may not be initially at equilibrium in the native host at the onset of migration. Yet, the dynamics are qualitatively similar to those in our Fig. 4.2, although the time resolution in the data is too limited to claim or test any quantitative agreement. An extension of the present work could be to allow for non-equilibrium source populations, which can readily be handled by the PDE (4.9) (reformulating $\phi(z) = \phi(z,t)$). However, our analytical result on $t_0$ does rely on an equilibrium source population. Note also that the four phases identified here are observed, in simulations, even in the low $d$ or low $U$ regimes where our analytical derivations can break down quantitatively. Therefore, while the model may provide qualitatively robust insight, quantitative analyses are necessary to really test its predictions. This would ideally include associated measures of decay rates $r_D$, mutation rate $U$ and ideally maximal possible growth rate $r_{\max}$, with a known immigration rate $d$.

Quite unexpectedly, the evolutionary dynamics (especially the waiting time $t_0$ to establishment) do not depend on the immigration rate. This emerges mathematically from the fact that the evolutionary dynamics only depend on the population size through the ratio $N(t)/d$ between the current population size and the immigration rate, this ratio itself remaining independent of $d$. This is confirmed by stochastic individual-based simulations (Fig. 4.6a): establishment time roughly decreases as $1/d$ when $d$ is small but indeed stabilizes as $d$ becomes larger. More precisely, the result on the independence of $t_0$ with respect to $d$ should always hold with an initially empty sink and when $dU/r_D \gg 1$ (see Section 4.3.3). In this case, the mutant input in the sink population is always large enough to enable our deterministic framework to accurately capture the evolution in the sink. This result *a priori* extends to any model where evolution and demography are density-independent. However density dependent effects on demography or evolution (including sexual reproduction) might alter this outcome. Yet, we argue that purely demographic effects due to a finite carrying capacity in the sink environment should have limited impact on the conclusions of our model, up until establishment time (as long as $K$ is large enough).

In a black-hole sink experiment Perron; Gonzalez; Buckling 2007 studied the evolution of resistance to two lethal doses of antibiotics and their combinations in the bacterium *Pseudomonas aeruginosa* (also asexual). Their experiment differs from our scenario in that the sink populations were initially filled with many “naive" individuals ($N_0 \gg 1$, amounting to an initial large single immigration event). The authors did notice that immigration rate $d$ affected population

densities, but this is not directly a test of our model: our deterministic model also predicts that $N(t)$ should depend on $d$, only the mean fitness and time to establishment do not.

The independence between $t_0$ and $d$ is counter-intuitive if we consider sink invasion as a repeated evolutionary rescue 'experiment'. Indeed, the immigration process in the sink could also be seen as a Poisson process of incoming new lineages (from the source), each having a given probability $p_R$ to yield a rescue in the future (in the absence of new immigration), hence to ultimately turn the sink into a source. This probability $p_R$ can be computed from evolutionary rescue theory, with various flavours: see (Orr; Unckless 2014) for a context-independent single allele rescue model or, in the case of the FGM, using results in (Anciaux; Chevin; Ronce; Martin 2018). By basic properties of Poisson processes, the waiting time $t_1$ to the first arrival, in the sink, of such a future rescue lineage should be exponential with mean $1/(d\,p_R)$, thus decreasing as $1/d$.

However, this waiting time is different from the one computed here. Our $t_0$ denotes the time at which the mean fitness of the sink population becomes positive in the absence of immigration, hence the time at which the sink has truly become a source. The evolutionary rescue approach above computes the time $t_1$ at which a lineage *ultimately* destined to produce a resistant genotype, enters the sink. This lineage may be very rare by $t = t_1$, it may even not be resistant itself but only destined to produce a mutant offspring that will be. The time at which the sink will *de facto* be a positively growing source can thus be far later. A study and comparison of both waiting times is interesting and feasible, but beyond the scope of the present paper. This remark, however, has one key implication: migration may be stopped long before $t_0$ and the sink may still ultimately become a source, with some probability (even if this will be 'visible' much later).

Some insight into the possible effects of management strategies, e.g. quarantine ($d$), lethal mutagenesis ($U$), prophylaxis ($m_D$ and $r_{\max}$), can be developed from the results presented here.

Migration (propagule pressure) is considered an important determinant of the success of biological invasions in ecology (Von Holle; Simberloff 2005; Lockwood; Cassey; Blackburn 2005). Consistently, it has been shown that the factors increasing potential contacts between human populations and an established animal pathogen or its host tend to increase the risk of emergence of infectious diseases (Morse 2001). Under the 'repeated rescue approach' above, it is indeed expected that emergence risk should increase as 1/contact rate. However, the present work shows that the time at which this emergence will be *de facto* effective (visible) may be unaffected by this contact rate. This means that care must be taken in the criteria chosen to evaluate strategies, and between the minimization of emergence risk *vs.* emergence time.

The use of a chemical mutagen to avoid the adaptation of a microbial pathogen and the breakdown of drugs is grounded in lethal mutagenesis theory (Bull; Sanjuan; Wilke 2007; Bull; Wilke 2008). Our approach successfully captures the

occurrence of this phenomenon: the establishment fails when the mutation rate $U$ exceeds a certain threshold, which depends on $r_{\max}$, on the mutational variance $\lambda$ and on the dimension of the phenotypic space. Additionally, once this threshold is reached, the equilibrium mean fitness ceases to depend linearly on the mutational parameter ($\mu = \sqrt{U\,\lambda}$), but rapidly decays (see Fig. 4.5). The existence of this negative “jump" in the equilibrium mean fitness, whose magnitude depends on the harshness of stress, leaves no room for evolutionary rescue. Conversely, our approach also reveals that below the lethal mutagenesis threshold, increasing the mutation rate decreases the establishment time as $1/\sqrt{U}$. Hence, the use of a mutagen may be a double-edged sword since it can both hamper or increase the potential for adaptation in the sink.

As expected, the establishment time $t_0$ increases with the harshness of stress $m_D$; the population simply needs more time to adapt to more stressful environmental conditions. Increasing $m_D$ or decreasing $r_{\max}$, whenever possible, are probably the safest ways to reduce the risks of biological invasions through adaptive processes or cross-species transmissions of pathogens (in both low and high $d$ regimes). The precise dependence of $t_0$ with respect to $m_D$ brings us further valuable information. As long as our approach is valid (not too large stresses, leading to finite establishment times), a linear dependence emerges. It suggests that, in a more complex environment with a source and several neighbouring sinks connected by a stepping stone model of migrations, the exact pathway before establishment occurs in a given sink does not really matter. Only the sum of the stresses due to habitat shifts has an effect on the overall time needed to establish in the whole system. Conversely, for larger stress values our analytical approach is not valid, and the numerical simulations indicate a convex (surlinear) dependence of $t_0$ with respect to $m_D$. In such case, for a fixed value of the cumulated stress, the establishment time in the sink could be drastically reduced by the presence of intermediate sink habitats.

This result, which needs to be confirmed by more realistic modelling approaches and empirical testing, might have applications in understanding the role of so-called “preadaptation" in biological invasions. Recent adaptation to one or more facets of the environment within the native range has been proposed as a factor facilitating invasions to similar environments (e.g. Hufbauer; Facon; Ravigné; Turgeon, et al. 2012, anthropogenically induced adaptation to invade). Our results suggest that preadaptation might only reduces the overall time to invasion (i.e., taking the preadapation period into account) only when invading highly stressful habitats.

The effect of a given environmental challenge, and thus their joint effects when combined (Rex Consortium 2013), might be modelled in various ways in a fitness landscape framework (see also discussions in Harmand; Gallet; Jabbour-Zahab; Martin, et al. 2017; Anciaux; Chevin; Ronce; Martin 2018). The first natural option is to consider that multiple stresses tend to pull the optimum further away, and possibly lower the fitness peak $r_{\max}$. In the simplified isotropic model studied

here, a larger shift in optimum amounts to increasing $m_D$. However, a possibly more realistic anisotropic version, with some directions favored by mutation or selection, might lead to directional effects (where two optima at the same distance are not equally easy to reach) and be particularly relevant to multiple stress scenarios. Such a more complex model could be handled by focusing on a single dominant direction (discussed in Anciaux; Chevin; Ronce; Martin 2018), or by following multiple fitness components (one per direction, Hamel et al. in prep).

Clearly, many developments are possible and could prove useful to understand how qualitative and quantitative aspects of environmental stresses may affect rescue and invasion. The present isotropic approach provides a simple, tractable null model for the latter, where all environmental effects are summarized by their measurable effects on $m_D$, $U\,\lambda$ and $r_{\max}$. We hope it will foster the empirical study of source-sinks with associated measurements of these key parameters.

## 4.5. Appendix

### 4.5.1. Fitness distribution of the migrants: derivation of formulae (4.5) and (4.6)

Consider an individual with phenotype $\mathbf{x}$. Its fitness in the source is $m_{source} = -\|\mathbf{x}-\mathbf{x}^\star\|^2/2$, where $\mathbf{x}^\star$ is the optimal phenotype in the source, whereas its fitness in the sink is $m_{migr} = -\|\mathbf{x}\|^2/2$. We observe that:

$$\begin{aligned} m_{migr} &= -\frac{\|\mathbf{x}-\mathbf{x}^\star+\mathbf{x}^\star\|^2}{2}, \\ &= -\frac{\|\mathbf{x}-\mathbf{x}^\star\|^2+\|\mathbf{x}^\star\|^2+2(\mathbf{x}-\mathbf{x}^\star)\cdot\mathbf{x}^\star}{2}, \\ &= m_{source} - \frac{\|\mathbf{x}^\star\|^2}{2} - \|\mathbf{x}-\mathbf{x}^\star\|\;\|\mathbf{x}^\star\|u, \\ &= m_{source} - m_D - 2\sqrt{m_D|m_{source}|}u, \end{aligned} \tag{4.13}$$

with $m_D = \|\mathbf{x}^\star\|^2/2$ and a constant $u \in [-1,1]$. As the source is assumed to be at the mutation-selection equilibrium, the distribution of fitness in the source satisfies $m_{source} \sim -\Gamma(n/2,\mu)$ (Martin; Roques 2016, equation (10)) and the corresponding moment generating function is $M_{m_{source}}(z) = (1+\mu z)^{-n/2}$. The results in (Martin; Lenormand 2015) show that $u$ is a random variable with moment generating function:

$$M_u(z) := \mathbb{E}[e^{uz}] = {}_0F_1(n/2,\ z^2/4),$$

with ${}_0F_1$ the hypergeometric function, defined by:

$$ {}_0F_1(\theta, z) = \sum_{k=0}^{\infty} \frac{1}{\theta(\theta+1)\cdots(\theta+k-1)} \frac{z^k}{k!}. $$

Let us first compute the moment generating function $M_{migr}(z) := \mathbb{E}[e^{m_{migr}z}]$. We have:

$$ M_{migr}(z) = \mathbb{E}[\mathbb{E}[e^{m_{migr}z}|m_{source}]], $$

and using (4.13),

$$ \begin{aligned} M_{migr}(z) &= \mathbb{E}\left[e^{m_{source}z} M_u\left(-2\sqrt{m_D|m_{source}|}z\right)\right] e^{-m_D z}, \\ &= \mathbb{E}\left[e^{m_{source}z}\ {}_0F_1\left(n/2,\ -m_D m_{source} z^2\right)\right] e^{-m_D z}. \end{aligned} $$

Thanks to the definition of the hypergeometric function ${}_0F_1(n/2, z)$, we get:

$$ \begin{aligned} M_{migr}(z) &= \sum_{k=0}^{\infty} \frac{(-m_D)^k}{n/2(n/2+1)\cdots(n/2+k-1)} \frac{z^{2k}}{k!} \mathbb{E}[e^{m_{source}z} m_{source}^k] e^{-m_D z}, \\ &= \sum_{k=0}^{\infty} \frac{(-m_D)^k}{n/2(n/2+1)\cdots(n/2+k-1)} \frac{z^{2k}}{k!} M^{(k)}_{m_{source}}(z) e^{-m_D z}, \end{aligned} $$

with $M^{(k)}_{m_{source}}(z)$ the k$^{\text{th}}$ derivative of $M_{m_{source}}(z)$ with respect to $z$. Thus,

$$ \begin{aligned} M_{migr}(z) &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{m_D \mu z^2}{1+\mu z}\right)^k (1+\mu z)^{-n/2} e^{-m_D z}, \\ &= \frac{1}{(1+\mu z)^{n/2}} \cdot \exp\left[-m_D z + \frac{m_D \mu z^2}{1+\mu z}\right]. \end{aligned} $$

Setting $\phi(z) = \ln\left(M_{migr}(z)\right)$, we obtain formula (4.5).

Let us now show that the distribution of the migrants in the sink satisfies (4.6). Let $p_{migr}$ be defined by (4.6). We just have to check that the moment generating function of $p_{migr}$ is $M_{migr}$:

$$ \begin{aligned} \int_{-\infty}^{0} e^{zx} p_{migr}(x)\mathrm{d}x &= \int_{-\infty}^{0} e^{zx} \frac{1}{\mu} \left(\frac{|x|}{m_D}\right)^{\frac{n/2-1}{2}} e^{\frac{x-m_D}{\mu}} I_{\frac{n}{2}-1}\left[\frac{2\sqrt{m_D|x|}}{\mu}\right] \mathrm{d}x, \\ &= e^{-m_D/\mu} \int_{-\infty}^{0} \sum_{p=0}^{\infty} e^{(z+1/\mu)x} \frac{m_D^p}{\mu^{2p+n/2}} \cdot \frac{1}{p!} \cdot \frac{|x|^{p+n/2-1}}{\mathbf{\Gamma}(p+n/2)} \mathrm{d}x, \\ &= e^{-m_D/\mu} \sum_{p=0}^{\infty} \frac{m_D^p}{\mu^{2p+n/2}} \cdot \frac{1}{p!} \cdot \frac{1}{\mathbf{\Gamma}(p+n/2)} \int_{-\infty}^{0} e^{(z+1/\mu)x} |x|^{p+n/2-1} \mathrm{d}x, \end{aligned} $$

where $I_\nu$ is the modified Bessel function of the first kind and $\mathbf{\Gamma}$ the gamma

function.

Now, for all positive numbers $a$ and $b$, we have:

$$\int_{-\infty}^{0} e^{ax}|x|^{b-1}\mathrm{d}x = \frac{1}{a^b}\int_{0}^{\infty} e^{-x}|x|^{b-1}\mathrm{d}x = \frac{\mathbf{\Gamma}(b)}{a^b}.$$

Therefore, we get, for $z > -1/\mu$:

$$\begin{aligned}
\int_{-\infty}^{0} e^{zx} p_{migr}(x)\mathrm{d}x &= e^{-m_D/\mu}\sum_{p=0}^{\infty}\frac{m_D^p}{\mu^{2p+n/2}}\cdot\frac{1}{p!}\cdot\frac{1}{\mathbf{\Gamma}(p+n/2)}\,\frac{\mathbf{\Gamma}(p+n/2)}{(z+1/\mu)^{p+n/2}},\\
&= \frac{e^{-m_D/\mu}}{(1+\mu z)^{n/2}}\sum_{p=0}^{\infty}\left(\frac{m_D/\mu}{1+\mu z}\right)^p\cdot\frac{1}{p!},\\
&= \frac{e^{-m_D/\mu}}{(1+\mu z)^{n/2}}\exp\left(\frac{m_D/\mu}{1+\mu z}\right),\\
&= \frac{1}{(1+\mu z)^{n/2}}\exp\left(-\frac{m_D z}{1+\mu z}\right).
\end{aligned}$$

This is consistent with formula (4.5), which proves that the expression (4.6) is correct.

### 4.5.2. PDE satisfied by the CGF of the fitness distribution

In the WSSM regime, and in the absence of immigration, Martin; Roques 2016 (see Appendix E, equation (E5)) have shown that the CGF of the fitness distribution satisfies the following equation:

$$\partial_t C_t(z) = \partial_z C_t(z) - \partial_z C_t(0) - \mu^2\left(z^2\,\partial_z C_t(z) + \frac{n}{2}z\right),\ z \geq 0.$$

We derive here the additional term in (4.9), which describes the effect of immigration on the CGF.

In that respect, we consider a discrete population of size $N(t) \in \mathbb{N}$ at time $t$, and the corresponding fitnesses $(m_1(t), \ldots, m_{N(t)}(t))$. We define the "empirical" moment generating function:

$$M_t(z) := \frac{1}{N(t)}\sum_{i=1}^{N(t)} e^{m_i(t)\,z}.$$

Assuming a Poisson number of immigration events, with rate $d$ per unit time (see Section 4.2.5), for $\Delta t$ small enough, the probability that a single immigration events occurs during $(t, t+\Delta t)$ is approximately $d\,\Delta t$. The probability that several immigration events occur during this time interval is close to $0$. Therefore, the expected change in the moment generating function during $\Delta t$, conditionally on

the fitness $m_{migr}$ of the unique migrant, is:

$$\begin{aligned}\Delta M_t(z|m_{migr}) &= d\ \Delta t \left[\frac{1}{N(t)+1}\left(\sum_{i=1}^{N(t)} e^{m_i(t)\ z} + e^{m_{migr}\ z}\right) - \frac{1}{N(t)}\sum_{i=1}^{N(t)} e^{m_i(t)\ z}\right], \\ &= d\ \Delta t \left[\frac{e^{m_{migr}\ z}}{N(t)+1} - \frac{M_t(z)}{N(t)+1}\right].\end{aligned}$$

Taking expectation over the distribution of $m_{migr}$ (see Appendix 4.5.1 for more details on the distribution of $m_{migr}$), we get:

$$\Delta M_t(z) = \frac{d\ \Delta t}{N(t)+1}\left(e^{\phi(z)} - M_t(z)\right),$$

with $\phi(z) = \ln\left(\mathbb{E}\left[e^{m_{migr}\ z}\right]\right)$. The corresponding change in the CGF $C_t(z) = \ln M_t(z)$ is $\Delta C_t(z) \approx \Delta M_t(z)/M_t(z)$. Thus, the approximation of large population size implies that:

$$\Delta C_t(z) \approx \frac{d\ \Delta t}{N(t)}\left(e^{\phi(z)-C_t(z)} - 1\right).$$

Finally, dividing the above expression by $\Delta t$ and passing to the limit $\Delta t \to 0$, we obtain the last term in (4.9), which describes the effect of immigration on the CGF:

$$\frac{d}{N(t)}\left(e^{\phi(z)-C_t(z)} - 1\right). \tag{4.14}$$

### 4.5.3. Solution of the system (4.1) & (4.9)

This section is devoted to the mathematical study of the system (4.1) & (4.9). We rewrite it in the following form:

$$\begin{cases} \partial_t C_t(z) = \alpha(z)\partial_z C_t(z) - \overline{m}(t) + \beta(z) + \frac{d}{N(t)}\left(e^{\phi(z)-C_t(z)} - 1\right), \\ N'(t) = N(t)\ (r_{\max} + \overline{m}(t)) + d, \\ C_t(0) = 0, \\ N(0) = 0, \end{cases} \tag{4.15}$$

with $t > 0$ and $z \geq 0$, and where $\overline{m}(t) = \partial_z C_t(0)$, $d \geq 0$, $\alpha(z) := 1 - \mu^2\ z^2$, $\beta(z) := -\mu\, n\, z/2$.

We can easily check that the sink is not empty at each time $t > 0$:

**Lemma 4.** *Assume that $\overline{m}$ is continuous over $[0, \infty)$. Then, at all time $t > 0$, we have $N(t) > 0$.*

*Proof.* For $\varepsilon > 0$ small enough, as $N'(0) = d > 0$, we have $N(t) > 0$ for all $t \in (0, \varepsilon]$. Additionally, for all $t \geq \varepsilon$,

$$N(t) = e^{\int_\varepsilon^t (r_{\max} + \overline{m}(s))\,\mathrm{d}s} \left( N(\varepsilon) + d \int_\varepsilon^t e^{-\int_\varepsilon^v (r_{\max} + \overline{m}(s))\,\mathrm{d}s}\, dv \right) > 0. \tag{4.16}$$

□

Let $N(t)$, $C_t(z)$ be a solution of (4.15), such that $\overline{m}$ is continuous over $[0, \infty)$. Set $D_t(z) = C_t(y(z))$, with $y(z) = \tanh(\mu z)/\mu$ which satisfies:

$$\begin{cases} y'(z) = \alpha(y(z)), \\ y(0) = 0, \end{cases}$$

so that:

$$\partial_t D_t(z) = \partial_t C_t(y(z)) \text{ and } \partial_z D_t(z) = \alpha(y(z)) \partial_z C_t(y(z)).$$

Thus, $D_t(z)$ satisfies the simpler equation:

$$\partial_t D_t(z) = \partial_z D_t(z) - \overline{m}(t) + \beta(y(z)) + \frac{d}{N(t)} \left( e^{\phi(y(z)) - D_t(z)} - 1 \right),$$

with $\overline{m}(t) = \partial_z D_t(0)$.

Using the method of characteristics, we derive an analytic expression for $D_t(z)$. Fix $z \geq 0$ and denote for all $z \geq t > 0$:

$$v(t) = \exp(D_t(z - t)).$$

The function $v \in C^1((0, z])$ satisfies for all $t \in (0, z)$:

$$\begin{aligned} v'(t) &= \left(\partial_t D_t(z-t) - \partial_z D_t(z-t)\right) v(t), \\ &= \left[ \beta(y(z-t)) - \overline{m}(t) - \frac{d}{N(t)} \right] v(t) + \frac{d}{N(t)}\, e^{\phi(y(z-t))}, \\ &= \left[ \beta(y(z-t)) - \frac{N'(t)}{N(t)} + r_{\max} \right] v(t) + \frac{d}{N(t)}\, e^{\phi(y(z-t))}, \end{aligned}$$

thanks to $N'(t) = (r_{\max} + \overline{m}(t))N(t) + d$. Let us fix times $0 < \varepsilon < t$. By Lemma 4, we know that $N(s) > 0$, for all $s \in [\varepsilon, t]$ and so $v(t)$ is given by:

$$\begin{aligned} v(t) = \exp &\left[ \int_\varepsilon^t \left( \beta(y(z-\tau)) - \frac{N'(\tau)}{N(\tau)} + r_{\max} \right) \mathrm{d}\tau \right] \\ &\left[ e^{C(\varepsilon, y(z))} + \int_\varepsilon^t \frac{d\, e^{\phi(y(z-\tau))}}{N(\tau)} \exp\left( - \int_\varepsilon^\tau \left( \beta(y(z-s)) - \frac{N'(s)}{N(s)} + r_{\max} \right) \mathrm{d}s \right) \mathrm{d}\tau \right]. \end{aligned}$$

As $\int_\varepsilon^t \frac{N'(s)}{N(s)}\mathrm{d}s = \ln N(t) - \ln N(\varepsilon)$, we can simplify the last expression to:

$$v(t) = \exp\left[-\ln N(t) + \int_\varepsilon^t (\beta(y(z-\tau)) + r_{\max})\,\mathrm{d}\tau\right] \left[N(\varepsilon)e^{C(\varepsilon,y(z))} + \int_\varepsilon^t d\, e^{\phi(y(z-\tau))} \exp\left(-\int_\varepsilon^\tau (\beta(y(z-s)) + r_{\max})\,\mathrm{d}s\right)\mathrm{d}\tau\right].$$

Taking the limit as $\varepsilon$ tends to 0 and using the fact that the initial population in the sink is $N(0) = 0$, the above expression can be simplified to:

$$v(t) = d\int_0^t e^{\phi(y(z-\tau)) - \int_0^\tau (\beta(y(z-s)) + r_{\max})\mathrm{d}s}\mathrm{d}\tau \cdot \exp\left[-\ln N(t) + \int_0^t (\beta(y(z-\tau)) + r_{\max})\,\mathrm{d}\tau\right].$$

Hence, by reversing the characteristics, we get:

$$D_t(z) = \int_0^t \beta(y(z+\tau))\mathrm{d}\tau - \ln\left(N(t)\right) + r_{\max}t + \ln\left[d\int_0^t e^{\phi(y(z+\tau)) - r_{\max}(t-\tau) - \int_\tau^t \beta(y(z+s))\mathrm{d}s}\mathrm{d}\tau\right].$$

This leads to an explicit but complex formula for $C_t(z)$ thanks to the relation:

$$C_t(z) = D_t\left(\frac{1}{\mu}\mathrm{atanh}(\mu z)\right). \tag{4.17}$$

Additionally, we have:

$$\partial_z D_t(z) = \beta(y(z+t)) - \beta(y(z)) + \frac{\int_0^t \partial_z\, g(t,z,\tau)\,\mathrm{d}\tau}{\int_0^t g(t,z,\tau)\,\mathrm{d}\tau},$$

with $g(t,z,\tau) = \exp\left[\phi(y(z+\tau)) + r_{\max}(\tau - t) - \int_\tau^t \beta(y(z+s))\mathrm{d}s\right]$. Using the fact that $\overline{m}(t) = \partial_z D_t(0)$, $y(0) = 0$ and $\beta(0) = 0$, we get:

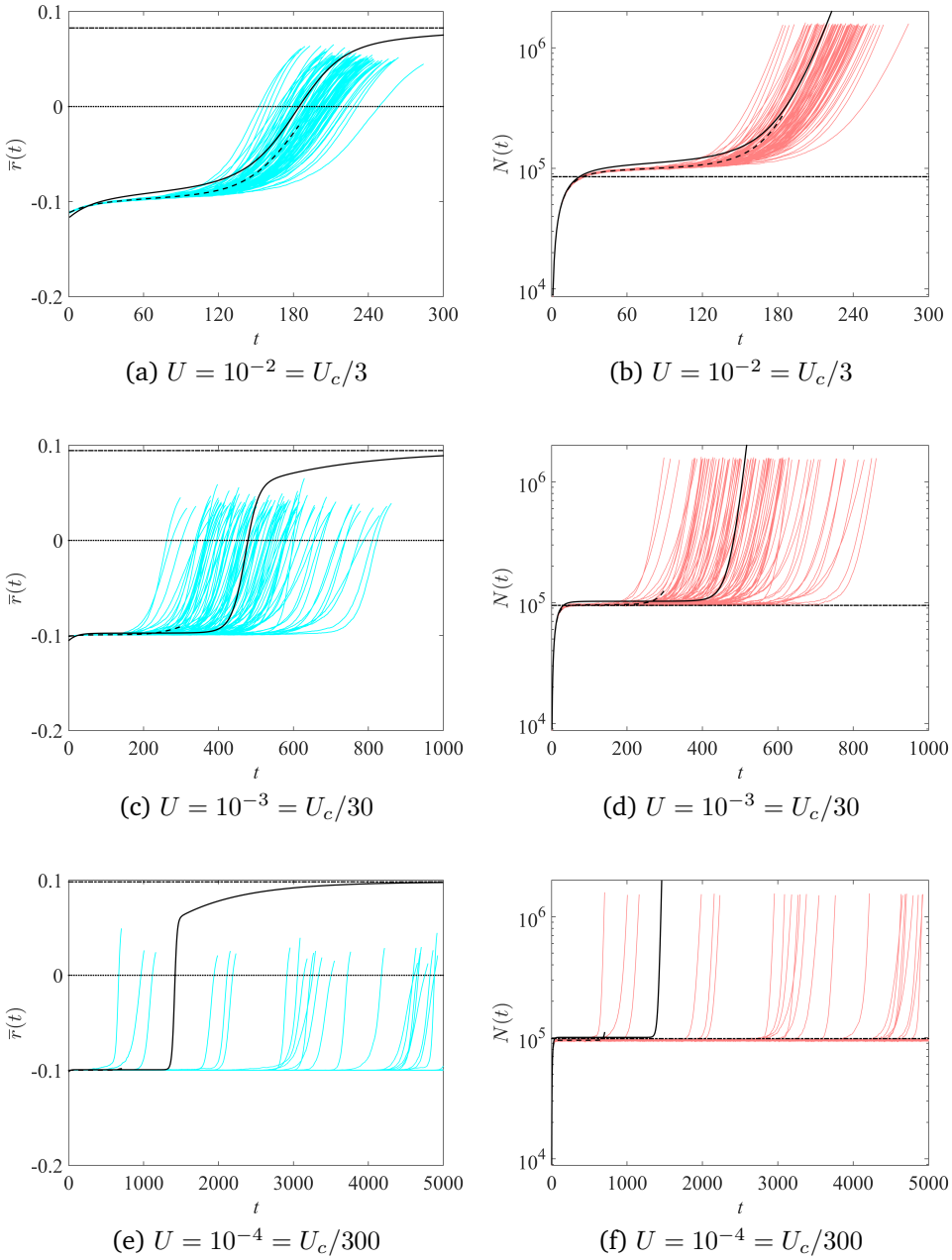
$$\begin{aligned}
\overline{m}(t) =& \beta(y(t)) + \frac{\int_0^t g(t,0,\tau)\,\left[y'(\tau)\phi'(y(\tau)) + \beta(y(\tau)) - \beta(y(t))\right]\,\mathrm{d}\tau}{\int_0^t g(t,0,\tau)\,\mathrm{d}\tau},\\
=& \frac{\int_0^t g(t,0,\tau)\,\left[y'(\tau)\phi'(y(\tau)) + \beta(y(\tau))\right]\,\mathrm{d}\tau}{\int_0^t g(t,0,\tau)\,\mathrm{d}\tau},\\
=& \frac{\int_0^t g(t,0,\tau)\,\left[y'(\tau)\phi'(y(\tau)) + \beta(y(\tau)) + r_{\max}\right]\,\mathrm{d}\tau}{\int_0^t g(t,0,\tau)\,\mathrm{d}\tau} - r_{\max},\\
=& \frac{\int_0^t g(t,0,\tau)\,\partial_\tau g(t,0,\tau)\,\mathrm{d}\tau}{\int_0^t g(t,0,\tau)\,\mathrm{d}\tau} - r_{\max},\\
=& \frac{g(t,0,t) - g(t,0,0)}{\int_0^t g(t,0,\tau)\,\mathrm{d}\tau} - r_{\max}.
\end{aligned}$$

Using the expression $g(t,0,\tau) = \exp\left[\phi(y(\tau)) + r_{\max}(\tau - t) - \int_\tau^t \beta(y(s))\mathrm{d}s\right]$, the formula (4.5) for $\phi$ and $y(z) = \tanh(\mu z)/\mu$, we finally get:

$$\overline{m}(t) = \frac{\exp\left[(r_{\max} - \mu\,\frac{n}{2})t + \frac{m_D}{2\mu}\,(e^{-2\mu t} - 1)\right] - 1}{\int_0^t \exp\left[(r_{\max} - \frac{n}{2}\mu)\tau + \frac{m_D}{2\mu}\,(e^{-2\mu\tau} - 1)\right]\mathrm{d}\tau} - r_{\max}. \tag{4.18}$$

As we have an explicit formula for $\overline{m}(t)$, we can also solve the ODE $N'(t) = N(t)\,(r_{\max} + \overline{m}(t)) + d$ (formula (4.16), with $\varepsilon = 0$ and $N(\varepsilon) = 0$). Finally, we can check that $N(t)$, $C_t(z)$ (defined by (4.17)) is a solution of (4.15) such that $\overline{m}$ (given by (4.18)) is continuous over $[0,\infty)$. Using the expression (4.18) with $\overline{r}(t) = r_{\max} + \overline{m}(t)$, we obtain the formula (4.10) in the main text.

### 4.5.4. Trajectories of mean fitness: $U < U_c$




(a) $U = 10^{-2} = U_c/3$ (b) $U = 10^{-2} = U_c/3$

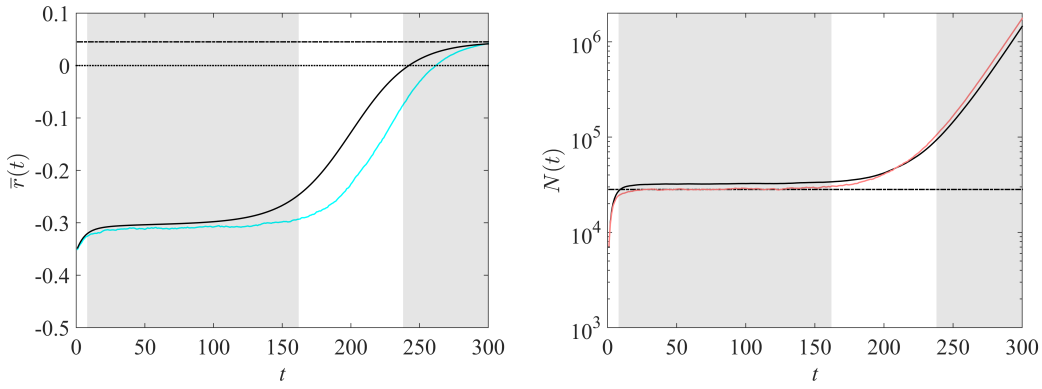
(c) $U = 10^{-3} = U_c/30$ (d) $U = 10^{-3} = U_c/30$

(e) $U = 10^{-4} = U_c/300$ (f) $U = 10^{-4} = U_c/300$

Figure 4.8. – **Trajectories of mean fitnesses and population sizes, low mutation rates.** Same legend as in Fig. 4.2. Other parameter values are $m_D = 0.2$, $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$, leading to $U_c = 0.03$.

### 4.5.5. Phenotype distribution in the sink: dynamics of $\overline{r}(t)$ and $N(t)$

The dynamics of mean fitness and population size corresponding to Fig. 4.3 are plotted in Fig. 4.9, to illustrate the occurrence of the four phases in this particular simulation.




(a) (b)

Figure 4.9. – **Trajectory of mean fitness and population size in the sink corresponding to the phenotype distribution in Fig. 4.3.** Same legend as in Fig. 4.2.

### 4.5.6. Independence of the evolutionary dynamics with respect to the immigration rate

The value of $\overline{r}(t)$ in formula (4.10) does not depend on $d$. Thus, only the population size dynamics are influenced by the immigration rate, but not the dynamics of adaptation. Actually, this phenomenon appears for a more general deterministic black-hole sink model, with a stable source and a constant immigration rate $d \geq 0$. In the sink, we have just to assume that the environment is initially empty ($N(0) = 0$), that both demography and evolution are density-independent (so that density dependence only arises in the migration effect). Apart from that, the proposed generalization may accommodate arbitrary forms of mutation and selection effects (possibly with changes in stress over time). The model then takes the following general form:

$$\begin{cases} \partial_t C_t(z) = \text{Selection}(t, z, C_t(z)) + \text{Mutation}(t, z, C_t(z)) + \frac{d}{N(t)} \left( e^{\phi(z) - C_t(z)} - 1 \right), \\ N'(t) = N(t)\, \overline{r}(t) + d, \\ C_t(0) = 0, \\ N(0) = 0, \end{cases}$$

with $\overline{r}(t) = \partial_z C_t(0)$ the coefficient of the exponential growth. Setting $P(t) = N(t)/d$, we observe that the above system can be written in the form:

$$\begin{cases} \partial_t C_t(z) = \text{Adaptation}(t, z, C_t(z)) + \text{Mutation}(t, z, C_t(z)) + \frac{1}{P(t)} \left( e^{\phi(z) - C_t(z)} - 1 \right), \\ P'(t) = P(t)\, \overline{r}(t) + 1, \\ C_t(0) = 0, \\ P(0) = 0, \end{cases}$$

with $\overline{r}(t) = \partial_z C_t(0)$. As this system does not depend on $d$, this implies that the dynamics of $P(t)$, of mean fitness $\overline{r}(t)$, and even of the full fitness distribution $(C_t(z))$ are all independent of $d$.

## 4.5.7. Large time behavior of $\overline{r}(t)$

We recall that, according to formula (4.10),

$$\overline{r}(t) = \frac{f(t) - 1}{\int_0^t f(\tau)\, \mathrm{d}\tau},$$

with $f(t) = \exp\left[ \left( r_{\max} - \mu \frac{n}{2} \right) t + \frac{m_D}{2\,\mu} (e^{-2\,\mu\, t} - 1) \right]$.

*We first show that $\overline{r}(t)$ is an increasing function of $t$.* First, we can check that:

$$f'(t) = f(t) \left( r_{\max} - \frac{\mu n}{2} - m_D e^{-2\mu t} \right).$$

Second, we have:

$$\begin{aligned} \overline{r}'(t) &= \frac{f'(t)}{\int_0^t f(\tau) \mathrm{d}\tau} - \frac{f(t) - 1}{\left( \int_0^t f(\tau) \mathrm{d}\tau \right)^2} f(t), \\ &= \frac{f(t)}{\left( \int_0^t f(\tau) \mathrm{d}\tau \right)^2} \left[ \left( r_{\max} - \frac{\mu n}{2} - m_D e^{-2\mu t} \right) \int_0^t f(\tau) \mathrm{d}\tau - (f(t) - 1) \right]. \end{aligned}$$

Let $h(t) = \left( r_{\max} - \frac{\mu n}{2} - m_D e^{-2\mu t} \right) \int_0^t f(\tau) \mathrm{d}\tau - (f(t) - 1)$. Thus we see that:

$$h'(t) = 2\,\mu\, m_D\, e^{-2\mu t} \int_0^t f(\tau) \mathrm{d}\tau \geq 0.$$

Therefore for all $t > 0$, $h(t) > h(0) = 0$, which shows that $\overline{r}$ is increasing.

Since $\overline{r}(0) = r_{\max} - \mu\, n/2 - m_D$, this implies that $\overline{r}(t) > r_{\max} - \mu\, n/2 - m_D$ for all $t > 0$. In particular, $\overline{r}(\infty) \geq r_{\max} - \mu\, n/2 - m_D$ which implies that $\delta(m_D) < m_D$ in (4.11).

*Next, we compute the limit of $\overline{r}(t)$ as $t \to \infty$.*

Case (i): we assume that $r_{\max} - \mu n/2 > 0$. Then, $f(t) \sim e^{-\frac{m_D}{2\mu}}\, e^{(r_{\max}-\mu n/2)t}$ and:

$$\int_0^t f(\tau)\,\mathrm{d}\tau \sim e^{-\frac{m_D}{2\mu}}\,\frac{e^{(r_{\max}-\mu n/2)t}}{r_{\max}-\mu n/2},\ \text{as } t\to\infty.$$

Thus,

$$\overline{r}(t) \to r_{\max} - \mu n/2 \text{ as } t\to\infty.$$

Case (ii): we assume that $r_{\max} - \mu n/2 = 0$. Then $f(t) = \exp\left[\frac{m_D}{2\mu}\,(e^{-2\mu t}-1)\right]$ and $\int_0^t f(\tau)\,\mathrm{d}\tau \sim t\, e^{-m_D/(2\mu)}$ as $t\to\infty$. Thus,

$$\overline{r}(t) \sim \frac{e^{-m_D/(2\mu)}-1}{e^{-m_D/(2\mu)}t} \to 0 \text{ as } t\to\infty.$$

Case (iii): we assume that $r_{\max} - \mu n/2 < 0$. Consider an arbitrary constant $\alpha\in(0,2)$. We can check that, for all $t < T_\alpha := \frac{1}{2\mu}\ln\frac{2}{\alpha}$, we have:

$$e^{-2\mu t} < 1 - \alpha\mu t.$$

In the sequel, we denote $X := r_{\max} - \mu n/2$. We get:

$$\begin{aligned}\int_0^\infty f(t)dt &= \int_0^{T_\alpha} f(t)dt + \int_{T_\alpha}^\infty f(t)dt,\\ &\le \int_0^{T_\alpha} \exp\left((X - m_D\alpha/2)t\right)dt,\\ &+ \int_{T_\alpha}^\infty \exp\left[Xt + \frac{m_D}{2\mu}\left(e^{-2\mu T_\alpha}-1\right)\right]dt.\end{aligned}$$

Using the assumption $X = r_{\max} - \mu n/2 < 0$, we obtain:

$$\int_0^\infty f(t)dt \le \frac{e^{(X-m_D\alpha/2)T_\alpha}-1}{X - m_D\,\alpha/2} - \exp\left[\frac{m_D}{2\mu}\left(e^{-2\mu T_\alpha}-1\right)\right]\ \frac{e^{XT_\alpha}}{X},$$

and using the definition of $T_\alpha = \frac{1}{2\mu}\ln\frac{2}{\alpha}$, we obtain:

$$\int_0^\infty f(t)dt \le -\left(\frac{\alpha}{2}\right)^{\frac{-X}{2\mu}}\left[\frac{\gamma}{X - \alpha\, m_D/2} + \frac{\rho}{X}\right],$$

with $\gamma := \left(\frac{\alpha}{2}\right)^{\frac{X}{2\mu}} - \left(\frac{\alpha}{2}\right)^{\alpha m_D/(4\mu)}$ and $\rho = \exp\left[\frac{m_D}{2\mu}\left(\frac{\alpha}{2}-1\right)\right]$. This leads to the

following inequality:

$$\overline{r}(\infty) = -\frac{1}{\int_0^\infty f(t)dt} \leq \left(\frac{\alpha}{2}\right)^{\frac{X}{2\mu}} \frac{X - \alpha m_D/2}{\gamma + \rho\left(1 - \frac{\alpha m_D}{2X}\right)},$$

which can be rewritten:

$$\overline{r}(\infty) \leq \frac{X - \alpha m_D/2}{1+\varepsilon},$$

with:

$$\varepsilon := \left(1 - \frac{\alpha\, m_D}{2\, X}\right) \rho \left(\frac{\alpha}{2}\right)^{-\frac{X}{2\,\mu}} - \left(\frac{\alpha}{2}\right)^{\frac{\alpha\, m_D}{4\,\mu} - \frac{X}{2\,\mu}}.$$

Next, to show that $\overline{r}(\infty) < X - \alpha m_D/2$, we only need to check that $\varepsilon < 0$. This is true for certain values of $\alpha$. As $\rho = \exp\left[\frac{m_D}{2\mu}\left(\frac{\alpha}{2} - 1\right)\right]$, we observe that $\varepsilon$ has the same sign as:

$$\varepsilon' = \left(1 - \frac{\alpha\, m_D}{2\, X}\right) \exp\left[\frac{m_D}{4\,\mu}(\alpha - 2)\right] - \exp\left[\frac{m_D}{4\,\mu}\alpha\, \ln(\alpha/2)\right].$$

Since $X = r_{\max} - \mu n/2$, we get:

$$\varepsilon' = \frac{m_D}{4\,\mu}\left[-\alpha\, \ln(\alpha/2) + \alpha\,(1 + 4/n) - 2\right] + o\left(\frac{1}{\mu}\right),$$

as $\mu \to \infty$. Thus, $\varepsilon < 0$ for $\mu$ large enough, if and only if:

$$n > \frac{4}{\ln(\alpha/2) - 1 + 2/\alpha}. \tag{4.19}$$

For $\alpha$ small enough, this inequality is true for any $n \geq 1$. However, higher values of $\alpha$ lead to sharper estimates of $\delta(m_D)$ in (4.11). With $\alpha = 1/4$ for instance, the inequality (4.19) is always satisfied (as $n \geq 1$). We obtain that $\overline{r}(\infty) \leq X - \frac{m_D}{8}$ and $\delta(m_D) \geq \frac{m_D}{8}$ for $\mu$ large enough. If $\alpha$ is increased, e.g., $\alpha = 1/2$, the inequality (4.19) is true for all $n \geq 3$, and consequently, $\overline{r}(\infty) \leq X - \frac{m_D}{4}$ for $\mu$ large enough ($\delta(m_D) \geq \frac{m_D}{4}$ , for $\mu$ large enough). In our numerical computations ($n = 6$), we can use $\alpha = 3/4$, which leads to $\overline{r}(\infty) \leq X - \frac{3\,m_D}{8}$ and $\delta(m_D) \geq \frac{3\,m_D}{8}$ for large $\mu$.

### 4.5.8. Establishment time $t_0$: formula (4.12)

We recall that $t_0$ is defined as the first zero of $\overline{r}(t)$. We note that, since $\overline{r}(t)$ is increasing, it admits at most one zero.

First, if $r_{\max} - \mu n/2 \leq 0$, as $\overline{r}(t)$ is increasing and $\overline{r}(\infty) < r_{\max} - \mu\, n/2$ (see (4.11) and Appendix 4.5.7), we have $\overline{r}(t) < 0$ for all $t \geq 0$. This implies that $t_0 = \infty$.

Second we assume that $r_{\max} - \mu n/2 > 0$. In this case, $\overline{r}(\infty) = r_{\max} - \mu n/2 > 0$

and the time $t_0$ is finite (and positive). Therefore, we can solve the equation $\overline{r}(t) = 0$, which is equivalent to:

$$(r_{\max} - \mu n/2)t + \frac{m_D}{2\mu}\left(e^{-2\mu t} - 1\right) = 0. \tag{4.20}$$

Let us set $c := m_D/(r_{\max} - \mu\, n/2)$. Since $\overline{r}(0) = r_{\max} - \mu\, n/2 - m_D < 0$, we observe that $c > 1$. The equation (4.20) is equivalent to:

$$2\mu\, t - c = -c\, e^{-2\,\mu\, t}.$$

Multiplying this expression by $e^{2\mu\, t - c}$, we get:

$$(2\mu\, t - c)e^{2\mu\, t - c} = -c\, e^{-c}.$$

Setting $X := 2\mu\, t - c$, we obtain:

$$X\, e^{X} = -c\, e^{-c}. \tag{4.21}$$

As $c > 1$, $-ce^{-c} \in (-e^{-1}, 0)$, thus the equation (4.21) admits two solutions, $X_0 = W_0(-c\, e^{-c})$ and $X_{-1} = W_{-1}(-c\, e^{-c}) < X_0$, with $W_0$ and $W_{-1}$ respectively the principal branch and the lower branch of the Lambert-W function. Thus, the equation (4.20) admits two solutions, $(c + X_0)/(2\mu)$ and $(c + X_{-1})/(2\mu) = 0$, but only the first one is positive. Finally, we obtain that:

$$t_0 = \frac{1}{2\mu}\left(c + W_0(-ce^{-c})\right). \tag{4.22}$$

As $t_0$ is an increasing function of $c$, we obtain that $t_0$ decreases as $r_{\max}$ is increased, and $t_0$ increases as $m_D$ and $n$ are increased. The dependence with respect to $\mu$ is more subtle. Differentiating the expression (4.22) with respect to $\mu$, we observe that $t_0'(\mu)$ has the same sign as:

$$\left(\frac{\mu\, n}{2\, r_{\max} - \mu\, n} - 1 - W_0(-ce^{-c})\right)\left(c + W_0(-ce^{-c})\right).$$

As the second factor in the above expression is always positive (since $c > 1$), we get that $t_0'(\mu)$ has the same sign as the function:

$$g(\mu) := \frac{\mu\, n}{2\, r_{\max} - \mu\, n} - 1 - W_0(-ce^{-c}).$$

Differentiating $g$ with respect to $\mu$, we observe that $g'(\mu)$ has the same sign as $r_{\max} + (\mu\, n/2 + m_D)\, W_0(-ce^{-c}) = r_{\max} - \mu\, n/2 + \mu\, n\,(1 - W_0(-ce^{-c}))/2 + m_D\, W_0(-ce^{-c})$. Thus $g'(\mu)$ has the same sign as $m_D\,(1/c + W_0(-ce^{-c})) + \mu\, n(1 - W_0(-ce^{-c}))/2 > 0$, as $1/c + W_0(-ce^{-c}) > 0$ (since $c > 1$) and $1 - W_0(-ce^{-c}) > 0$.

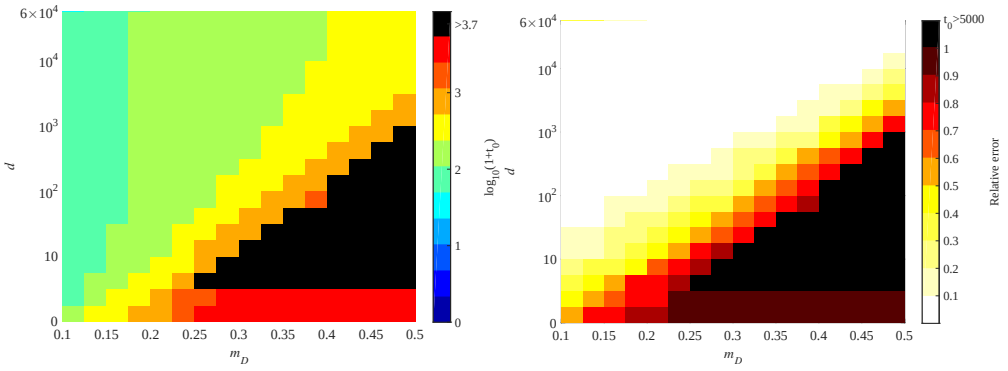
Finally, $g$ is increasing, with:

$$g(0) = -1 - W_0\left(-\frac{m_D}{r_{\max}} e^{-\frac{m_D}{r_{\max}}}\right) \leq 0,$$

(and the sign is strict unless $m_D = r_{\max}$). Additionally, we have $g(2r_{\max}/n) = +\infty$ (corresponding to $\mu_{lethal}$). This means that, unless $m_D = r_{\max}$, $t_0(\mu)$ first decreases until $\mu$ reaches an optimal value, and then increases as $\mu$ is increased.

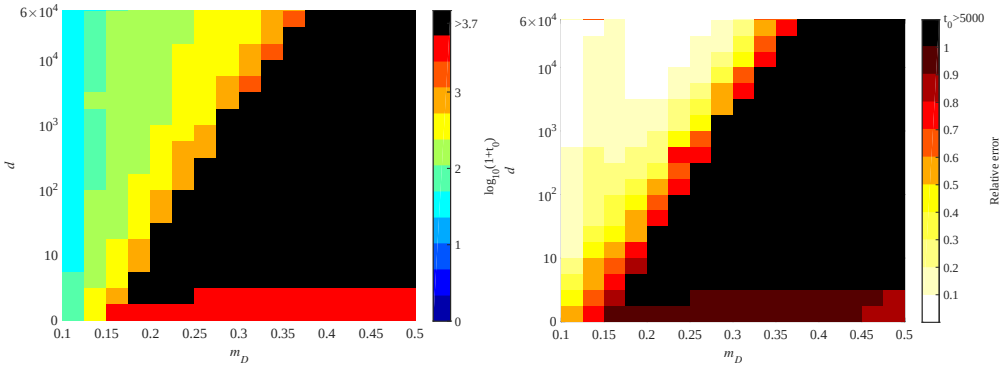
### 4.5.9. Establishment time $t_0$: dependence with the harshness of stress $m_D$ and the immigration rate $d$

Using the stochastic individual-based model of Section 4.2.5, we analysed the dependence of the establishment time $t_0$ with respect to $m_D$ and $d$ for a wide range of parameter values. Namely, taking $U = 0.1$, $r_{\max} = 0.1$, $\lambda = 1/300$ and $n = 6$ as in Fig. 4.6, $m_D$ was varied between $0.1$ and $0.5$. The results are presented in Fig. 4.10a. It shows that, for each value of $m_D$, there is a threshold value of the immigration rate above which the establishment time $t_0$ becomes almost independent of $d$. This threshold tends to increase as the harshness of stress $m_D$ takes higher values. Additionally, we measured the relative error between the theoretical value of $t_0$ given by formula (4.12) and the value given by individual-based simulations; see Fig. 4.10b. As soon as the parameters are far from the black region in Fig. 4.10, (a,b), the approximation is accurate (relative error $< 0.1$). This black region corresponds to values of $t_0 > 5000$, for which individual-based simulations were stopped before establishment, and where we can expect that the final outcome is establishment failure. This means that there is only a narrow region where formula (4.12) is not accurate; it is located close to the region where establishment fails, and describes a rapid increase in $t_0$ which is not captured by our analytical approach.




(a) $U = 0.1$ Establishment time

(b) $U = 0.1$ Relative error




(c) $U = 0.01$, establishment time

(d) $U = 0.01$, relative error

Figure 4.10. – **Establishment time $t_0$, dependence with the harshness of stress $m_D$ and the immigration rate $d$.** (a,c): Average value of $t_0$ over 100 individual-based simulations. The color legend corresponds to $\log_{10}(1 + t_0)$. (b,d): relative error between the theoretical value of $t_0$ given by formula (4.12) and the average value obtained by individual-based simulations. The black regions correspond to parameter values for which at least one simulation led to $t_0 > 5000$; in that case, the average value of $t_0$ was not computed numerically. In all cases, the parameter values are $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$.

Fig. 4.10 (c,d) depicts comparable simulations, with $U = U_c/3 = 0.01$, i.e., outside of the WSSM regime. The conclusions are similar to the case $U = 0.1$, but with a larger region corresponding to establishment failure, and a lower accuracy (panel d).

## 4.5.10. Dynamics in the absence of mutation in the sink

To get a better understanding of the four phases described in Section 4.3.1, we considered the case where the mutation rate $U = 0$ in the sink (while it remains positive in the source).

First, using the same arguments as in Appendix 4.5.3, we can derive a formula for $\overline{r}(t)$ in that case. The formula can be expressed in the same form as (4.10), with:

$$f(t) = \exp\left[\phi(t) + r_{\max}\, t\right],$$

with $\phi$ given by (4.5).

An example of trajectory of fitness is given in Fig. 4.11, where we observe that the four phases are still present. The corresponding phenotype distribution is presented in Fig. 4.12. A video file of the phenotype distribution is also available as Supplementary File 3.

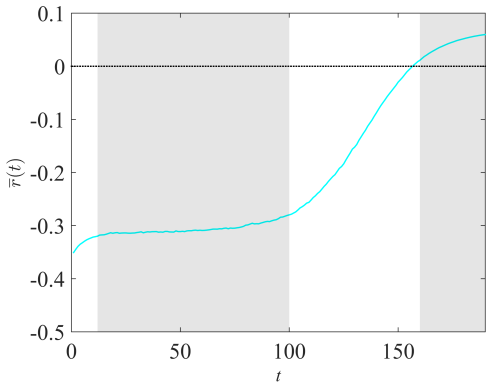
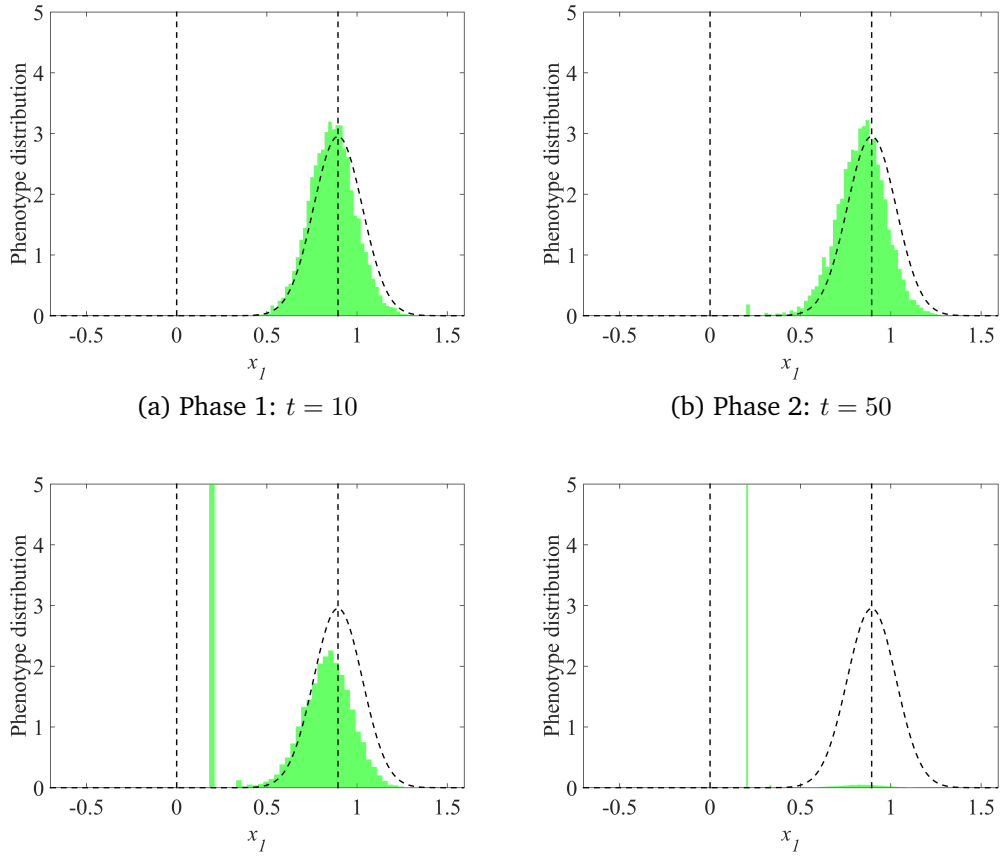



Figure 4.11. – **Dynamics of $\overline{r}(t)$ in the absence of mutation in the sink.** The blue curve corresponds to the trajectory of $\overline{r}(t)$ given by a single individual-based simulation, in the absence of mutation in the sink. The parameter values are $m_D = 0.4$, $U = 0.1$, $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$.




(a) Phase 1: $t = 10$

(b) Phase 2: $t = 50$

(c) Phase 3: $t = 115$

(d) Phase 4: $t = 190$

Figure 4.12. – **Phenotype distribution in the sink, along the direction $x_1$, in the absence of mutation.** The vertical dotted lines correspond to the sink ($x_1 = 0$) and source ($x_1 = \sqrt{2\,m_D}$) optima. The black dotted curve corresponds to the theoretical distribution of migrant's phenotypes in the sink (Gaussian distribution, centered at $x_1 = \sqrt{2\,m_D}$, and with variance $\mu = \sqrt{U\,\lambda}$). The parameter values are $m_D = 0.4$, $U = 0.1$, $r_{\max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$.

5

# Adaptation in a heterogeneous environment. Persistence versus extinction

**F. Hamel[a], F. Lavigne[a,b,c] and L. Roques[b]**

[a] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
[b] BioSP, INRA, 84914, Avignon, France
[c] ISEM (UMR 5554), CNRS, 34095, Montpellier, France

**Abstract**

Understanding how a diversity of plants in agroecosystems affects the adaptation of pathogens in a key issue in agroecology. We analyze PDE systems describing the dynamics of adaptation of two phenotypically structured populations, under the effects of mutation, selection and migration in a two-patches environment, each patch being associated with a different phenotypic optimum.

We consider two types of growth functions that depend on the $n-$dimensional phenotypic trait: either local and linear or nonlocal nonlinear. In both cases, we obtain existence and uniqueness results as well as a characterization of the large-time behaviour of the solution (persistence or extinction) based on the sign of a principal eigenvalue. We show that migration between the two environments decreases the chances of persistence, with in some cases a 'lethal migration threshold' above which persistence is not possible.

Comparison with stochastic individual-based simulations shows that the PDE approach accurately captures this threshold. Our results illustrate the importance of cultivar mixtures for disease prevention and control.

# Sommaire

## 5.1. Introduction

Phenotypic differences between populations generally appear as a consequence of differential selection regimes (Orr 1998). For instance, in the absence of migration, the adaptation of a population to local habitat conditions leads to a particular phenotypic distribution. In asexual populations, a standard way to describe the gene – environment interaction is to use Fisher's geometrical model (FGM) (Martin; Lenormand 2015; Tenaillon 2014). In this approach, each individual in the population is characterized by a multivariate phenotype at a set of $n$ traits, *i.e.*, a vector $\mathbf{x} \in \mathbb{R}^n$. This vector $\mathbf{x}$ determines the fitness $r(\mathbf{x})$ (the reproductive success of the individual) through its quadratic distance with respect to an optimum $\mathbf{O} \in \mathbb{R}^n$ associated with the considered environment:

$$r(\mathbf{x}) = r_{\max} - \frac{\|\mathbf{x} - \mathbf{O}\|^2}{2},$$

with $r_{\max} > 0$ the fitness of the optimal phenotype. Throughout the paper, $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^n$.

**PDE models.** Under the assumption of the FGM, recent models of asexual adaptation based on partial differential equations (PDE) (Alfaro; Carles 2017; Alfaro; Veruete 2018; Hamel; Lavigne; Martin; Roques 2020) typically describe the dynamics of the phenotype distribution $q$ of a population in a single environment, with equations of the form:

$$\forall t > 0,\ \forall \mathbf{x} \in \mathbb{R}^n, \quad \partial_t q(t,\mathbf{x}) = \frac{\mu^2}{2}\Delta q(t,\mathbf{x}) + [r(\mathbf{x}) - \overline{r}(t)]\, q(t,\mathbf{x}),$$

where the Laplace operator describes the mutation effects on the phenotype (see Hamel; Lavigne; Martin; Roques 2020, Appendix for the derivation of this term in this framework), and the term $[r(\mathbf{x}) - \overline{r}(t)]\, q(t,\mathbf{x})$ describes the effects of selection (Tsimring; Levine; Kessler 1996), with $\overline{r}(t) = \int_{\mathbb{R}^n} r(\mathbf{x})\, q(t,\mathbf{x})\, \mathrm{d}\mathbf{x}$ the mean fitness in the population at time $t$. Extensions to temporally changing environments (with an optimum $\mathbf{O}(t)$) have also been proposed (Roques; Patout; Bonnefon; Martin 2020). In all those cases, it was possible to describe the full dynamics of adaptation, by deriving explicit expressions for $\overline{r}(t)$.

Here, we consider a spatially heterogeneous environment, made of two habitats, each of them corresponding to a different phenotype optimum, $\mathbf{O}_1$ and $\mathbf{O}_2$. The main issue that we are going to deal with is to determine the respective effect of the migration between the two habitats and of the phenotypic distance between the two habitats on the faith (persistence or extinction) of the total population. This type of model has already been considered in Mirrahimi; Gandon 2020, in a particular regime of parameters such that the effect of the mutation is low, while the mutation rate is large enough, and in dimension $n = 1$. The authors have used

a specific method based on constrained Hamilton-Jacobi equations (*e.g.*, Barles; Mirrahimi; Perthame 2009; Diekmann; Jabin; Mischler; Perthame 2005; Gandon; Mirrahimi 2017b; Lorz; Mirrahimi; Perthame 2011; Perthame; Barles 2008 for more details on this method), to find an accurate analytic approximation of the equilibrium phenotype distribution and the population size in each habitat. They found that, when the two environments are symmetric (same mutation parameters, same selection pressure, same competition intensity and same migration rates), there exists an explicit threshold for the migration rate, which depends on the phenotypic distance between the two habitats. When the migration rate is above this threshold, the two subpopulations are well-mixed so that the total equilibrium population is monomorphic or 'generalist'. On the contrary, when the migration is below the threshold, the two subpopulations stay different, causing dimorphism in the phenotype density for the global population: the equilibrium population is made of two 'specialists'. They also obtained some results in the general case, without the symmetry assumption.

As we focus here on persistence/extinction issues, instead of dealing with the phenotype distribution $q(t,\mathbf{x})$, we are interested in the *phenotype density* $u(t,\mathbf{x})$, *i.e.*, $u(t,\mathbf{x}) = q(t,\mathbf{x})\,N(t)$, with $N(t)$ the population size at time $t$. We therefore deal with systems of the form:

$$\forall\, t \geq 0,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\ \begin{cases} \partial_t u_1(t,\mathbf{x}) &= \dfrac{\mu^2}{2}\,\Delta u_1(t,\mathbf{x}) + f_1(t,\mathbf{x},u_1) + \delta\,[u_2(t,\mathbf{x}) - u_1(t,\mathbf{x})], \\ \partial_t u_2(t,\mathbf{x}) &= \dfrac{\mu^2}{2}\,\Delta u_2(t,\mathbf{x}) + f_2(t,\mathbf{x},u_2) + \delta\,[u_1(t,\mathbf{x}) - u_2(t,\mathbf{x})], \end{cases} \tag{5.1}$$

with $u_i$ the phenotype density in the habitat $i \in \{1,2\}$, $\delta > 0$ the migration rate, and $\mu > 0$ a mutational parameter. Note that the migration and mutation parameters are assumed to be identical in the two habitats. This is a simplifying assumption which leads to symmetry properties of the solutions that are important for our analysis.

We may assume two different types of growth functions $f_i$. We first state that, in both cases, the fitness of a phenotype $\mathbf{x}$ in the habitat $i$ is given by:

$$r_i(\mathbf{x}) = r_{\max} - \frac{\|\mathbf{x} - \mathbf{O}_i\|^2}{2}, \tag{5.2}$$

Notice in particular that the fitnesses $r_i$ are unbounded in the phenotypic space $\mathbb{R}^n$ and, since they are involved in the definition of the growth functions $f_i$ for both types listed below, the system (5.1) of unknowns $(u_1,u_2)$ then has unbounded coefficients.

1. *The first type* (*Malthusian*):

$$f_i(t,\mathbf{x},u_i) = r_i(\mathbf{x})\,u_i(t,\mathbf{x}), \tag{5.3}$$

corresponds to the standard assumption of Malthusian population growth:

$$\forall\, t \geq 0, \quad \begin{cases} N_1'(t) = \overline{r}_1(t)\, N_1(t) + \delta\, [N_2(t) - N_1(t)], \\ N_2'(t) = \overline{r}_2(t)\, N_2(t) + \delta\, [N_1(t) - N_2(t)], \end{cases} \tag{5.4}$$

with $N_i(t)$ the population size in habitat $i$ at time $t$:

$$N_i(t) = \int_{\mathbb{R}^n} u_i(t, \mathbf{x})\, \mathrm{d}\mathbf{x}, \tag{5.5}$$

and $\overline{r}_i(t)$ the mean fitness of the individuals located in habitat $i$ at time $t$:

$$\overline{r}_i(t) = \frac{1}{N_i(t)} \int_{\mathbb{R}^n} r_i(\mathbf{x})\, u_i(t, \mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{5.6}$$

Note that, with $f_i$ of the type (5.3), the system (5.1) is a local cooperative system since the right-hand side of the equation of each component is nondecreasing with respect to the other component, and since the right-hand side only depends on the densities for the phenotype $\mathbf{x}$. As a consequence, the maximum principle holds for (5.1) in this first type, that is, if $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ are two classical solutions of (5.1) which are locally bounded in time and are such that $\mathbf{u}(0, \cdot) \leq \mathbf{v}(0, \cdot)$ in $\mathbb{R}^n$ (in the sense of componentwise inequalities), then $\mathbf{u}(t, \cdot) \leq \mathbf{v}(t, \cdot)$ in $\mathbb{R}^n$ for all $t > 0$.

2. *The second type* (*density-dependent*):

$$f_i(t, \mathbf{x}, u_i) = \left( r_i(\mathbf{x}) - \int_{\mathbb{R}^n} u_i(t, \mathbf{y})\, \mathrm{d}\mathbf{y} \right) u_i(t, \mathbf{x}), \tag{5.7}$$

corresponds to the standard assumption of logistic population growth:

$$\forall\, t \geq 0, \quad \begin{cases} N_1'(t) = \overline{r}_1(t)\, N_1(t) - N_1(t)^2 + \delta\, [N_2(t) - N_1(t)], \\ N_2'(t) = \overline{r}_2(t)\, N_2(t) - N_2(t)^2 + \delta\, [N_1(t) - N_2(t)], \end{cases} \tag{5.8}$$

with $N_i(t)$ and $\overline{r}_i(t)$ as in (5.5)-(5.6). Note that, with $f_i$ of the type (5.7), the system (5.1) is a nonlocal cooperative system, the nonlocality in the form of an internal competition. As a consequence, the maximum principle does not hold for (5.1) in this second type.

In this work, we will mainly focus our attention on the effects of the migration parameter $\delta$ and of the *habitat difference* defined by:

$$m_D := \frac{\|\mathbf{O}_1 - \mathbf{O}_2\|^2}{2} > 0. \tag{5.9}$$

**Stochastic model.** Before stating our main results, and in order to underline the interest of the PDE approach, we compare its accuracy with a standard Wright-

Fisher individual-based stochastic model (IBM), with mutation, selection and migration.

In this IBM, each individual is characterized by a phenotype $\mathbf{x} \in \mathbb{R}^n$, and a corresponding fitness $r_i(\mathbf{x})$, depending on the position (*i.e.*, the habitat $i = 1$ or $i = 2$) of the individual. The populations in the two habitats are initially clonal (with all of the phenotypes set at $(\mathbf{O}_1 + \mathbf{O}_2)/2$), and of size $N_i(0) = N^0$. Then, at each time step (the model is discrete in time), the *reproduction-selection step* is simulated by drawing a Poisson number of offspring, for each individual, with rate $\exp(r_i(\mathbf{x}))$ (Darwinian fitness, the discrete-time counterpart of $r_i(\mathbf{x})$). Then, the *mutation step* is simulated by randomly drawing, for each individual, a Poisson number of mutations, with rate $U > 0$. Each single mutation has a random phenotypic effect $\mathrm{d}\mathbf{x} \in \mathbb{R}^n$ drawn into a multivariate Gaussian distribution: $\mathrm{d}\mathbf{x} \sim \mathcal{N}(0, \lambda I_n)$, where $\lambda > 0$ is the mutational variance at each trait, and $I_n$ is the identity matrix of size $n \times n$. Multiple mutations in a single individual have additive effects on phenotype. Lastly, the *migration step* consists in sending individuals from the first habitat into the second (resp. from the second into the first): the numbers of migrants are drawn in a Poisson law with parameter $\delta N_1(t)$ (resp. $\delta N_2(t)$), and the migrants are randomly sampled in the populations.

**Numerical comparison between the PDE and stochastic models.** We simulated the IBM until a time $t = 300$, and compared the result with the numerical solution of the PDE model (5.1) with $\mu^2 = \lambda U$ (see Hamel; Lavigne; Martin; Roques 2020, Appendix for a justification of this parameter choice), and with the first type of growth function (Malthusian), as the IBM does not take density-dependence into account. The solution of the PDE was computed using the method of lines coupled with the Runge-Kutta ODE solver Matlab® *ode45*. The results are presented in Fig. 5.1. We observe a very good agreement between the results obtained with the IBM and the PDE, with in both cases a strong dependence of the persistence/extinction behaviour with respect to the parameters $\delta$ and $m_D$.

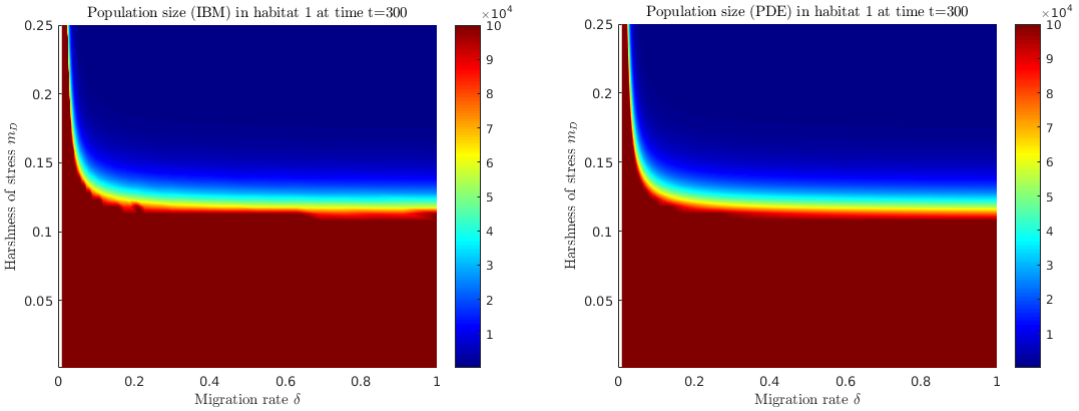



Figure 5.1. – **Persistence vs extinction: effect of the migration rate $\delta$ and the habitat difference $m_D$.** Total population size $N_1 + N_2$, given (a) by simulation of the stochastic model (average result over 50 replicate simulations) (b) by numerically solving (5.1) with $f_1, f_2$ given by (5.3). The parameters are $U = 1/6$, $\lambda = 1/300$, $n = 2$, $r_{\max} = 1/18$ and $\mu^2 = \lambda U$, and the results are computed at $t = 300$. Initially, each habitat $i \in \{1, 2\}$ has $N^0 = 10^4$ individuals, all of them with the phenotype $(\mathbf{O}_1 + \mathbf{O}_2)/2$.

**Aim of this paper.** Our main goal is to set on a firm mathematical basis the behaviour observed in Fig. 5.1, based on the sign of the principal eigenvalue of a system of linear elliptic equations, and to study the dependence of this eigenvalue with respect to the model parameters. The main results are presented in the next section, and discussed in Section 5.3.

## 5.2. Main results

Without loss of generality, we assume that the optima $\mathbf{O}_1$ and $\mathbf{O}_2$ are located along the $x_1$-axis and are symmetric with respect to the origin, *i.e.*, there exists $\beta > 0$ such that:

$$\mathbf{O}_1 = (-\beta, 0, \dots, 0), \quad \text{and} \quad \mathbf{O}_2 = (\beta, 0, \dots, 0). \tag{5.10}$$

We also assume that the two densities $u_1$ and $u_2$ are initially symmetric with respect to the hyperplane $\left\{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n,\ x_1 = 0\right\}$:

$$\forall\, \mathbf{x} \in \mathbb{R}^n,\ \ u_1^0(\mathbf{x}) = u_2^0(\iota(\mathbf{x})) =: u^0(\mathbf{x}),\ \mathbf{u}^0(\mathbf{x}) := (u_1^0(\mathbf{x}), u_2^0(\mathbf{x})) = (u^0(\mathbf{x}), u^0(\iota(\mathbf{x}))), \tag{SH}$$

with:

$$\forall\, \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n,\ \iota(\mathbf{x}) = (-x_1, x_2, \dots, x_n). \tag{5.11}$$

**The Cauchy problem.** We first show that the Cauchy problem associated with (5.1) admits a unique solution, under some assumptions on the initial condition $u^0$ given in (SH):

(H1) $u^0 \in C^{2,\alpha}(\mathbb{R}^n)$ for some $\alpha \in (0,1)$;

(H2) $N^0 := N(0) = \displaystyle\int_{\mathbb{R}^n} u^0(\mathbf{x})\,\mathrm{d}\mathbf{x} > 0$ and $N^0 < +\infty$;

(H3) there exists a nonincreasing function $g : \mathbb{R}_+ \to \mathbb{R}$ (with $\mathbb{R}_+ = [0,+\infty)$) such that:

(i) $0 \le u^0(\mathbf{x}) \le g(\|\mathbf{x} - \mathbf{O}_1\|)$ (and therefore $0 \le u^0(\iota(\mathbf{x})) \le g(\|\mathbf{x} - \mathbf{O}_2\|)$) for all $\mathbf{x} \in \mathbb{R}^n$;

(ii) the function $r \mapsto r^{n+1} g(r)$ belongs to $L^1(\mathbb{R}_+)$ and converges to $0$ as $r \to +\infty$.

Hereafter, unless otherwise specified, we always make the assumptions (SH) and (H1)-(H3).

Our first main result provides the existence and uniqueness of the density of phenotypes, for both types (5.3) and (5.7) of growth functions $f_1, f_2$.

**Theorem 32.**
*Assume that $f_1, f_2$ are either both of the first type (5.3) or both of the second type (5.7), and that $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$ satisfies (SH) and (H1)-(H3). Then, there exists a unique solution $\mathbf{u} = (u_1, u_2) \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ of (5.1), such that $\mathbf{u} \in L^\infty((0,T) \times \mathbb{R}^n)$ for all $T > 0$, $u_i(t,\mathbf{x}) \to 0$ as $\|\mathbf{x}\| \to +\infty$ locally uniformly in $t \in \mathbb{R}_+$, $u_i > 0$ in $(0,+\infty) \times \mathbb{R}^n$, the population sizes $N_i : \mathbb{R}_+ \to (0,+\infty)$ are of class $C^1$, the mean fitnesses $\overline{r}_i : \mathbb{R}_+ \to \mathbb{R}$ are continuous, $\mathbf{u}$ is symmetric in the sense that:*

$$u_1(t,\mathbf{x}) = u_2(t,\iota(\mathbf{x})), \quad \text{for all } t \ge 0 \text{ and } \mathbf{x} = (x_1,\dots,x_n) \in \mathbb{R}^n, \tag{5.12}$$

*with $\iota$ defined in (5.11), and:*

$$N_1(t) = N_2(t) =: N(t) \quad \text{and} \quad \overline{r}_1(t) = \overline{r}_2(t) =: \overline{r}(t), \quad \text{for all } t \ge 0. \tag{5.13}$$

*Moreover, the population sizes $N_1 = N_2$ satisfy (5.4) if $f_1, f_2$ are of the first type (5.3), whereas $N_1 = N_2$ satisfy (5.8) if $f_1, f_2$ are of the second type (5.7). In both cases, the functions $u_i$ satisfy the nonlocal parabolic equation:*

$$\partial_t u_i(t,\mathbf{x}) = \frac{\mu^2}{2}\,\Delta u_i(t,\mathbf{x}) + f_i(t,\mathbf{x},u_i) + \delta[u_i(t,\iota(\mathbf{x})) - u_i(t,\mathbf{x})], \tag{5.14}$$

*for all $t \ge 0$ and $\mathbf{x} \in \mathbb{R}^n$.*

**Remark 2.** *The existence and uniqueness result of Theorem 32 in the first type (5.3) can easily be extended to the non-symmetric case, i.e., without assumption (SH) (but with initial conditions $\mathbf{u}^0 = (u_1^0, u_2^0)$ such that both functions $u_1^0$ and $u_2^0$ still satisfy the assumptions (H1)-(H3)). In that case, the population sizes $N_1$ and $N_2$ still satisfy (5.4) but the $N_i$ and $\overline{r}_i$ do not satisfy (5.13) in general and the subsequent analysis then becomes more involved. In the second type (5.7), the existence and uniqueness is established only for symmetric solutions satisfying (5.12), since the proof, which is based on a change of functions amounting to a system with the first type (5.3), uses as a key ingredient the equality of the corresponding population sizes.*

**Persistence vs extinction.** Before going further on, we give a precise meaning to the notions of persistence and extinction. By extinction, we mean that the total population size $2\,N(t) = N_1(t) + N_2(t)$ converges to $0$ as $t \to +\infty$. By persistence, we mean that the population does not get extinct at large times. To analyze the effect of the parameter values on the persistence/extinction behaviour of the system (5.1), we consider an eigenvalue problem (see Cantrell; Cosner 2003 for several other examples of persistence/extinction results *via* eigenvalue problems in bounded domains).

For any $R > 0$, we denote by $\mathcal{A}$ the self-adjoint differential operator:

$$\mathcal{A} := -\frac{\mu^2}{2}\Delta - \begin{pmatrix} r_1(\mathbf{x}) - \delta & \delta \\ \delta & r_2(\mathbf{x}) - \delta \end{pmatrix}, \tag{5.15}$$

acting here on functions in $[W^{2,n}_{loc}(B(0,R)) \cap C_0(\overline{B(0,R)})]^2$, with $B(0,R)$ the open Euclidean ball of $\mathbb{R}^n$ of center $0$ and radius $R > 0$, and $C_0(\overline{B(0,R)})$ the space of continuous functions in $\overline{B(0,R)}$ which vanish on $\partial B(0,R)$. It follows from Sweers 1992, Theorem 1.1 that there exists a unique principal eigenvalue $\lambda^{\delta,R} \geq -r_{\max}$ and a unique (up to multiplication by a positive constant) pair of positive (in $B(0,R)$) eigenfunctions $(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) \in [W^{2,n}_{loc}(B(0,R)) \cap C_0(\overline{B(0,R)})]^2$, satisfying:

$$\mathcal{A}\,(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) = \lambda^{\delta,R}\,(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) \ \text{ in } B(0,R).$$

Moreover, the functions $\varphi_i^{\delta,R}$ are of class $C_0^\infty(\overline{B(0,R)}) = C^\infty(\overline{B(0,R)}) \cap C_0(\overline{B(0,R)})$ by standard elliptic estimates, and the eigenvalue $\lambda^{\delta,R}$ is characterized by the following minmax formula:

$$\lambda^{\delta,R} = \sup_{(\psi_1,\psi_2)\in E} \ \inf_{\mathbf{x}\in B(0,R),\, i\in\{1,2\}} \frac{(\mathcal{A}\,(\psi_1,\psi_2))_i(\mathbf{x})}{\psi_i(\mathbf{x})},$$

with:

$$E = \Big\{(\psi_1, \psi_2) \in [C^2(B(0,R)) \cap C(\overline{B(0,R)})]^2, \\ \psi_i(\mathbf{x}) > 0 \text{ for all } \mathbf{x} \in B(0,R) \text{ and } i \in \{1,2\}\Big\}.$$

This formula readily implies that the map $R \mapsto \lambda^{\delta,R}$ is nonincreasing. Since $\lambda^{\delta,R} \geq -r_{\max}$, the quantity $\lambda^{\delta,R}$ admits a finite limit as $R \to +\infty$:

$$\lambda^{\delta} := \lim_{R \to +\infty} \lambda^{\delta,R} \geq -r_{\max}. \tag{5.16}$$

The eigenfunctions $\varphi_1^{\delta,R}$ and $\varphi_2^{\delta,R}$ also satisfy a symmetry property.

**Lemma 5. (SYMMETRY PROPERTY OF THE EIGENFUNCTIONS)** *For every $R > 0$ and $\delta > 0$, the eigenfunction $(\varphi_1^{\delta,R}, \varphi_2^{\delta,R})$ satisfies $\varphi_1^{\delta,R}(\mathbf{x}) = \varphi_2^{\delta,R}(\iota(\mathbf{x}))$ for all $\mathbf{x} \in \overline{B(0,R)}$, with $\iota$ defined in (5.11).*

*Proof.* Set $\widetilde{\varphi}_1(\mathbf{x}) = \varphi_2^{\delta,R}(\iota(\mathbf{x}))$ and $\widetilde{\varphi}_2(\mathbf{x}) = \varphi_1^{\delta,R}(\iota(\mathbf{x}))$ for $\mathbf{x} \in \overline{B(0,R)}$. Then, from the symmetry assumption (5.10), one has $\mathcal{A}(\widetilde{\varphi}_1, \widetilde{\varphi}_2) = \lambda^{\delta,R}(\widetilde{\varphi}_1, \widetilde{\varphi}_2)$. By uniqueness (up to multiplication) of the pair of principal eigenfunctions, there exists $K > 0$ such that $(\widetilde{\varphi}_1, \widetilde{\varphi}_2) = K(\varphi_1^{\delta,R}, \varphi_2^{\delta,R})$. At $\mathbf{x} = 0$, we get that $\varphi_2^{\delta,R}(0) = \widetilde{\varphi}_1(0) = K\,\varphi_1^{\delta,R}(0)$ and $\varphi_1^{\delta,R}(0) = \widetilde{\varphi}_2(0) = K\,\varphi_2^{\delta,R}(0)$. Therefore $K = 1$ and the result is proved. □

Thus, we may (and we have to) take the same normalization condition for $\varphi_1^{\delta,R}$ and $\varphi_2^{\delta,R}$. In the proofs, we either assume that $\|\varphi_i^{\delta,R}\|_{L^1(B(0,R))} = 1$ for both $i \in \{1,2\}$, or $\varphi_i^{\delta,R}(0) = 1$ for both $i \in \{1,2\}$.

The large time behaviour of the population size is closely related to the sign of the quantity $\lambda^{\delta}$ defined in (5.16). We treat separately the first and second types (5.3) and (5.7).

**Theorem 33. (MALTHUSIAN GROWTH: BLOW UP VS EXTINCTION)**
*Assume that $f_1, f_2$ are of the first type (5.3), let $\delta > 0$ and $\lambda^\delta$ be given by (5.16). Let $\mathbf{u}$ be the solution of (5.1) given by Theorem 32, with initial condition $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$, and let $N(t) = N_1(t) = N_2(t)$ be its population size in each habitat.*

*(i) If $\lambda^\delta < 0$, then $N(t) \to +\infty$ as $t \to +\infty$ (blow up of the population).*

*(ii) If $\lambda^\delta = 0$ and if $u^0$ is compactly supported, then:*

$$\limsup_{t\to+\infty} N(t) < +\infty \quad \text{(boundedness of the population)}.$$

*Furthermore, there exist bounded positive stationary solutions of (5.1) with finite population sizes.*

*(iii) If $\lambda^\delta > 0$ and if $u^0$ is compactly supported, then $N(t) \to 0$ as $t \to +\infty$ (extinction of the population).*

**Theorem 34. (LOGISTIC GROWTH: PERSISTENCE VS EXTINCTION)**
*Assume that $f_1, f_2$ are of the second type (5.7), let $\delta > 0$ and $\lambda^\delta$ given by (5.16). Let $\mathbf{u}$ be the solution of (5.1) given by Theorem 32, with initial condition $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$, and let $N(t) = N_1(t) = N_2(t)$ be its population size in each habitat.*

*(i) If $\lambda^\delta < 0$, then:*

$$0 < \liminf_{t\to+\infty} N(t) \le \limsup_{t\to+\infty} N(t) < +\infty \quad \text{(persistence of the population)}, \tag{5.17}$$

*for some initial conditions $\mathbf{u}^0$.*

*(ii) If $\lambda^\delta \ge 0$ and if $u^0$ is compactly supported, then $N(t) \to 0$ as $t \to +\infty$ (extinction of the population).*

As a consequence of Theorems 33-34, the faith of the population is determined by the sign of $\lambda^\delta$, *i.e.*, by the linear stability of the steady state $\mathbf{u} = (0, 0)$, whether the growth functions $f_i$ be of the first or second type. The main differences between the Malthusian case and the logistic case arise when this steady state $(0, 0)$ is unstable ($\lambda^\delta < 0$). Although persistence occurs with both types of growth functions, the population size remains bounded with type 2 growth functions, due to the nonlocal competition term. We conjecture that it converges to $-\lambda^\delta$, as $t \to +\infty$. Interestingly, the threshold case $\lambda^\delta = 0$ leads to very different behaviours, depending on the type of growth functions: in the absence of competition (Malthusian growth), persistence is still possible in this case, although it is not in the logistic case. Biologically, however, the particular case

$\lambda^\delta = 0$ is presumably not relevant.

**Remark 3.** *In part (i) of Theorem 34 with $\lambda^\delta < 0$, the initial conditions $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$ such that (5.17) holds are those which are trapped between two multiples of the principal eigenfunctions associated to the operator $\mathcal{A}$ given by (5.15) but acting this time on $[W^{2,n}_{loc}(\mathbb{R}^n) \cap C_0(\mathbb{R}^n)]^2$, where $C_0(\mathbb{R}^n)$ is the space of continuous functions in $\mathbb{R}^n$ converging to $0$ at infinity. Such eigenfunctions are introduced in Lemma 6 below.*

In the following results, we now use the above persistence/extinction criteria to study the effect of the parameters, especially the habitat difference $m_D$ given by (5.9) and the migration rate $\delta$, on the faith of the population.

**Proposition 35.** *The map $\delta \mapsto \lambda^\delta$ is continuous, increasing and concave in $(0, +\infty)$. Moreover,*

$$\lim_{\delta \to 0^+} \lambda^\delta = -r_{\max} + \frac{\mu\, n}{2} =: \lambda^0, \quad \text{and} \quad \lim_{\delta \to +\infty} \lambda^\delta = \frac{m_D}{4} - r_{\max} + \frac{\mu\, n}{2} =: \lambda^\infty.$$

Two corollaries of Theorems 33-34 and Proposition 35 follow immediately with straightforward proof.

**Corollary 7.**
*Assume that $\lambda^0 < 0$, i.e., $r_{\max} > \mu\, n/2$. Let $\delta > 0$, let $\mathbf{u}$ be the solution of (5.1) given by Theorem 32, with initial condition $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$, and let $N(t) = N_1(t) = N_2(t)$ be its population size in each habitat.*

*(i) If $\lambda^\infty \leq 0$, i.e., if $m_D \leq 4\,(r_{\max} - \mu\, n/2)$, then $\lim_{t\to+\infty} N(t) = +\infty$ for the first type (5.3), whereas (5.17) is satisfied for some initial conditions $\mathbf{u}^0$ for the second type (5.7).*

*(ii) If $\lambda^\infty > 0$, i.e., if $m_D > 4\,(r_{\max} - \mu\, n/2)$, then there exists $\delta_{crit} > 0$, independent of $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$, such that:*

*(ii-a) if $\delta < \delta_{crit}$, then $\lim_{t\to+\infty} N(t) = +\infty$ for the first type (5.3), whereas (5.17) is satisfied for some initial conditions $\mathbf{u}^0$ for the second type (5.7);*

*(ii-b) if $\delta = \delta_{crit}$ and if $u^0$ is compactly supported, then $\limsup_{t\to+\infty} N(t) < +\infty$ for the first type (5.3) and $N(t) \to 0$ as $t \to +\infty$ for the second type (5.7);*

*(ii-c) if $\delta > \delta_{crit}$ and if $u^0$ is compactly supported, then $\lim_{t\to+\infty} N(t) = 0$ for both types (5.3) and (5.7).*

**Corollary 8.**
*Assume that $\lambda^0 \geq 0$, i.e., $r_{\max} \leq \mu\, n/2$. Let $\delta > 0$, let $\mathbf{u}$ be the solution of (5.1) given by Theorem 32, with initial condition $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$ and $u^0$ compactly supported, and let $N(t) = N_1(t) = N_2(t)$ be its population size in each habitat. Then $N(t) \to 0$ as $t \to +\infty$, for both types (5.3) and (5.7).*

An interpretation of Corollary 7 is that, when the maximal fitness $r_{\max}$ is large enough, namely $r_{\max} > \mu\, n/2$, and the environmental stress $m_D$ is low, namely $m_D \leq 4\,(r_{\max} - \mu\, n/2)$, the population can adapt to the global environment, whichever migration rate $\delta$. However, when the stress is high, namely $m_D > 4(r_{\max} - \mu\, n/2)$, the population can only survive if the migration rate is low ($\delta \leq \delta_{crit}$). These results are coherent with the numerical simulations of Figure 5.1. Corollary 8 says that, on the other hand, when the maximal fitness is small enough, namely $r_{\max} \leq \mu\, n/2$, the population, if initially compactly supported, can never adapt to the global environment, whichever migration rate $\delta$ and environmental stress $m_D$.

The last result is related to a unification result at each time $t > 0$ in the limit of infinite migration rates $\delta$.

**Theorem 36.**
*Let* $\mathbf{u}_\delta = (u_{\delta,1}, u_{\delta,2})$ *be the solution of* (5.1) *given by Theorem* 32 *and Remark* 2*, for growth functions* $f_1, f_2$ *of the first type* (5.3)*, with a fixed initial condition* $\mathbf{u}^0 = (u_1^0, u_2^0)$ *independent of* $\delta$ *and such that both functions* $u_1^0$ *and* $u_2^0$ *satisfy the assumptions (H1)-(H3). Then:*

$$\lim_{\delta \to +\infty} \|u_{\delta,1}(t,\cdot) - u_{\delta,2}(t,\cdot)\|_{L^\infty(\mathbb{R}^n)} = 0, \quad \textit{locally uniformly in } t \in (0, +\infty).$$

In other words, a strong migration rate $\delta$ unifies the two populations into one global population, since the exchanges between them are very large. The population then goes to be generalist at every time $t > 0$, even if it is not initially.

## 5.3. Discussion

*On the biological interpretation of the main results.* Proposition 35 together with Theorems 33-34 show that the more the two environments are connected by migration (*i.e.*, when $\delta$ is increased), the lower are the chances of persistence. In the absence of migration, when the two habitats are not connected ($\delta = 0$), it was already known that persistence occurs if $r_{\max} > \mu\, n/2$ Hamel; Lavigne; Martin; Roques 2020; Martin; Roques 2016, *i.e.*, $\lambda^0 < 0$, whereas $r_{\max} < \mu\, n/2$ leads to extinction (for both types of growth functions). In the case $\delta = 0$, at large times, the mean fitness $\overline{r}(t)$ converges to $r_{\max} - \mu\, n/2$. Thus, $-\mu\, n/2$ corresponds to the *mutation load*: the amount of maladaptation due to mutations. More precisely, if the mutation load exceeds the fitness of the optimal phenotype $r_{\max}$, the population is doomed to extinction. This corresponds to *lethal mutagenesis* Anciaux; Lambert; Ronce; Roques, et al. 2019.

When $\delta$ becomes positive, some individuals migrate between the two environments. Generally these individuals are better adapted to their environment of origin. Thus, as shown by Proposition 35, increasing the migration rate increases the global maladaptation. Ultimately, when $\delta \to +\infty$, the condition for persistence becomes $\lambda^\infty = -r_{\max} + m_D/4 + \mu\, n/2 < 0$: in this case, as shown by Theorem 36, the two phenotypic populations merge into a single one, centered at the origin, in-between the two optima. We observe that in addition to the mutation load $-\mu\, n/2$, a *migration load* $-m_D/4$ appears. It is proportional to the habitat difference $m_D$.

If $\lambda^0 < 0$ and $\lambda^\infty = -r_{\max} + m_D/4 + \mu\, n/2 > 0$, populations are doomed to extinction for large migration rates, but survive for small migration rates. Corollary 7 shows that there exists a migration threshold such that persistence is possible if the migration rate is below this threshold, but not if the migration rate is above this threshold. Thus, increasing the migration rate may imply a 'lethal migration effect', comparable to lethal mutagenesis.

*Implications in agroecology.* One of the fundamental principles in agroecology is to promote diversified agroecosystems rather than uniform cultures Caquet; Gascuel; Tixier-Boichard 2020; Food and Agriculture Organization of the United Nations 2018. Some empirical study already illustrated the higher resilience of such diversified agroecosystems Borg; Kiær; Lecarpentier; Goldringer, et al. 2018 to plant diseases. In our case, the two environments can be interpreted as two different types of host plants (different species, or different genetic variants) and the populations of phenotypes $u_1$, $u_2$ describe the density of a pathogen over these two types of host plants. With this interpretation, our study advocates for more diversified cultures, with strong migration of the pathogens between the host plants: it should reduce the chances of persistence of the pathogen over the agroecosystem. This is consistent, therefore, with the above-mentioned principle of plant diversification. However, we point out that this conclusion may not be valid for three environments or more: as discussed in Lavigne; Martin; Anciaux; Papaix, et al. 2020, the presence of a third environment associated with a phenotype optimum between the two others may lead to higher chances of persistence of the pathogen, compared to two environments, due to a 'springboard' effect. By now, and up to our knowledge, there is no rigorous mathematical proof of this result.

*On the derivation of quantitative estimates.* The methods used in our paper do not allow for a computation of the dispersal load: when $\delta = 0$, as discussed above, the mean growth rate $\overline{r}(t)$ converges to $r_{\max} - \mu\, n/2$. With positive values of $\delta$, it should converge to some value $r_{\max} - \mu\, n/2 + \text{Load}_{migr}(\delta)$, with $\text{Load}_{migr}(\delta) \in (-m_D/4, 0)$, the migration load. The determination of $\text{Load}_{migr}(\delta)$ would help disentangling the respective effect of mutation and migration on the persistence of a population. Additionally, Theorem 36 shows that when the migration rate is increased the two population merge into a single one, which may be qualified as 'generalist'. This is consistent with the results that have been obtained by Mirrahimi; Gandon 2020 in the case $n = 1$ with methods based on constrained Hamilton-Jacobi equations. This means that the mean phenotype in each environment converges to $\mathbf{x} = 0$. With smaller migration rates, the two populations should behave as 'specialists', with mean phenotypes that converges to $\mathbf{O}_1$ and $\mathbf{O}_2$ respectively as $\delta \to 0$. In a forthcoming work, using the methods in Hamel; Lavigne; Martin; Roques 2020 based on the analysis of moment generating functions associated with the distribution of fitness, we will aim to derive quantitative estimates for the migration load, the lethal migration threshold $\delta_{crit}$ and the respective distributions of phenotypes in the two environments.

## 5.4. Proofs

This section is devoted to the proofs of the results stated in Section 5.2. Section 5.4.1 is devoted to the proof of Theorem 32 on the well-posedness of the Cauchy problem (5.1). Section 5.4.2 is concerned with the proof of Theorems 33-34 on the large time behaviour of the population size, and Section 5.4.3 with the dependence of the faith of the population with respect to the parameters.

### 5.4.1. The Cauchy problem (5.1)

*Proof of Theorem 32.* We begin by assuming that $f_1, f_2$ are of the first type (5.3). As we will see later in the proof, the results in the case where $f_1, f_2$ are the second type (5.7) are then straightforward thanks to a change of functions. Note that the proof of the existence and uniqueness of the solution of the Cauchy problem (5.1) for the first type (5.3) actually does not require the symmetry property (SH) (but the proof still uses the same sign, smoothness and decay assumptions of each component $u_i^0$ of the initial condition $\mathbf{u}^0 = (u_1^0, u_2^0)$, and the positivity of the initial population size in each habitat).

So, let us first assume that $f_1, f_2$ are of the first type (5.3). Thanks to the assumptions (H1)-(H3) and owing to the definition (5.2) of the fitnesses $r_i$, it follows from Besala 1979, Theorem 3 that, for any $T > 0$, the Cauchy problem:

$$\left\{\begin{array}{rcl} \partial_t v_1(t,\mathbf{x}) & = & \dfrac{\mu^2}{2}\,\Delta v_1(t,\mathbf{x}) + [r_1(\mathbf{x}) - r_{\max}]\,v_1(t,\mathbf{x}) + \delta[v_2(t,\mathbf{x}) - v_1(t,\mathbf{x})], \quad t \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \\ \partial_t v_2(t,\mathbf{x}) & = & \dfrac{\mu^2}{2}\,\Delta v_2(t,\mathbf{x}) + [r_2(\mathbf{x}) - r_{\max}]\,v_2(t,\mathbf{x}) + \delta[v_1(t,\mathbf{x}) - v_2(t,\mathbf{x})], \quad t \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \\ \mathbf{v}(0,\mathbf{x}) & = & \mathbf{u}^0(\mathbf{x}) = (u_1^0(\mathbf{x}), u_2^0(\mathbf{x})),\ \mathbf{x} \in \mathbb{R}^n, \end{array}\right.$$

admits a solution $\mathbf{v} = (v_1, v_2) \in [C^{1,2}([0,T] \times \mathbb{R}^n) \cap L^\infty((0,T) \times \mathbb{R}^n)]^2$, such that $\mathbf{v}(t,\mathbf{x}) \to (0,0)$ as $\|\mathbf{x}\| \to +\infty$ uniformly in $t \in [0,T]$. Thus, the function $\mathbf{u} : (t,\mathbf{x}) \mapsto e^{r_{\max} t}\,\mathbf{v}(t,\mathbf{x})$, defined in $[0,T] \times \mathbb{R}^n$, is a bounded classical solution of (5.1) satisfying the same properties as $\mathbf{v}$. Moreover, this solution is nonnegative (componentwise) from the comparison principle (Weinberger 1975, Lemma 1) applied to this linear cooperative system. This maximum principle also yields the uniqueness of this solution $\mathbf{u}$. Since the initial population density in each habitat is not identically equal to $0$ by assumption (H2), the nonnegativity of each component $u_i$ and the strong parabolic maximum principle applied to each linear operator $\partial_t - (\mu^2/2)\Delta - r_i(\mathbf{x}) + \delta$ (for $i \in \{1,2\}$) yield the positivity of each component $u_i$ in $(0,T] \times \mathbb{R}^n$. As $T > 0$ can be chosen arbitrarily, these existence, uniqueness and positivity results extend to $t \in (0, +\infty)$, with local boundedness in $t$.

Still for the first type (5.3), in order to show that the population sizes and mean fitnesses $N_i(t)$ and $\overline{r}_i(t)$ defined by (5.5)-(5.6) are real valued, continuous and satisfy (5.4), we first establish some bounds and, to do so, we construct a

super-solution for $\mathbf{u} = (u_1, u_2)$. Let us set, for all $t > 0$ and $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{h}(t,\mathbf{x}) := \begin{pmatrix} h_1(t,\mathbf{x}) \\ h_2(t,\mathbf{x}) \end{pmatrix} := e^{(r_{\max}-\delta)t} \left[K_t * u_1^0\right](\mathbf{x}) \begin{pmatrix} \cosh(\delta t) \\ \sinh(\delta t) \end{pmatrix} + e^{(r_{\max}-\delta)t} \left[K_t * u_2^0\right](\mathbf{x}) \begin{pmatrix} \sinh(\delta t) \\ \cosh(\delta t) \end{pmatrix}, \tag{5.18}$$

with:

$$\forall\, t > 0,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\quad K_t(\mathbf{x}) = \frac{e^{-\|\mathbf{x}\|^2/(2\mu^2 t)}}{(2\pi\mu^2 t)^{n/2}},$$

and $\mathbf{h}(0,\mathbf{x}) = \mathbf{u}^0(\mathbf{x}) = \mathbf{u}(0,\mathbf{x})$. The function $\mathbf{h}$ is of class $[C^\infty((0,+\infty)\times\mathbb{R}^n) \cap C([0,+\infty)\times\mathbb{R}^n)]^2$, it is locally bounded in time, it converges to $(0,0)$ as $\|\mathbf{x}\| \to +\infty$ locally uniformly in $t \in \mathbb{R}_+$, and it satisfies:

$$\partial_t \mathbf{h}(t,\mathbf{x}) = \frac{\mu^2}{2}\Delta\mathbf{h}(t,\mathbf{x}) + \begin{pmatrix} r_{\max}-\delta & \delta \\ \delta & r_{\max}-\delta \end{pmatrix} \mathbf{h}(t,\mathbf{x}),$$

for all $t > 0$ and $\mathbf{x} \in \mathbb{R}^n$. Let $\boldsymbol{\psi}(t,\mathbf{x}) := \mathbf{u}(t,\mathbf{x}) - \mathbf{h}(t,\mathbf{x})$. We see that $\boldsymbol{\psi}(0,\mathbf{x}) = (0,0)$ for all $\mathbf{x} \in \mathbb{R}^n$, and:

$$\partial_t \boldsymbol{\psi}(t,\mathbf{x}) - \frac{\mu^2}{2}\Delta\boldsymbol{\psi}(t,\mathbf{x}) - \begin{pmatrix} r_{\max}-\delta & \delta \\ \delta & r_{\max}-\delta \end{pmatrix} \boldsymbol{\psi}(t,\mathbf{x}) = \begin{pmatrix} m_1(\mathbf{x})\, u_1(t,\mathbf{x}) \\ m_2(\mathbf{x})\, u_2(t,\mathbf{x}) \end{pmatrix} \le \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{5.19}$$

for all $t > 0$ and $\mathbf{x} \in \mathbb{R}^n$, with:

$$m_i(\mathbf{x}) := r_i(\mathbf{x}) - r_{\max} = -\frac{\|\mathbf{x}-\mathbf{O}_i\|^2}{2} \le 0. \tag{5.20}$$

Again, the comparison principle (Weinberger 1975, Lemma 1) implies that $\boldsymbol{\psi} \le 0$ (componentwise) in $\mathbb{R}_+ \times \mathbb{R}^n$, hence:

$$0 \le \mathbf{u}(t,\mathbf{x}) \le \mathbf{h}(t,\mathbf{x}) \ \text{ for all } (t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n. \tag{5.21}$$

The strong parabolic maximum principle actually implies that the second inequality, as is the first one, is strict in $(0,+\infty)\times\mathbb{R}^n$, as follows from (5.19) together with the positivity of $u_1$ and $u_2$. Moreover, for $i \in \{1,2\}$ and $t > 0$,

$$\int_{\mathbb{R}^n} \left[K_t * u_i^0\right](\mathbf{x})\, d\mathbf{x} = \int_{\mathbb{R}^n} u_i^0(\mathbf{x})\, d\mathbf{x} =: N_i^0 < +\infty.$$

Thus, $\int_{\mathbb{R}^n} h_i(t,\mathbf{x})\, d\mathbf{x} \le (N_1^0 + N_2^0)\, e^{r_{\max} t}$ for all $t > 0$, and, from (5.21) and the positivity of $u_i$ in $(0,+\infty)\times\mathbb{R}^n$, there holds:

$$0 < N_i(t) = \int_{\mathbb{R}^n} u_i(t,\mathbf{x})\, d\mathbf{x} \le (N_1^0 + N_2^0)\, e^{r_{\max} t}, \tag{5.22}$$

for all $t > 0$, as well as for $t = 0$ trivially.

Consider now any time $t \ge 0$ and let us prove that $\overline{r}_i(t)$ defined in (5.6) is finite, for $i \in \{1,2\}$. First, the hypotheses (H2)-(H3) imply that $\overline{r}_i(0)$ is finite. Assume

then that $t > 0$. From (5.20)-(5.21) and the positivity of $u_i$, we have:

$$r_{\max} N_i(t) \geq \int_{\mathbb{R}^n} r_i(\mathbf{x})\, u_i(t,\mathbf{x})\, \mathrm{d}\mathbf{x} \geq r_{\max} N_i(t) + \int_{\mathbb{R}^n} m_i(\mathbf{x})\, h_i(t,\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{5.23}$$

Thus, to show that $\overline{r}_i(t)$ is finite, we only have to show that the last term in the right-hand side of the above equation is finite. First, we note that:

$$0 \leq h_i(t,\cdot) \leq e^{r_{\max} t}\, K_t * (u_1^0 + u_2^0) \ \text{ in } \mathbb{R}^n. \tag{5.24}$$

Then, still using the assumption (H3), we have:

$$\begin{aligned}
0 \leq & \int_{\mathbb{R}^n} -m_i(\mathbf{x})\, [K_t * u_i^0](\mathbf{x})\, \mathrm{d}\mathbf{x}, \\
& \leq \frac{1}{2(2\pi\mu^2 t)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{x}-\mathbf{y}\|^2/(2\mu^2 t)} g(\|\mathbf{y}-\mathbf{O}_i\|)\, \mathrm{d}\mathbf{y}\, \mathrm{d}\mathbf{x}, \\
& = \frac{1}{2\pi^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{z}\|^2} g(\|\mathbf{x}-\mu\sqrt{2t}\,\mathbf{z}-\mathbf{O}_i\|)\, \mathrm{d}\mathbf{z}\, \mathrm{d}\mathbf{x}, \\
& \leq \frac{1}{2\pi^{n/2}} \int_{\mathbb{R}^n} \int_{\|\mathbf{z}\| \leq \|\mathbf{x}-\mathbf{O}_i\|/(2\mu\sqrt{2t})} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{z}\|^2} g\left(\frac{\|\mathbf{x}-\mathbf{O}_i\|}{2}\right) \mathrm{d}\mathbf{z}\, \mathrm{d}\mathbf{x} \\
& \quad + \frac{1}{2\pi^{n/2}} \int_{\mathbb{R}^n} \int_{\|\mathbf{z}\| > \|\mathbf{x}-\mathbf{O}_i\|/(2\mu\sqrt{2t})} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{z}\|^2} g(0)\, \mathrm{d}\mathbf{z}\, \mathrm{d}\mathbf{x}, \\
& \leq \frac{1}{2\pi^{n/2}} \left[ \pi^{n/2} \int_{\mathbb{R}^n} \|\mathbf{x}-\mathbf{O}_i\|^2\, g\left(\frac{\|\mathbf{x}-\mathbf{O}_i\|}{2}\right) \mathrm{d}\mathbf{x} + g(0) \int_{\mathbb{R}^n} \zeta_t(\|\mathbf{x}-\mathbf{O}_i\|) \|\mathbf{x}-\mathbf{O}_i\|^2\, \mathrm{d}\mathbf{x} \right],
\end{aligned}$$

where:

$$\zeta_t(r) := \int_{\|\mathbf{z}\| \geq r/(2\mu\sqrt{2t})} e^{-\|\mathbf{z}\|^2}\, \mathrm{d}\mathbf{z} = O\Big(e^{-r}\Big), \ \text{ as } r \to +\infty. \tag{5.25}$$

The assumption (H3) thus implies that:

$$0 \leq \int_{\mathbb{R}^n} -m_i(\mathbf{x}) [K_t * u_i^0](\mathbf{x})\, \mathrm{d}\mathbf{x} < +\infty, \tag{5.26}$$

for every $t > 0$. Let us now check that $-\int_{\mathbb{R}^n} m_i(\mathbf{x}) [K_t * u_j^0](\mathbf{x})\, \mathrm{d}\mathbf{x} < +\infty$ for

$i \neq j \in \{1,2\}$:

$$\begin{aligned}
&\int_{\mathbb{R}^n} -m_i(\mathbf{x})[K_t * u_j^0](\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&\qquad \le \frac{1}{2(2\pi\mu^2 t)^{n/2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{x}-\mathbf{y}\|^2/(2\mu^2 t)} g(\|\mathbf{y}-\mathbf{O}_j\|)\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}, \\
&\qquad = \frac{1}{2\pi^{n/2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{z}\|^2} g(\|\mathbf{x}-\mu\sqrt{2t}\,\mathbf{z}-\mathbf{O}_j\|)\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{x}, \\
&\qquad \le \frac{1}{2\pi^{n/2}} \int_{\mathbb{R}^n}\int_{\|\mathbf{z}\|\le\|\mathbf{x}-\mathbf{O}_j\|/(2\mu\sqrt{2t})} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{z}\|^2} g\left(\frac{\|\mathbf{x}-\mathbf{O}_j\|}{2}\right)\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{x} \\
&\qquad\quad + \frac{1}{2\pi^{n/2}} \int_{\mathbb{R}^n}\int_{\|\mathbf{z}\|>\|\mathbf{x}-\mathbf{O}_j\|/(2\mu\sqrt{2t})} \|\mathbf{x}-\mathbf{O}_i\|^2\, e^{-\|\mathbf{z}\|^2} g(0)\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{x}, \\
&\qquad \le \frac{1}{2\pi^{n/2}} \left[\pi^{n/2}\int_{\mathbb{R}^n} \left(2\|\mathbf{x}-\mathbf{O}_j\|^2+8\beta^2\right) g\left(\frac{\|\mathbf{x}-\mathbf{O}_j\|}{2}\right)\mathrm{d}\mathbf{x}\right. \\
&\qquad\qquad \left. + g(0)\int_{\mathbb{R}^n} \zeta_t(\|\mathbf{x}-\mathbf{O}_j\|)\left(2\|\mathbf{x}-\mathbf{O}_j\|^2+8\beta^2\right)\mathrm{d}\mathbf{x}\right],
\end{aligned}$$

where we recall that $\beta$ is defined in (5.10). Thus, (H3) implies that:

$$0 \le \int_{\mathbb{R}^n} -m_i(\mathbf{x})[K_t * u_j^0](\mathbf{x})\,\mathrm{d}\mathbf{x} < +\infty. \tag{5.27}$$

Adding (5.26) and (5.27), and using (5.24), we obtain that:

$$0 \le \int_{\mathbb{R}^n} -m_i(\mathbf{x})\, h_i(t,\mathbf{x})\,\mathrm{d}\mathbf{x} < +\infty,$$

and, together with (5.23), we infer that $-\infty < \overline{r}_i(t) \le r_{\max}$ for $i \in \{1,2\}$ and $t > 0$ (and also for $t = 0$ as already emphasized).

Finally, since the quantities $\zeta_t(r)$ given in (5.25) are nondecreasing with respect to $t > 0$, the same arguments as above together with Lebesgue's dominated convergence theorem yield the continuity of the maps $t \mapsto N_i(t)$, $t \mapsto \int_{\mathbb{R}^n} m_i(\mathbf{x})\, u_i(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$ and $t \mapsto \overline{r}_i(t)$, in $\mathbb{R}_+$ (up to $t = 0$), for $i \in \{1,2\}$. Now, for any $i \neq j \in \{1,2\}$, $0 < \epsilon < t$ and $R > 0$, integrating (5.1) over $(\epsilon, t) \times B(0,R)$ yields:

$$\begin{aligned}
\int_{B(0,R)} u_i(t,\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_{B(0,R)} u_i(\epsilon,\mathbf{x})\,\mathrm{d}\mathbf{x} \;=\; & \frac{\mu^2}{2}\int_\epsilon^t \int_{\partial B(0,R)} \nu\cdot\nabla u_i(s,\mathbf{x})\,\mathrm{d}\sigma(\mathbf{x})\,\mathrm{d}s \\
& + \int_\epsilon^t \int_{B(0,R)} r_i(\mathbf{x})\, u_i(s,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}s \\
& + \delta\int_\epsilon^t \int_{B(0,R)} \left(u_j(s,\mathbf{x}) - u_i(s,\mathbf{x})\right)\mathrm{d}\mathbf{x}\,\mathrm{d}s,
\end{aligned}$$

where $\nu$ and $\mathrm{d}\sigma(\mathbf{x})$ denote the outward normal and surface measure on $\partial B(0,R)$. From (5.1), (5.21) and (5.24), together with (H3) and standard parabolic estimates, it follows that $\|\mathbf{x}\|^{n+1} u_i(s,\mathbf{x}) \to 0$ and $\|\mathbf{x}\|^{n-1}\|\nabla u_i(s,\mathbf{x})\| \to 0$ as $\|\mathbf{x}\| \to +\infty$, uniformly for $s \in [\epsilon, t]$. Therefore, by passing to the limit $R \to +\infty$

in the above displayed equality, one gets that:

$$N_i(t) - N_i(\epsilon) = \int_\epsilon^t \overline{r}_i(s)\, N_i(s)\, \mathrm{d}s + \delta \int_\epsilon^t \Big(N_j(s) - N_i(s)\Big)\, \mathrm{d}s,$$

where we also used Lebesgue's dominated convergence theorem, formula (5.20) and the continuity of the map $s \mapsto \int_{\mathbb{R}^n} m_i(\mathbf{x})\, u_i(s,\mathbf{x})\, \mathrm{d}\mathbf{x}$ in $\mathbb{R}_+$. Using the continuity of $N_i$, $N_j$ and $\overline{r}_i$ in $\mathbb{R}_+$, the passage to the limit $\epsilon \to 0^+$ yields:

$$N_i(t) - N_i(0) = \int_0^t \overline{r}_i(s)\, N_i(s)\, \mathrm{d}s + \delta \int_0^t \Big(N_j(s) - N_i(s)\Big)\, \mathrm{d}s.$$

Hence, each function $N_i$ is of class $C^1(\mathbb{R}_+)$ and the pair $(N_1, N_2)$ satisfies (5.4).

We now show the symmetry property of the solutions of (5.1), still for the first type (5.3). With $(u_1, u_2)$ given as above and satisfying (SH), it follows that the pair of functions $(U_1, U_2)$ defined by:

$$\forall\, t \in \mathbb{R}_+,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\ \ (U_1(t,\mathbf{x}), U_2(t,\mathbf{x})) = (u_2(t,\iota(\mathbf{x})), u_1(t,\iota(\mathbf{x}))),$$

with $\iota$ as in (5.11), is a $C^{1,2}([0,+\infty)\times\mathbb{R}^n)^2$ solution of the Cauchy problem (5.1). Furthermore, each component $U_i$ is positive in $(0,+\infty)\times\mathbb{R}^n$, bounded in $(0,T)\times\mathbb{R}^n$ for every $T>0$, and converges to 0 as $\|\mathbf{x}\| \to +\infty$ locally uniformly in $t \in \mathbb{R}_+$. By uniqueness of such solutions and by (SH), one gets that $U_1(t,\mathbf{x}) = u_1(t,\mathbf{x})$ and $U_2(t,\mathbf{x}) = u_2(t,\mathbf{x})$ for all $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n$, and so $u_1(t,\mathbf{x}) = u_2(t,\iota(\mathbf{x}))$. The equation (5.14) then readily follows from this equality. Moreover the population sizes at time $t \geq 0$ satisfy:

$$N_1(t) = \int_{\mathbb{R}^n} u_1(t,\mathbf{x})\, \mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} u_2(t,\mathbf{x})\, \mathrm{d}\mathbf{x} = N_2(t),$$

and the mean fitnesses are also such that $\overline{r}_1(t) = \overline{r}_2(t)$ for all $t \geq 0$.

In order to complete the proof of Theorem 32, we now derive an equivalence between the problem (5.1) in the symmetric case with $f_1, f_2$ of the first type (5.3), and the problem (5.1) with $f_1, f_2$ of the second type (5.7), still in the symmetric case. Firstly, assume that $f_1, f_2$ are of the first type (5.3), and let $u_i$, $N_i$ and $\overline{r}_i$ be defined by the first part of the present proof, for $i \in \{1,2\}$. From (SH) and the previous paragraph, we know that $\overline{r}_1(t) = \overline{r}_2(t) =: \overline{r}(t)$ and $N_1(t) = N_2(t) =: N(t) > 0$, with $N'(t) = \overline{r}(t)\, N(t)$, for all $t \geq 0$. Let $\widetilde{N}(t)$ be the solution of the ODE:

$$\widetilde{N}'(t) = \overline{r}(t)\, \widetilde{N}(t) - \widetilde{N}(t)^2,$$

with $\widetilde{N}(0) = N(0) > 0$. Since $\overline{r}$ is continuous in $\mathbb{R}_+$, the function $\widetilde{N}$ is well defined, positive, and of class $C^1$ in $\mathbb{R}_+$. Define, for $i \in \{1,2\}$, the functions:

$$\forall\, t \in \mathbb{R}_+,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\ \ \widetilde{u}_i(t,\mathbf{x}) = \frac{\widetilde{N}(t)}{N(t)}\, u_i(t,\mathbf{x}),$$

where the functions $u_i$ are recalled to satisfy (5.1) with $f_1, f_2$ of the first type (5.3). The pair $(\widetilde{u}_1, \widetilde{u}_2)$ is of class $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)^2$, it is locally bounded in time, it converges to $(0,0)$ as $\|\mathbf{x}\| \to +\infty$ locally uniformly in $t \in \mathbb{R}_+$, and it has the same initial condition as the pair $(u_1, u_2)$. Moreover, for all $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$, we have:

$$\frac{\widetilde{N}(t)}{N(t)}\,\partial_t u_i(t,\mathbf{x}) = \frac{\mu^2}{2}\Delta\widetilde{u}_i(t,\mathbf{x}) + r_i(\mathbf{x})\widetilde{u}_i(t,\mathbf{x}) + \delta\left[\widetilde{u}_j(t,\mathbf{x}) - \widetilde{u}_i(t,\mathbf{x})\right], \text{ [1]}$$

and:

$$\begin{aligned}\partial_t\widetilde{u}_i(t,\mathbf{x}) &= \frac{\widetilde{N}(t)}{N(t)}\,\partial_t u_i(t,\mathbf{x}) + \left(\frac{\widetilde{N}'(t)\,N(t) - \widetilde{N}(t)\,N'(t)}{N^2(t)}\right) u_i(t,\mathbf{x}),\\ &= \frac{\widetilde{N}(t)}{N(t)}\,\partial_t u_i(t,\mathbf{x}) - \frac{\widetilde{N}(t)^2}{N(t)}\,u_i(t,\mathbf{x}) = \frac{\widetilde{N}(t)}{N(t)}\,\partial_t u_i(t,\mathbf{x}) - \widetilde{N}(t)\widetilde{u}_i(t,\mathbf{x}).\end{aligned}$$

The functions $\widetilde{u}_i$ thus satisfy (with $i,j \in \{1,2\}$ and $i \neq j$):

$$\partial_t\widetilde{u}_i(t,\mathbf{x}) = \frac{\mu^2}{2}\Delta\widetilde{u}_i(t,\mathbf{x}) + \left[r_i(\mathbf{x}) - \widetilde{N}(t)\right]\,\widetilde{u}_i(t,\mathbf{x}) + \delta[\widetilde{u}_j(t,\mathbf{x}) - \widetilde{u}_i(t,\mathbf{x})],$$

for all $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$, and, as:

$$\int_{\mathbb{R}^n} \widetilde{u}_i(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = \widetilde{N}(t),$$

for all $t \geq 0$ and $i \in \{1,2\}$, the functions $\widetilde{u}_i$ then solve (5.1), with $f_1, f_2$ of the second type (5.7). These solutions $\widetilde{u}_i$ are also symmetric, in the sense that $\widetilde{u}_1(t,\mathbf{x}) = \widetilde{u}_2(t,\iota(\mathbf{x}))$ for all $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$, and they are positive in $(0,+\infty)\times\mathbb{R}^n$. Notice finally that:

$$\widetilde{r}(t) := \frac{1}{\widetilde{N}(t)}\int_{\mathbb{R}^n} r_i(\mathbf{x})\,\widetilde{u}_i(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = \overline{r}(t),$$

for all $t \geq 0$ and $i \in \{1,2\}$.

Conversely, assume that $(\widetilde{u}_1, \widetilde{u}_2)$ is a symmetric $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)^2$ locally bounded in time solution of (5.1) and converging to $(0,0)$ as $\|\mathbf{x}\| \to +\infty$ locally uniformly in $t \in \mathbb{R}_+$, with $f_1, f_2$ of the second type (5.7) and with a continuous associated population size $\widetilde{N}(t)$ in each habitat, such that $\widetilde{N}(0) > 0$. Since the system satisfied by $(\widetilde{u}_1 1, \widetilde{u}_2)$ can also be viewed as a linear cooperative system (with additional diagonal term $-\widetilde{N}_i(t)\,\widetilde{u}_i(t,x)$), the weak and strong comparison principle applied with respect to the trivial solution $(0,0)$ imply that the functions $\widetilde{u}_i$ are then positive in $(0,+\infty)\times\mathbb{R}^n$. Therefore, the population size $\widetilde{N}(t)$ is positive and $f_i(t,\mathbf{x},\widetilde{u}_i) \leq r_i(\mathbf{x})\,\widetilde{u}_i(t,\mathbf{x})$ for all $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$. As a consequence, the

1. We use in the last term the fact that the proportionality factor between $\widetilde{u}_i$ and $u_i$ is the same for $i \in \{1,2\}$.

pair $(\widetilde{u}_1, \widetilde{u}_2)$ is then a subsolution of the cooperative system (5.1) with growth functions of the first type (5.3). Since the maximum principle holds for the latter system, one infers that the functions $\widetilde{u}_i$ satisfy similar bounds as (5.21) and (5.24) above for the solutions $u_i$ in the first type (5.3). By arguing as above, it follows that the mean fitness $t \mapsto \widetilde{r}(t) = \widetilde{N}(t)^{-1} \int_{\mathbb{R}^n} r_i(\mathbf{x})\, \widetilde{u}_i(t,\mathbf{x})\, \mathrm{d}\mathbf{x}$ is continuous in $\mathbb{R}_+$ and independent of $i \in \{1,2\}$, and that population size $\widetilde{N}$ is of class $C^1(\mathbb{R}_+)$ and satisfies (5.8) (due to the additional term $-\widetilde{N}(t)\, \widetilde{u}_i(t,\mathbf{x})$ in the right-hand side of the equation satisfied by $\widetilde{u}_i$). Finally, by inverting all the calculations of the previous paragraph and by defining $N(t)$ as the solution of $N'(t) = \widetilde{r}(t)\, N(t)$ with $N(0) = \widetilde{N}(0)$, one gets that the pair $(u_1, u_2)$ defined by:

$$\forall\, t \in \mathbb{R}_+,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\quad u_i(t,\mathbf{x}) = \frac{N(t)}{\widetilde{N}(t)}\, \widetilde{u}_i(t,\mathbf{x}),$$

is a symmetric solution of (5.1) satisfying the conditions of Theorem 32 with growth functions $f_1, f_2$ of the first type (5.3). The uniqueness result for the solutions in the first type (5.3) then leads to the uniqueness of the symmetric solutions of (5.1) for growth functions of the second type (5.7). The proof of Theorem 32 is thereby complete. □

### 5.4.2. Large time behavior

This section is devoted to the proof of Theorems 33-34. Before that, we state an auxiliary lemma on the existence of positive eigenfunctions of the operator $\mathcal{A}$ defined in (5.15).

**Lemma 6.** *There exists a pair of symmetric positive eigenfunctions $(\varphi_1^\delta, \varphi_2^\delta) \in [C_0^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)]^2$ such that $\mathcal{A}\,(\varphi_1^\delta, \varphi_2^\delta) = \lambda^\delta(\varphi_1^\delta, \varphi_2^\delta)$ in $\mathbb{R}^n$, with $\mathcal{A}$ defined by (5.15) and $\lambda^\delta$ by (5.16). Furthermore, this pair $(\varphi_1^\delta, \varphi_2^\delta)$ is unique up to multiplication by a positive constant.*

The proof of Lemma 6 is postponed after that of Theorem 33.

*Proof of Theorem 33.* Let $\mathbf{u} = (u_1, u_2)$ be the solution of (5.1) given by Theorem 32 with an initial condition $\mathbf{u}^0 = (u^0, u^0 \circ \iota)$ satisfying (SH) and (H1)-(H3), for $f_1, f_2$ of the first type (5.3). Recall that the symmetry of the problem implies that $N_1(t) = N_2(t) =: N(t)$ for all $t \geq 0$. For $R > 0$, let $(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) \in C_0^\infty(\overline{B(0,R)})^2$ and $\lambda^{\delta,R}$ be the principal eigenfunctions and eigenvalue of the operator $\mathcal{A}$ defined by (5.15). Finally, let $\lambda^\delta$ be given by (5.16). We consider the cases $\lambda^\delta < 0$ and $\lambda^\delta \geq 0$ separately, with $\lambda^\delta$ given by (5.15).

*First case: Assume that $\lambda^\delta < 0$.* From assumptions (H2)-(H3), we know that $u^0 \geq 0$ and $u^0 \not\equiv 0$ in $\mathbb{R}^n$ and, from Theorem 32, $u_i(1,\cdot) > 0$ in $\mathbb{R}^n$ for each $i \in$

$\{1,2\}$. As $\lim_{R\to+\infty} \lambda^{\delta,R} = \lambda^\delta < 0$, we can fix $R > 0$ such that $\lambda^{\delta,R} < 0$. Let $\underline{K} > 0$ be such that $\underline{K}\, e^{-\lambda^{\delta,R}}(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) \le \mathbf{u}(1,\cdot)$ in $\overline{B(0,R)}$. Set $\underline{H}(t,\mathbf{x}) = (\underline{H}_1, \underline{H}_2)(t,\mathbf{x}) := \underline{K}\, e^{-\lambda^{\delta,R}t}(\varphi_1^{\delta,R}(\mathbf{x}), \varphi_2^{\delta,R}(\mathbf{x}))$ for $t \ge 1$ and $\mathbf{x} \in \overline{B(0,R)}$. In particular, $\underline{H}(1,\cdot) \le \mathbf{u}(1,\cdot)$ in $\overline{B(0,R)}$. We have, for all $t \ge 1$, and $i \ne j \in \{1,2\}$,

$$\partial_t \underline{H}_i = \frac{\mu^2}{2}\Delta \underline{H}_i + r_i(\mathbf{x})\, \underline{H}_i + \delta\, (\underline{H}_j - \underline{H}_i) \text{ in } \overline{B(0,R)}$$
$$\text{and } (\underline{H}_1, \underline{H}_2)(t,\cdot) = (0,0) \text{ on } \partial B(0,R). \quad (5.28)$$

As the pair $(u_1(t,\cdot), u_2(t,\cdot))$ satisfies the same equation in $\overline{B(0,R)}$ and is positive in $\mathbb{R}^n$ for each $t \ge 1$ and therefore on $\partial B(0,R)$, the maximum principle applied to this cooperative system implies that $u_i(t,\mathbf{x}) \ge \underline{H}_i(t,\mathbf{x})$ for all $t \ge 1$, $\mathbf{x} \in \overline{B(0,R)}$ and $i \in \{1,2\}$. Integrating over $B(0,R)$ the above inequality and using the positivity of $u_i$, we get:

$$N(t) \ge \underline{K}\, e^{-\lambda^{\delta,R}t}\, \|\varphi_i^{\delta,R}\|_{L^1(B(0,R))}, \quad \text{for all } t \ge 1 \text{ and } i \in \{1,2\}.$$

Since $\lambda^{\delta,R} < 0$, this implies that $N(t) \to +\infty$ as $t \to +\infty$ and this shows part (i) of Theorem 33.

*Second case: Assume that* $\lambda^\delta \ge 0$. Assume also that the initial condition $u^0$ is compactly supported. Then, there is $\overline{K} > 0$ large enough, one has $\overline{K}\, (\varphi_1^\delta, \varphi_2^\delta) \ge \mathbf{u}(0,\cdot)$ in $\mathbb{R}^n$. Set $\overline{H}(t,\mathbf{x}) = (\overline{H}_1, \overline{H}_2)(t,\mathbf{x}) := \overline{K}\, e^{-\lambda^\delta t}(\varphi_1^\delta(\mathbf{x}), \varphi_2^\delta(\mathbf{x}))$ for $t \ge 0$ and $\mathbf{x} \in \mathbb{R}^n$. As for (5.28), the function $\overline{H}$ satisfies the same cooperative system (5.1) as $\mathbf{u}$ in $\mathbb{R}_+ \times \mathbb{R}^n$, but with a larger initial condition. The comparison principle thus implies that, for $i \in \{1,2\}$ and $t \ge 0$:

$$0 \le u_i(t,\mathbf{x}) \le \overline{H}_i(t,\mathbf{x}) = \overline{K}\, e^{-\lambda^\delta t}\varphi_i^\delta(\mathbf{x}), \quad \text{for all } \mathbf{x} \in \mathbb{R}^n. \quad (5.29)$$

As the functions $\varphi_i^\delta$ belong to $L^1(\mathbb{R}^n)$, integrating (5.29) over $\mathbb{R}^n$ yields $\lim_{t\to+\infty} N(t) = 0$ if $\lambda^\delta > 0$. If $\lambda^\delta = 0$, (5.29) implies that:

$$\limsup_{t\to+\infty} N(t) < +\infty.$$

Furthermore, in that case, for every $C > 0$, $C\,(\varphi_1^\delta, \varphi_2^\delta)$ is a pair of positive stationary solutions of (5.1). That shows parts (ii) and (iii) of Theorem 33 and the proof of Theorem 33 is thereby complete. □

*Proof of Lemma 6.* For $R > 0$, the functions $(\varphi_1^{\delta,R}, \varphi_2^{\delta,R}) \in C_0^\infty(\overline{B(0,R)})^2$ satisfy:

$$\frac{\mu^2}{2}\Delta\varphi_i^{\delta,R} + (\lambda^{\delta,R} - \delta + r_i)\, \varphi_i^{\delta,R} + \delta\, \varphi_j^{\delta,R} = 0 \text{ in } \overline{B(0,R)},$$

with $i \ne j \in \{1,2\}$. As the eigenvalues $\lambda^{\delta,R}$ are nonincreasing with respect to $R$ and not smaller than $-r_{\max}$, we have $-r_{\max} \le \lambda^{\delta,R} \le \lambda^{\delta,2}$ for all $R \ge 2$. For every

$R' \geq 1$, it then follows from the Harnack inequality in Sirakov 2009, Theorem 2 (applied here with $\Omega = B(0, 2R')$) that there is a positive constant $C(R')$ such that:

$$\max_{\mathbf{x}\in\overline{B(0,R')},\, i\in\{1,2\}} \varphi_i^{\delta,R}(\mathbf{x}) \leq C(R') \min_{\mathbf{x}\in\overline{B(0,R')},\, i\in\{1,2\}} \varphi_i^{\delta,R}(\mathbf{x}),$$

for all $R \geq 2R'$. Without loss of generality, we assume the normalization condition $\varphi_i^{\delta,R}(0) = 1$ (recall that $\varphi_1^{\delta,R}(0) = \varphi_2^{\delta,R}(0)$ by the symmetry property of Lemma 5). Thus, we get:

$$0 < \varphi_i^{\delta,R}(\mathbf{x}) \leq C(R'), \quad \text{for all } \mathbf{x} \in \overline{B(0,R')},\ i \in \{1,2\}, \text{ and } R \geq 2R' \geq 2.$$

Standard elliptic estimates then imply that, for every $\theta \in [0,1)$, and for every $R' \geq 1$, the functions $\varphi_i^{\delta,R}$ are bounded in $C^{2,\theta}(\overline{B(0,R')})$, independently of $R \in [2R', +\infty)$. Thus Sobolev's injections imply that, up to the extraction of a subsequence, $\varphi_i^{\delta,R} \to \varphi_i^{\delta}$ in $C^2_{loc}(\mathbb{R}^n)$ as $R \to +\infty$, where the functions $\varphi_i^{\delta}$ satisfy $\mathcal{A}\,(\varphi_1^{\delta}, \varphi_2^{\delta}) = \lambda^{\delta}(\varphi_1^{\delta}, \varphi_2^{\delta})$, are nonnegative and such that $\varphi_i^{\delta}(0) = 1$ for $i \in \{1,2\}$. From the (scalar) strong elliptic maximum principle, the functions $\varphi_i^{\delta}$ are positive in $\mathbb{R}^n$. Furthermore, they satisfy the same symmetry property as the functions $\varphi_i^{\delta,R}$, and, again from standard elliptic estimates, they are of class $C^{\infty}(\mathbb{R}^n)$.

To show that the eigenfunctions $\varphi_i^{\delta}$ are in $L^1(\mathbb{R}^n)$ and converge to $0$ at infinity, we use the fact that the potentials $r_i$ are confining. In particular, we fix $R_0' \geq 1$ large enough such that, all $R \geq 2R_0'$, there holds $\lambda^{\delta,R} + r_i(\mathbf{x}) < -\|\mathbf{x}\|^2/4$ for all $\mathbf{x} \in \overline{B(0,R)} \setminus B(0,R_0')$ and $i \in \{1,2\}$, hence:

$$-\frac{\mu^2}{2}\Delta(\varphi_1^{\delta,R} + \varphi_2^{\delta,R})(\mathbf{x}) + \frac{\|\mathbf{x}\|^2}{4}\,(\varphi_1^{\delta,R} + \varphi_2^{\delta,R})(\mathbf{x}) < 0 \ \text{ in } \overline{B(0,R)} \setminus B(0,R_0').$$

For any such $R$, since $\max_{\partial B(0,R_0')} \varphi_i^{\delta,R} \leq C(R_0')$ and $\varphi_1^{\delta,R} + \varphi_2^{\delta,R} = 0$ on $\partial B(0,R)$, the maximum principle implies that $\varphi_1^{\delta,R} + \varphi_2^{\delta,R} \leq w$ in $\overline{B(0,R)} \setminus B(0,R_0')$, where $w$ denotes the solution of the equation $-(\mu^2/2)\Delta w(\mathbf{x}) + (\|\mathbf{x}\|^2/4)w(x) = 0$ in $\overline{B(0,R)} \setminus B(0,R_0')$ with the boundary conditions $w = 0$ on $\partial B(0,R)$ and $w = 2\,C(R_0')$ on $\partial B(0,R_0')$. Consequently,

$$\varphi_1^{\delta,R}(x) + \varphi_2^{\delta,R}(x) \leq w(x) \leq 2\,C(R_0')\,e^{({R_0'}^2 - \|\mathbf{x}\|^2)/\sqrt{8\mu^2}}, \quad \text{for all } \mathbf{x} \in \overline{B(0,R)} \setminus B(0,R_0'),$$

and for all $R \geq 2R_0'$. Thus, the same inequality holds for the functions $\varphi_1^{\delta} + \varphi_2^{\delta}$ in $\mathbb{R}^n \setminus B(0,R_0')$. This implies in particular that the eigenfunctions $\varphi_i^{\delta}$ belong to $L^1(\mathbb{R}^n)$ and converge to $0$ at infinity.

Lastly, since for any $\lambda \in \mathbb{R}$ the weak maximum principle holds outside a large ball for the system $\mathcal{A}(\varphi_1, \varphi_2) = \lambda(\varphi_1, \varphi_2)$ in the class of $C^2_0(\mathbb{R}^n)$ functions (namely, there is $\rho > 0$ such that, if $\varphi_1, \varphi_2 \in C^2_0(\mathbb{R}^n)$ satisfy $\mathcal{A}(\varphi_1, \varphi_2) \leq \lambda(\varphi_1, \varphi_2)$ in $\mathbb{R}^n \setminus B(0,\rho)$ and $(\varphi_1, \varphi_2) \leq (0,0)$ on $\partial B(0,\rho)$, then $(\varphi_1, \varphi_2) \leq (0,0)$ in $\mathbb{R}^n \setminus B(0,\rho)$) and since the strong maximum principle holds as well in any connected open subset $\Omega \subset \mathbb{R}^n$ (namely, if $\varphi_1, \varphi_2 \in C^2(\Omega)$ are such that $\mathcal{A}(\varphi_1, \varphi_2) \leq \lambda(\varphi_1, \varphi_2)$

and $(\varphi_1, \varphi_2) \leq (0,0)$ in $\Omega$ with $\varphi_i(x_0) = 0$ for some $i \in \{1,2\}$ and $x_0 \in \Omega$, then $(\varphi_1, \varphi_2) \equiv (0,0)$ in $\Omega$), it follows with similar arguments as in Berestycki; Nirenberg; Varadhan 1994 that the pair of eigenfunctions $(\varphi_1^\delta, \varphi_2^\delta)$ constructed above is unique, up to multiplication by a positive constant, in the class of $C_0^2(\mathbb{R}^n)$ eigenfunctions. Moreover, the eigenvalue $\lambda^\delta$ is the unique eigenvalue associated with a pair of positive eigenfunctions. The proof of Lemma 6 is thereby complete. □

*Proof of Theorem 34.* Let $\mathbf{u} = (u_1, u_2)$ be the unique symmetric solution of (5.1) given by Theorem 32, for $f_1, f_2$ of the second type (5.7). Let $N(t) := N_1(t) = N_2(t)$ be its population size given by (5.5) and $\overline{r}(t) := \overline{r}_1(t) = \overline{r}_2(t)$ be its mean fitness given by (5.6), at each time $t \geq 0$. From Theorem 32, the densities $u_i$ are positive in $(0, +\infty) \times \mathbb{R}^n$, the function $\overline{r}$ is continuous in $\mathbb{R}_+$, the function $N$ is positive and of class $C^1$ in $\mathbb{R}_+$, and $N'(t) = \overline{r}(t)N(t) - N(t)^2$ for all $t \in \mathbb{R}_+$.

Let also $\widetilde{\mathbf{u}} = (\widetilde{u}_1, \widetilde{u}_2)$ be the unique symmetric solution of (5.1) given by Theorem 32, for $f_1, f_2$ of the first type (5.3), with the same initial condition $\mathbf{u}^0$ as $\mathbf{u}$. Let $\widetilde{N}(t) := \widetilde{N}_1(t) = \widetilde{N}_2(t)$ be its population size and $\widetilde{r}(t) := \widetilde{r}_1(t) = \widetilde{r}_2(t)$ be its mean fitness, at each time $t \geq 0$. From Theorem 32, the densities $\widetilde{u}_i$ are positive in $(0, +\infty) \times \mathbb{R}^n$, the function $\widetilde{r}$ is continuous in $\mathbb{R}_+$, the function $\widetilde{N}$ is positive and of class $C^1$ in $\mathbb{R}_+$, and $\widetilde{N}'(t) = \widetilde{r}(t)\widetilde{N}(t)$ for all $t \in \mathbb{R}_+$.

The correspondence between the symmetric solutions of (5.1) for both types (5.3) and (5.7), shown in the last part of the proof of Theorem 32, implies that:

$$\widetilde{\mathbf{u}}(t, \mathbf{x}) = \frac{\widetilde{N}(t)}{N(t)}\, \mathbf{u}(t, \mathbf{x}), \quad \text{for all } t \geq 0 \text{ and } \mathbf{x} \in \mathbb{R}^n,$$

hence $\widetilde{r}(t) = \overline{r}(t)$ for all $t \geq 0$. Therefore, we have:

$$\frac{\widetilde{N}'(t)}{\widetilde{N}(t)} = \frac{N'(t)}{N(t)} + N(t),$$

for all $t \geq 0$. Integrating this equality and using $\widetilde{N}(0) = N(0)$ yields:

$$N(t) = \frac{\widetilde{N}(t)}{1 + \displaystyle\int_0^t \widetilde{N}(s)\, \mathrm{d}s}, \quad \text{for all } t \geq 0. \tag{5.30}$$

Let now $(\varphi_1^\delta, \varphi_2^\delta)$ be defined by Lemma 6 with the normalization $\|\varphi_i^\delta\|_{L^1(\mathbb{R}^n)} = 1$. Set:

$$H(t, \mathbf{x}) = (H_1(t, \mathbf{x}), H_2(t, \mathbf{x})) := e^{-\lambda^\delta t}(\varphi_1^\delta(\mathbf{x}), \varphi_2^\delta(\mathbf{x})),$$

for $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$. As in the proof of Theorem 33, the function $H$ satisfies (5.1) with growth functions $f_i$ of the first type (5.3). We then treat separately the cases $\lambda^\delta \geq 0$ and $\lambda^\delta < 0$.

*First case: Assume that* $\lambda^\delta \geq 0$. Assume also in this case that $u^0$ is compactly

supported. Then there is $K > 0$ such that $\tilde{\mathbf{u}}(0,\cdot) = \mathbf{u}(0,\cdot) = \mathbf{u}^0 \leq K\, H(0,\cdot)$ in $\mathbb{R}^n$ and the maximum principle applied to the cooperative system (5.1) with growth functions of the first type (5.3) implies that $\tilde{\mathbf{u}}(t,\cdot) \leq K\, H(t,\cdot)$ in $\mathbb{R}^n$ for all $t \geq 0$, hence $\widetilde{N}(t) \leq K\, e^{-\lambda^\delta t}$ for all $t \geq 0$. From (5.30) and the positivity of $N$ and $\widetilde{N}$, one immediately infers that $N(t) \to 0$ as $t \to +\infty$ if $\lambda^\delta > 0$.

Consider now the sub-case $\lambda^\delta = 0$. The previous observations imply that $\widetilde{N}$ is bounded in $\mathbb{R}_+$. Furthermore, on the one hand, if the integral $\int_0^{+\infty} \widetilde{N}(s)\,\mathrm{d}s$ diverges, then formula (5.30) and the boundedness of $\widetilde{N}$ imply that $N(t) \to 0$ as $t \to +\infty$. On the other hand, if the integral $\int_0^{+\infty} \widetilde{N}(s)\,\mathrm{d}s$ converges, then the boundedness of the function $\widetilde{N}' = \widetilde{r}\,\widetilde{N}$ in $\mathbb{R}_+$ (which itself follows from the inequalities $0 \leq \tilde{\mathbf{u}}(t,\cdot) \leq K\, H(t,\cdot) = K\,(\varphi_1^\delta, \varphi_2^\delta)$ in $\mathbb{R}^n$ and the exponential decay at infinity of the eigenfunctions $\varphi_i^\delta$ given the proof of Lemma 6) implies that $\widetilde{N}(t) \to 0$ as $t \to +\infty$, and finally $N(t) \to 0$ as $t \to +\infty$ by (5.30).

*Second case: Assume that* $\lambda^\delta < 0$. Assume also in this case that $\mathbf{u}^0$ is trapped between two positive multiples of the eigenfunctions $(\varphi_1^\delta, \varphi_2^\delta)$, namely, there exist $0 < K_1 \leq K_2$ such that:

$$K_1\,(\varphi_1^\delta, \varphi_2^\delta) \leq \mathbf{u}^0 \leq K_2\,(\varphi_1^\delta, \varphi_2^\delta) \ \text{ in } \mathbb{R}^n.$$

Thus, $K_1\, H(0,\cdot) \leq \tilde{\mathbf{u}}(0,\cdot) = \mathbf{u}^0 \leq K_2\, H(0,\cdot)$ in $\mathbb{R}^n$ and the maximum principle applied to the cooperative system (5.1) with growth functions of the first type (5.3) implies that:

$$K_1\, H(t,\cdot) \leq \tilde{\mathbf{u}}(t,\cdot) \leq K_2\, H(t,\cdot) \ \text{ in } \mathbb{R}^n, \text{ for all } t \geq 0.$$

In particular, $K_1\, e^{-\lambda^\delta t} \leq \widetilde{N}(t) \leq K_2\, e^{-\lambda^\delta t}$ for all $t \geq 0$. Together with (5.30) and the negativity of $\lambda^\delta$, one concludes that:

$$0 < \frac{K_1}{K_2}\,|\lambda^\delta| \leq \liminf_{t\to+\infty} N(t) \leq \limsup_{t\to+\infty} N(t) \leq \frac{K_2}{K_1}\,|\lambda^\delta| < +\infty.$$

The proof of Theorem 34 is thereby complete. □

### 5.4.3. Dependence with respect to the parameters

*Proof of Proposition* 35. We start with the concavity and the monotonicity in $(0,+\infty)$ of the map $\delta \mapsto \lambda^\delta$ defined in (5.16). Using the confining properties of the fitnesses $r_i(\mathbf{x})$, it follows from Lemma 6 and elementary arguments that, for any $\delta > 0$,

$$\lambda^\delta = \min_{\substack{\varphi \in H^1(\mathbb{R}^n)\setminus\{0\} \\ \mathbf{x}\mapsto \|\mathbf{x}\|\,\varphi(\mathbf{x}) \in L^2(\mathbb{R}^n)}} \mathcal{R}(\delta,\varphi), \tag{5.31}$$

with,

$$\mathcal{R}(\delta,\varphi)=\frac{\dfrac{\mu^2}{2}\displaystyle\int_{\mathbb{R}^n}\|\nabla\varphi(\mathbf{x})\|^2\,\mathrm{d}\mathbf{x}-\int_{\mathbb{R}^n}r_1(\mathbf{x})\,\varphi(\mathbf{x})^2\,\mathrm{d}\mathbf{x}+\delta\displaystyle\int_{\mathbb{R}^n}\left(\varphi(\mathbf{x})^2-\varphi(\mathbf{x})\,\varphi(\iota(\mathbf{x}))\right)\mathrm{d}\mathbf{x}}{\displaystyle\int_{\mathbb{R}^n}\varphi(\mathbf{x})^2\,\mathrm{d}\mathbf{x}},$$

and the minimum of the Rayleigh quotient $\mathcal{R}(\delta,\cdot)$ in (5.31) is reached only by multiples of the function $\varphi_1^\delta$ given in Lemma 6. For each $\varphi\in H^1(\mathbb{R}^n)\setminus\{0\}$ such that $\mathbf{x}\mapsto\|\mathbf{x}\|\,\varphi(\mathbf{x})\in L^2(\mathbb{R}^n)$, the map $\delta\mapsto\mathcal{R}(\delta,\varphi)$ is affine in $(0,+\infty)$ and nondecreasing in $(0,+\infty)$ (from the Cauchy-Schwarz inequality). Therefore, the map $\delta\mapsto\lambda^\delta$ is concave and nondecreasing in $(0,+\infty)$.

From this characterization, it also follows that the map $\delta\mapsto\lambda^\delta$ is not only nondecreasing but also increasing in $(0,+\infty)$. Indeed, to do so, assume by way of contradiction that there are two migration rates $0<\delta<\delta'$ such that $\lambda^\delta=\lambda^{\delta'}$. The function $\varphi_1^{\delta'}$ is a minimum of $\mathcal{R}(\delta',\varphi)$ among the functions $\varphi\in H^1(\mathbb{R}^n)\setminus\{0\}$ such that $\mathbf{x}\mapsto\|\mathbf{x}\|\,\varphi(\mathbf{x})\in L^2(\mathbb{R}^n)$. Thus, the monotonicity of $\mathcal{R}(\cdot,\varphi_1^{\delta'})$ in $(0,+\infty)$ yields:

$$\lambda^\delta\le\mathcal{R}(\delta,\varphi_1^{\delta'})\le\mathcal{R}(\delta',\varphi_1^{\delta'})=\lambda^{\delta'}=\lambda^\delta,$$

hence $\lambda^\delta=\mathcal{R}(\delta,\varphi_1^{\delta'})$, that is, $\varphi_1^{\delta'}$ also minimizes $\mathcal{R}(\delta,\varphi)$ among the same set of functions $\varphi$. Therefore, there is a constant $C>0$ such that $\varphi_1^{\delta'}\equiv C\,\varphi_1^\delta$ in $\mathbb{R}^n$, hence:

$$\delta'\,(\varphi_1^\delta-\varphi_1^\delta\circ\iota)\equiv\delta\,(\varphi_1^\delta-\varphi_1^\delta\circ\iota)\ \text{ in }\mathbb{R}^n,$$

from the equations satisfied by $\varphi_1^\delta$ and $\varphi_1^{\delta'}=C\,\varphi_1^\delta$. As a consequence, $\varphi_1^\delta\equiv\varphi_1^\delta\circ\iota$ in $\mathbb{R}^n$, that is, $\varphi_1^\delta\equiv\varphi_2^\delta$ in $\mathbb{R}^n$ by Lemma 6. Finally, the system $\mathcal{A}\,(\varphi_1^\delta,\varphi_2^\delta)=\lambda^\delta(\varphi_1^\delta,\varphi_2^\delta)$ yields $r_1\,\varphi_1^\delta\equiv r_2\,\varphi_1^\delta$ in $\mathbb{R}^n$, which is clearly impossible since $\varphi_1^\delta>0$ in $\mathbb{R}^n$ and $\mathbf{O}_1\neq\mathbf{O}_2$. Therefore, the map $\delta\mapsto\lambda^\delta$ is increasing in $(0,+\infty)$.

Let us now investigate the limits of $\lambda^\delta$ as $\delta\to0$ and $\delta\to+\infty$. First of all, one knows from (5.16) that $\lambda^\delta\ge-r_{\max}$ for all $\delta>0$ (this property can also be viewed as a consequence of (5.31) since $-r_1(\mathbf{x})=-r_{\max}+\|\mathbf{x}-\mathbf{O}_1\|^2/2\ge-r_{\max}$ for all $\mathbf{x}\in\mathbb{R}^n$). Furthermore, by choosing a symmetric test function, such as $\varphi_0(x)=e^{-\|\mathbf{x}\|^2}$ for instance, one has $\lambda^\delta\le\mathcal{R}(\delta,\varphi_0)$, and the quantity $\mathcal{R}(\delta,\varphi_0)$ is independent of $\delta$, hence $\sup_{\delta>0}\lambda^\delta<+\infty$. Therefore, there are two real numbers $\ell^0<\ell^\infty$ in $[-r_{\max},+\infty)$ such that $\lambda^\delta\to\ell^0$ as $\delta\to0$ and $\lambda^\delta\to\ell^\infty$ as $\delta\to+\infty$.

By defining $\mathcal{R}(0,\varphi)$ as above by deleting the (nonnegative) third term of the numerator of the Rayleigh quotient $\mathcal{R}(\delta,\varphi)$, one has:

$$\mathcal{R}(0,\varphi)\le\mathcal{R}(\delta,\varphi)\le\mathcal{R}(0,\varphi)+2\delta,$$

for every function $\varphi\in H^1(\mathbb{R}^n)\setminus\{0\}$ such that $\mathbf{x}\mapsto\|\mathbf{x}\|\,\varphi(\mathbf{x})\in L^2(\mathbb{R}^n)$. Thus, as $\delta\to0$, the minimum $\lambda^\delta$ of $\mathcal{R}(\delta,\varphi)$ over this set of functions $\varphi$ converges to the minimum $\ell^0$ of $\mathcal{R}(0,\varphi)$ over the same set, and this last minimum $\ell^0$ corresponds

to the principal eigenvalue of the Schrödinger operator,

$$-\frac{\mu^2}{2}\,\Delta - r_1(\mathbf{x}) = -\frac{\mu^2}{2}\,\Delta - r_{\max} + \frac{\|\mathbf{x}-\mathbf{O}_1\|^2}{2},$$

acting on the same set of functions. Since the principal eigenvalue of the operator $-\Delta + \|\mathbf{x}\|^2$ is equal to $n$ (with ground state, namely the principal eigenfunction, $\varphi_{GS}(\mathbf{x}) = e^{-\|\mathbf{x}\|^2/2}$ up to multiplicative constants), it easily follows by translation and scaling that $\ell^0 = -r_{\max} + \mu\, n/2 =: \lambda^0$, with principal eigenfunction $\varphi^0(\mathbf{x}) = e^{-\|\mathbf{x}-\mathbf{O}_1\|^2/(2\mu)}$ up to multiplicative constants.

In order to identity the real number $\ell^\infty = \lim_{\delta\to+\infty}\lambda^\delta = \lim_{k\to+\infty}\lambda^k$, we consider a sequence of (positive) principal eigenfunctions $(\varphi_1^k,\varphi_2^k)_{k\in\mathbb{N}} = (\varphi_1^k,\varphi_1^k\circ\iota)_{k\in\mathbb{N}}$ given by Lemma 6 (with $\delta = k \in \mathbb{N}$), normalized by $\|\varphi_1^k\|_{L^2(\mathbb{R}^n)} = 1$. For each $k\in\mathbb{N}$, there holds $\lambda^k = \mathcal{R}(k,\varphi_1^k)$, hence:

$$\frac{\mu^2}{2}\int_{\mathbb{R}^n}\|\nabla\varphi_1^k(\mathbf{x})\|^2\,\mathrm{d}\mathbf{x} + \int_{\mathbb{R}^n}\frac{\|\mathbf{x}-\mathbf{O}_1\|^2}{2}\,\varphi_1^k(\mathbf{x})^2\,\mathrm{d}\mathbf{x} + k\int_{\mathbb{R}^n}\Big(\varphi_1^k(\mathbf{x})^2 - \varphi_1^k(\mathbf{x})\varphi_1^k(\iota(\mathbf{x}))\Big)\,\mathrm{d}\mathbf{x} = r_{\max} + \lambda^k. \quad (5.32)$$

Notice that the right-hand side is bounded as $k\to+\infty$, while the left-hand side is the sum of three nonnegative terms. Therefore, the sequence $(\varphi_1^k)_{k\in\mathbb{N}}$ is bounded in $H^1(\mathbb{R}^n)$ and, up to extraction of a subsequence, there exists a nonnegative function $\varphi_1\in H^1(\mathbb{R}^n)$ such that $\varphi_1^k\to\varphi_1$ in $L^2_{loc}(\mathbb{R}^n)$ strongly, in $H^1_{loc}(\mathbb{R}^n)$ weakly, and almost everywhere in $\mathbb{R}^n$. Furthermore, since $\|\mathbf{x}-\mathbf{O}_1\|\to+\infty$ as $\|\mathbf{x}\|\to+\infty$, one has $\sup_{k\in\mathbb{N}}\|\varphi_1^k\|_{L^2(\mathbb{R}^n\setminus B(0,R))}\to 0$ as $R\to+\infty$, hence $\varphi_1^k\to\varphi_1$ in $L^2(\mathbb{R}^n)$ as $k\to+\infty$, and $\|\varphi_1\|_{L^2(\mathbb{R}^n)} = 1$. Fatou's lemma also implies that the function $\mathbf{x}\mapsto\|\mathbf{x}-\mathbf{O}_1\|\,\varphi_1(\mathbf{x})$ belongs to $L^2(\mathbb{R}^n)$, and so does the function $\mathbf{x}\mapsto\|\mathbf{x}\|\,\varphi_1(\mathbf{x})$. Moreover,

$$\int_{\mathbb{R}^n}\Big(\varphi_1^k(\mathbf{x})^2-\varphi_1^k(\mathbf{x})\varphi_1^k(\iota(\mathbf{x}))\Big)\,\mathrm{d}\mathbf{x}\to\int_{\mathbb{R}^n}\Big(\varphi_1(\mathbf{x})^2-\varphi_1(\mathbf{x})\varphi_1(\iota(\mathbf{x}))\Big)\,\mathrm{d}\mathbf{x},\qquad\text{as } k\to+\infty.$$

But since the left-hand side is $O(1/k)$ as $k\to+\infty$ by (5.32), one gets that:

$$\int_{\mathbb{R}^n}\Big(\varphi_1(\mathbf{x})^2-\varphi_1(\mathbf{x})\varphi_1(\iota(\mathbf{x}))\Big)\,\mathrm{d}\mathbf{x} = 0.$$

Since both functions $\varphi_1$ and $\varphi_1\circ\iota$ are nonnegative and with the same $L^2(\mathbb{R}^n)$ norm (equal to 1), the case of equality in the Cauchy-Schwarz inequality implies that:

$$\varphi_1 = \varphi_1\circ\iota,$$

almost everywhere in $\mathbb{R}^n$. Since each $C_0^\infty(\mathbb{R}^n)$ function $\varphi_1^k+\varphi_2^k = \varphi_1^k+\varphi_1^k\circ\iota$ obeys:

$$-\frac{\mu^2}{2}\,\Delta(\varphi_1^k+\varphi_2^k) - r_1\,\varphi_1^k - r_2\,\varphi_2^k = \lambda^k(\varphi_1^k+\varphi_2^k)\ \text{ in } \mathbb{R}^n,$$

and since $\varphi_1^k \to \varphi_1$ and $\varphi_2^k = \varphi_1^k \circ \iota \to \varphi_1 \circ \iota = \varphi_1$ in $L^2(\mathbb{R}^n)$ strongly and in $H^1_{loc}(\mathbb{R}^n)$ weakly, it then follows from a passage to the limit in the weak sense and from standard elliptic regularity theory that the function $\varphi_1$ is a $C^\infty(\mathbb{R}^n)$ solution of:

$$-\frac{\mu^2}{2}\,\Delta\varphi_1 - \frac{r_1+r_2}{2}\,\varphi_1 = \ell^\infty \varphi_1 \ \text{ in } \mathbb{R}^n.$$

Furthermore, since $\|\varphi_1\|_{L^2(\mathbb{R}^n)} = 1$ and since $\varphi_1$ is nonnegative, the elliptic strong maximum principle implies that $\varphi_1 > 0$ in $\mathbb{R}^n$. The $H^1(\mathbb{R}^n)$ function $\varphi_1$ is then a ground state of the Schrödinger operator $-(\mu^2/2)\Delta - (r_1+r_2)/2 = -(\mu^2/2)\Delta - r_{\max} + m_D/4 + \|\mathbf{x}\|^2/2$, with $m_D$ defined in (5.9). As a consequence, $\ell^\infty$ is the principal eigenvalue of this operator and $\varphi_1$ is its principal eigenfunction. In other words, $\ell^\infty = -r_{\max} + m_D/4 + \mu\, n/2 =: \lambda^\infty$ and $\varphi_1(\mathbf{x}) = (\pi\mu)^{-n/4}\, e^{-\|\mathbf{x}\|^2/(2\mu)}$. The proof of Proposition 35 is thereby complete. □

*Proof of Theorem 36.* Let $\mathbf{u}_\delta = (u_{\delta,1}, u_{\delta,2})$ be the unique $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)^2$ solution of (5.1) given by Theorem 32 and Remark 2, for growth functions $f_1, f_2$ of the first type (5.3), with a fixed initial condition $\mathbf{u}^0 = (u_1^0, u_2^0)$ independent of $\delta$ and such that both functions $u_1^0, u_2^0$ satisfy the assumptions (H1)-(H3). Let us fix two positive times $0 < T' \le T$ and let us show that $\sup_{t\in[T',T]} \|u_{\delta,1}(t,\cdot) - u_{\delta,2}(t,\cdot)\|_{L^\infty(\mathbb{R}^n)} \to 0$ as $\delta \to +\infty$.

From the first part of the proof of Theorem 32, especially from (5.21), (5.24)-(5.25) and similar calculations as the ones between (5.24) and (5.25), it follows that there exists a constant $K \ge 0$ (independent of $\delta > 0$) such that, for all $\delta > 0$,

$$|x_1 u_{\delta,2}(t,\mathbf{x})| \le |x_1 h_2(t,\mathbf{x})| \le K, \quad \text{for all } t \in [0,T] \text{ and } \mathbf{x} = (x_1,\dots,x_n) \in \mathbb{R}^n, \quad (5.33)$$

with $h_2$ defined by (5.18) (notice that the function $h_2$ actually depends on $\delta$, but the upper bound (5.24) is independent of $\delta > 0$). For each $\delta > 0$, one infers from (5.1)-(5.2) and (5.10) that the function $v_\delta := u_{\delta,1} - u_{\delta,2}$ is a classical $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ solution of:

$$\partial_t v_\delta(t,\mathbf{x}) = \frac{\mu^2}{2}\Delta v_\delta(t,\mathbf{x}) + r_1(\mathbf{x})\, v_\delta(t,\mathbf{x}) - 2\delta\, v_\delta(t,\mathbf{x}) - 2\beta x_1 u_{\delta,2}(t,\mathbf{x}),$$

such that $v_\delta$ is locally bounded in time and $v_\delta(t,x) \to 0$ as $\|\mathbf{x}\| \to +\infty$ locally uniformly in $t \in \mathbb{R}_+$. The previous relation, together with (5.20) and (5.33), implies that:

$$\begin{cases} -2\beta K \le \partial_t v_\delta(t,\mathbf{x}) - \dfrac{\mu^2}{2}\Delta v_\delta(t,\mathbf{x}) - (r_{\max} + m_1(\mathbf{x}))\, v_\delta(t,\mathbf{x}) + 2\delta\, v_\delta(t,\mathbf{x}) \le 2\beta K, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t \in [0,T],\ \mathbf{x} \in \mathbb{R}^n, \\ |v_\delta(0,\mathbf{x})| \le \max(\|u_1^0\|_{L^\infty(\mathbb{R}^n)}, \|u_2^0\|_{L^\infty(\mathbb{R}^n)}) =: M, \quad \mathbf{x} \in \mathbb{R}^n. \end{cases}$$

Since the potential $m_1(\mathbf{x}) = -\|\mathbf{x} - \mathbf{O}_1\|^2/2$ is nonpositive, there exists a

$C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ solution $V : \mathbb{R}_+ \times \mathbb{R}^n \to [0, M]$ of:

$$\begin{cases} \partial_t V(t,\mathbf{x}) = \dfrac{\mu^2}{2}\Delta V(t,\mathbf{x}) + m_1(\mathbf{x})\, V(t,\mathbf{x}), \quad t \geq 0,\ \mathbf{x} \in \mathbb{R}^n, \\ V(0,\mathbf{x}) = M, \quad \mathbf{x} \in \mathbb{R}^n. \end{cases}$$

Such a function $V$, which is independent of $\delta > 0$, can be obtained as the nondecreasing local limit as $R \to +\infty$ of $C^{1,2}(\mathbb{R}_+ \times \overline{B(0,R)})$ solutions $V^R : \mathbb{R}_+ \times \overline{B(0,R)} \to [0, M]$ of the same equation in $\mathbb{R}_+ \times \overline{B(0,R)}$, with Dirichlet boundary conditions $V^R = 0$ on $\mathbb{R}_+ \times \partial B(0,R)$ and initial conditions of the type $V^R(0,\mathbf{x}) = M\,\phi(\|\mathbf{x}\|/R)$ in $\overline{B(0,R)}$, where $\phi : [0,1] \to [0,1]$ is a $C^\infty([0,1])$ nonincreasing function such that $\phi = 1$ in $[0, 1/3]$ and $\phi = 0$ in $[2/3, 1]$.

Consider now any $\delta > r_{\max}/2$ and let $V_\delta$ be the $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ function defined in $\mathbb{R}_+ \times \mathbb{R}^n$ by:

$$V_\delta(t,\mathbf{x}) = v_\delta(t,\mathbf{x})\, e^{(2\delta - r_{\max})t} - \frac{2\beta K}{2\delta - r_{\max}} \Big( e^{(2\delta - r_{\max})t} - 1 \Big).$$

A straightforward calculation shows that:

$$\partial_t V_\delta(t,\mathbf{x}) - \frac{\mu^2}{2}\,\Delta V_\delta(t,\mathbf{x}) - m_1(\mathbf{x})\, V_\delta(t,\mathbf{x}) \leq m_1(\mathbf{x})\, \frac{2\beta K}{2\delta - r_{\max}} \Big( e^{(2\delta - r_{\max})t} - 1 \Big) \leq 0,$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$. Furthermore, $V_\delta(0,\mathbf{x}) = v_\delta(0,\mathbf{x}) \leq M = V(0,\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, and $\limsup_{\|\mathbf{x}\| \to +\infty} V_\delta(t,\mathbf{x}) \leq 0$ uniformly in $t \in [0,T]$. It follows from the maximum principle that $V_\delta(t,\mathbf{x}) \leq V(t,x)$ for all $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^n$, hence:

$$v_\delta(t,\mathbf{x}) \leq e^{(r_{\max} - 2\delta)t} V(t,\mathbf{x}) + \frac{2\beta K}{2\delta - r_{\max}} \Big( 1 - e^{(r_{\max} - 2\delta)t} \Big), \qquad \text{for all } (t,\mathbf{x}) \in [0,T] \times \mathbb{R}^n.$$

Since the function $V$ is bounded (by $M$), one gets that:

$$\limsup_{\delta \to +\infty} \Big( \sup_{[T',T] \times \mathbb{R}^n} v_\delta \Big) \leq 0,$$

recalling that $0 < T' \leq T$. The same argument applied to the functions $-V_\delta$ and $-v_\delta$ implies that, for all $\delta > r_{\max}/2$ and $(t,\mathbf{x}) \in [0,T] \times \mathbb{R}^n$,

$$v_\delta(t,\mathbf{x}) \geq -e^{(r_{\max} - 2\delta)t} V(t,\mathbf{x}) - \frac{2\beta K}{2\delta - r_{\max}} \Big( 1 - e^{(r_{\max} - 2\delta)t} \Big),$$

hence $\liminf_{\delta \to +\infty} \Big( \inf_{[T',T] \times \mathbb{R}^n} v_\delta \Big) \geq 0$. As a conclusion, $\sup_{[T',T] \times \mathbb{R}^n} |v_\delta| \to 0$ as $\delta \to +\infty$ and the proof of Theorem 36 is thereby complete. □

# 6

# Asexual evolution and demography over two islands: a weak selection strong mutation limit

**Abstract**

The evolution of an isolated asexual population is ruled by selection, mutation and drift. Several theoretical tools exist to model this dynamics, but they are not always accurate when the population is distributed over two habitats connected by migration. As seen in Chapter 5, the parabolic system associated with the phenotypic distributions in two habitats connected by migration is not usual. However, thanks to different mathematical tools, equilibria have been already approximated in several ways: if the migration is weak and the selection is strong, the phenotype distribution over the two habitats is bimodal (specialization), while if the migration is stronger, this distribution is unimodal (generalism). However, describing the transient behavior is a more open problem.

In this chapter, we focus on the description of the evolution of the mean fitness in each of two habitats, using the same framework proposed for single populations in Martin; Roques 2016; Hamel; Lavigne; Martin; Roques 2020; Lavigne; Martin; Anciaux; Papaix, et al. 2020; Roques; Patout; Bonnefon; Martin 2020. This framework, based on the generating function of the fitness distribution, cannot be directly applied and must be extended to a bivariate distribution of fitness across two habitats, among individuals from each habitat. The associated differential system is weakly coupled, which makes it difficult finding an explicit solution.

Thus we study a specific case, with symmetric migration, which simplifies the system into one nonlocal first order partial differential equation. A linearization of this operator allows deriving explicit solutions. Our results are compared with individual based simulations and with a numeric approximation of the solution of the initial problem. This suggests that the approach must be refined for better accuracy although it provides the correct qualitative behavior.

# Sommaire

## 6.1. Introduction

In a given habitat, an asexual population evolves under the actions of selection, mutation and genetic drift. At wider scale, because of heterogeneity over a given geographic landscape, individuals may migrate between habitats to which they are more or less adapted. The resulting dynamics can be fairly complex. Mutation typically has pleiotropic effects in different environments and this will strongly affect the joint evolutionary dynamics across habitats (see Hietpas; Bank; Jensen; Bolon 2013 for a detailed description of single nucleotide effects across environments in yeast). Drift also has more complex effects in metapopulations, both on the fixation of new mutations or the loss of standing genetic variance (Barton; Whitlock 1997).

In sexual populations, this is further complexified by recombination between migrant and local genotypes. However, even in asexuals, migration between different habitats impacts the demography and the evolution of a local population. This phenomenon has consequences on the mean fitness, which depends on the current local population size – how much the migrant genotypes dilute within the population they reach (Débarre; Ronce; Gandon 2013).

On the one hand, the interplay of mutation, selection and migration yields complex evolutionary dynamics, even with fixed demographic parameters (population sizes and migration rates). Mutation and migration both increase the local variance available for selection, but generate mutation and migration loads, due to the adverse effects of deleterious mutations and to maladapted migrant inflow respectively (Blows; Hoffmann 2005; Lenormand 2002). On the other hand, heterogeneous habitats can show complex non-stationary demographic dynamics, even ignoring evolutionary forces: local extinction/recolonisation events, local population expansions or decays.

These various feedbacks imply that modelling, analytical modelling especially, must rely on key simplifications. Many papers have been devoted to the dynamics of adaptation to heterogeneous conditions with stable demography (Gandon; Mirrahimi 2017b; Meszéna; Czibula; Geritz 1997; Mirrahimi; Gandon 2020) or to purely demographic dynamics in heterogeneous conditions but without evolution (*e.g.*,Brown; Pavlovic 1992; Freedman; Waltman 1977; Hastings 1983; Holt 1985; Holt; Gomulkiewicz 1997; Kawecki 1995; Kawecki 2000; Kawecki; Holt 2002; Levin 1974; Pulliam 1988 where demographic rates are fixed). Some models have tackled the interplay of the two in sexuals (Ronce; Kirkpatrick 2001) and in asexuals (Débarre; Ronce; Gandon 2013). In these models in particular, key simplifications were made to allow analytical treatment of the eco-evolutionary feedback, and we will indeed follow a similar strategy. These models use a maximally simplified demographic/ecological model, with non-overlapping generations, exponential or logistic growth or decay, two discrete habitats connected by a fixed symmetric probability of migrating per individual per generation.

The evolutionary dynamics are also simplified by considering deterministic dynamics (thus ignoring the effect of drift except in individual based simulation checks) and a simplified genetic basis for adaptation across environments. Let us detail this key aspect. Because we are dealing with polymorphic populations and new genotypes produced by recombination (in the case of sexuals) or mutation (in the case of sexuals and asexuals) a simple model must be taken to describe the joint fitness of each genotype across habitats. By far the most widely used strategy for this, in evolutionary ecology in general (Anciaux; Chevin; Ronce; Martin 2018; Débarre; Ronce; Gandon 2013; Gandon; Mirrahimi 2017b; Gomulkiewicz; Holt; Barfield 1999; Gomulkiewicz; Holt; Barfield; Nuismer 2010; Mirrahimi 2012; Mirrahimi; Gandon 2020) is to consider that each genotype corresponds to a given phenotype (possibly multivariate) for which fitness is maximized at some particular phenotypic optimum, which depends on the habitat. This very classic model allows mathematical treatment, intuitive graphical representations, yet it includes pervasive pleiotropy for fitness effects across habitats with a wealth of patterns ranging from antagonism (some mutations are beneficial in one environment and deleterious in another) to synergy (some are deleterious or beneficial in both habitats). This framework was also shown to generate patterns of mutation fitness effects across environments that were qualitatively consistent with experimental evolution data from various model species (Martin; Lenormand 2006b). In the present chapter we stick with all these simplifications.

Yet, even in this simple phenotype-fitness landscape with one optimum per habitat, the evolutionary dynamics are complex, even more so when coupled with demography. Further simplification have thus been necessary to give general analytical insight. In the case of sexuals, a very wide genetic basis for phenotype is assumed, with a fixed variance (Ronce; Kirkpatrick 2001), obtained as a limit when the number of independent loci increases, by invoking Fisher's infinitesimal model (Fisher 1930). In asexuals, obviously, such limit cannot be invoked. Previous work on asexuals have thus focused on another form of fixed variance equilibrium: when mutation, selection and migration are at equilibrium and selection is weak. In this limit, the phenotypic distribution is close to normal in each habitat with some fixed variance that can be determined analytically (Débarre; Ronce; Gandon 2013).

On the contrary, as migration becomes weaker and selection stronger, substantial deviations from normality, even at equilibrium, can arise because of local adaptation. The phenotype distribution over the two habitats is then bimodal, with local phenotypes tending to be closer to their corresponding local optimum (Débarre; Ronce; Gandon 2013). Migrants from the alternative habitat then introduce an excess of phenotypes lying at one side of the local phenotypic distribution, introducing asymmetry in the distribution (skewness of opposite sign in each habitat). When mild, the resulting deviations form normality have been treated analytically by various methods in recent papers, with a coupling to demography (Débarre; Ronce; Gandon 2013; Mirrahimi; Gandon 2020).

A key simplification required by these methods is the study of equilibrium states. Yet, the transient dynamics that lead to this equilibrium can be of equal importance, in terms of applications, and are more easily observable, empirically (as an equilibrium can take quite long to set). Our aim in this chapter is thus to explore an approach that would allow to deal with the transient dynamics of the same model, with a coupling of demographic and evolutionary dynamics, from arbitrary initial conditions, and allow for deviations from normality. To do so, we follow the lead of previous work (Débarre; Ronce; Gandon 2013) and use the same assumptions described above with a simple phenotype-fitness landscape, deterministic demography and evolution, symmetric migration, *etc.* However, we aim to a description of the transient dynamics of the system (and of course to retrieve previous results on equilibria).

More precisely, we model adaptation and mutation, in each habitat, according to a somewhat classic version of Fisher's Geometric Model (FGM) (Martin; Lenormand 2015; Tenaillon 2014), with gaussian mutation effects on phenotypes and quadratic fitness function relating growth rate and the phenotypic distance to the optimum. It has been shown recently (Anciaux; Chevin; Ronce; Martin 2018) that this model is handy when dealing with the demographic and evolutionary consequences of maladaptation, extinction and evolutionary rescue (the rebounce of an initially decaying isolated population, by the emergence and rise of beneficial mutants). Previously, we also studied the dynamics of a 'black hole sink' (Lavigne; Martin; Anciaux; Papaix, et al. 2020) under this model: a model with a recurrent input of migrants from a stable source into a habitat for which the mean fitness of source phenotypes is initially negative (a sink). The proposed work is thus a generalisation of these approaches when migrants can be exchanged both ways between the two habitats, so that none can be deemed stable and dynamics are coupled (between habitats and between demography and evolution). Beyond its mathematical convenience, note that this particular version of the FGM yields distributions of mutation effects on fitness that were shown to accurately predict or reproduce several qualitative and quantitative patterns from empirical mutant distributions (see Hietpas; Bank; Jensen; Bolon 2013; Tenaillon 2014; Trindade; Sousa; Gordo 2012 for more details).

We now turn to the details of the model, starting with its basic assumptions and the approximations used for the analytical developments.

## Individual-based model

We first describe the generic assumptions used in the stochastic individual based simulations that were used to check the range of validity of the analytical approximations.

*Phenotype-fitness relationships in the two habitats.* Under to the FGM, each individual is characterized by a given breeding value for a phenotype at $n$ biological traits, namely a vector $\mathbf{x} \in \mathbb{R}^n$. Furthermore, this model assumes the existence of

a phenotypic optimum for each habitat. We will denote by $\mathbf{O}_1$ and $\mathbf{O}_2$ the optima in respectively habitat 1 and 2. Up to translation and rotation, we can make the assumption that:

$$\mathbf{O}_1 = (-\beta, 0, \ldots, 0) \in \mathbb{R}^n \qquad \text{and} \qquad \mathbf{O}_2 = (\beta, 0, \ldots, 0) \in \mathbb{R}^n, \tag{6.1}$$

with $\beta \geq 0$. Thus the distance $\|\mathbf{O}_1 - \mathbf{O}_2\| = 2\beta$ characterizes how distinct (adaptively, so to speak) the two habitats are, and it determines the level of maladaptation induced by a change of habitat. In fact, we quantify the 'harshness of the environmental stress' imposed by habitat change by the quantity:

$$m_D = \frac{\|\mathbf{O}_1 - \mathbf{O}_2\|^2}{2} = 2\beta^2. \tag{6.2}$$

which, as we will see below, is the relative fitness (difference in growth rates) between the phenotypic optimum in habitat 1 and the phenotypic optimum in habitat 2, when measured in either of the two habitats.

While the phenotypic optimum in habitat 1 has maximal growth rate (or *absolute fitness*) $r_{\max,1}$, the growth rate of an individual will decrease away from this optimum. Under our classic version of the FGM, this decrease is a quadratic function: the *absolute fitness* of $\mathbf{x}$ is:

$$r_1(\mathbf{x}) = r_{\max,1} - \frac{\|\mathbf{x} - \mathbf{O}_1\|^2}{2}. \tag{6.3}$$

Hereafter, we will call *relative fitness* the quantity:

$$m_1(\mathbf{x}) = r_1(\mathbf{x}) - r_1(\mathbf{O}_1) = -\frac{\|\mathbf{x} - \mathbf{O}_1\|^2}{2} = -\frac{\beta^2}{2} - \frac{\|\mathbf{x}\|^2}{2} - \beta x_1. \tag{6.4}$$

which is the fitness of $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ relative to the phenotypic optimum in habitat 1. Correspondingly, in the second habitat with optimum $\mathbf{O}_2 \in \mathbb{R}^n$, the absolute fitness is $r_2(\mathbf{x}) = r_{\max,2} - \|\mathbf{x} - \mathbf{O}_2\|^2/2$ and the relative fitness is:

$$m_2(\mathbf{x}) = r_2(\mathbf{x}) - r_2(\mathbf{O}_2) = -\frac{\|\mathbf{x} - \mathbf{O}_2\|^2}{2} = -\frac{\beta^2}{2} - \frac{\|\mathbf{x}\|^2}{2} + \beta x_1. \tag{6.5}$$

*Reproduction, Selection and Drift*. The number of offspring from a given phenotype $\mathbf{x}$, in habitat $i$, is drawn into a Poisson distribution with parameter $e^{r_i(\mathbf{x})}$.

*Mutations*. In the first (resp. second) habitat, mutations occur at rate $U_1 > 0$ (resp. $U_2 > 0$). The offspring phenotype is given by independent and identically distributed (iid) random variations $\mathrm{d}\mathbf{x}$ around the phenotype of its parent, for each trait. We assume that the mutation phenotypic effects are distributed according to a standard Gaussian distribution (Kimura 1965; Lande 1980): $\mathrm{d}\mathbf{x} \sim \mathcal{N}(0, \lambda I_n)$, where $\lambda$ is the mutational variance at each trait, and $I_n$ is the identity matrix in $n$ dimensions. These mutations have some effect on fitness, which

distribution, conditional on the relative fitness $m_p < 0$ of the parent, has stochastic representation (Martin 2014; Tenaillon 2014) $s \sim -m_p - \frac{\lambda}{2}\chi^2_n(-2\,m_p/\lambda)$, where $\chi^2_n(-2\,m_p/\lambda)$ denotes the noncentral chi-square distribution with $n$ degrees of freedom and noncentrality $-2m_p/\lambda$.

*Migration events.* Every generation and simultaneously, a number of individuals emigrates from the first habitat into the second (resp. from the second into the first). This number is drawn into a Poisson law with parameter $\delta N_1(t)$ (resp. $\delta N_2(t)$), where $\delta > 0$ is the *per capita per generation* migration rate.

Note that because of the discrete time scheme here, the number of migrants must be drawn into a binomial law Bin$(N_i(t), 1 - e^{-\delta})$, where $\delta$ is the per capita per generation migration rate (in continuous time) so that $0 < 1 - e^{-\delta} < 1$ is a probability of migrating over one generation, for each individual, which is equal in both habitats. Under the continuous time approximation used in all analytical derivations, these two formulations become equivalent as we require $\delta \ll 1$ for the continuous time approximation to apply, so that $1 - e^{-\delta} \approx \delta$.

Note finally that we consider that each individual leaves its whole life in one habitat (per generation), while in the continuous time version (same model under our approximation) they may be found in any of the two habitats depending on their history over continuous time.

### Differential system satisfied by the phenotypic density

As long as changes per generation are small, we can approximate the discrete time process by continuous time dynamics. This requires the rates per generation to be small, namely $\delta,\, r_{\max,i},\, U_i \ll 1$. The biological system may also be in itself following continuous time dynamics (birth-death process), in which case no approximation is involved (although we used discrete time simulations for code efficiency and to check the model in its less favorable context). Furthermore, we use a deterministic approximation to these dynamics which *a priori* requires, on the other hand, that $N_i(t)\delta,\, N_i(t)r_{\max i},\, N_i(t)U_i \gg 1$, and hence that $N_i(t)$ be large enough at all times.

Under this deterministic continuous time approximation, we derive the effect of the three processes (population growth, migration, mutation) in terms of differential operators. We denote $u_i(t, \mathbf{x})$ the density of individuals with phenotype $\mathbf{x} \in \mathbb{R}^n$ at time $t > 0$ in the habitat $i$. First, the habitat and phenotype-specific population growth (reproduction and survival) is determined by the phenotype's absolute fitness in this habitat, *via* the following linear operator (exponential growth):

$$L_G(t, \mathbf{x}, u_i) = r_i(\mathbf{x})u_i(t, \mathbf{x}) = [r_{\max,i} + m_i(\mathbf{x})]u_i(t, \mathbf{x}).$$

Second, let us describe the mutation effects on phenotypes, under the standard FGM. When the mutational variance $\lambda$ is small, we can approximate the mutational effects on the densities by a linear elliptic operator (diffusion approximation

in phenotype space):

$$L_U(t, \mathbf{x}, u_i) = \frac{\mu_i^2}{2} \Delta u_i(t, \mathbf{x}),$$

with $\mu_i^2 = \lambda U_i > 0$, $\Delta = \sum_{i=1}^n \partial_{x_i x_i}$ and $\partial_{x_i x_i}$ denoting the second order partial derivative with respect to the $i$-th coordinate of $\mathbf{x}$ (*e.g.*, see Appendix of Hamel; Lavigne; Martin; Roques 2020, for more details).

Finally, migration from habitat $i$ to $j$ induces another linear operator:

$$L_M(t, \mathbf{x}, u_i, u_j) = \delta(u_j(t, \mathbf{x}) - u_i(t, \mathbf{x})).$$

Overall, the evolution of the habitat-specific densities induced by the effects of growth, migration and mutation can be approximated by a differential system:

$$\begin{cases} \partial_t u_1(t, \mathbf{x}) = \frac{\mu_1^2}{2} \Delta u_1(t, \mathbf{x}) + [r_{\max,1} + m_1(\mathbf{x})] u_1(t, \mathbf{x}) + \delta[u_2(t, \mathbf{x}) - u_1(t, \mathbf{x})], \\ \partial_t u_2(t, \mathbf{x}) = \frac{\mu_2^2}{2} \Delta u_2(t, \mathbf{x}) + [r_{\max,2} + m_2(\mathbf{x})] u_2(t, \mathbf{x}) + \delta[u_2(t, \mathbf{x}) - u_1(t, \mathbf{x})], \end{cases} \tag{6.6}$$

where $u_1(t, \mathbf{x})$ (resp. $u_2(t, \mathbf{x})$) is the density of individuals with phenotype $\mathbf{x} \in \mathbb{R}^n$ at time $t > 0$ in the first (resp. second) habitat, and for $i \in \{1, 2\}$, $\mu_i^2 = \lambda U_i$, $r_{\max i}$ the maximal growth rate in habitat $i$, and $m_i(\mathbf{x})$ defined by (6.4) and (6.5) (see Holt 1996b for details about the migration term). The population size is define as the integral of the density over all phenotypes:

$$N_1(t) = \int_{\mathbb{R}^n} u_1(t, \mathbf{x}) \, \mathrm{d}\mathbf{x}, \qquad \text{and} \qquad N_2(t) = \int_{\mathbb{R}^n} u_2(t, \mathbf{x}) \, \mathrm{d}\mathbf{x}. \tag{6.7}$$

Note that because these dynamics involve linear operators, they are in fact exactly describing the expectation of the densities, at all times (no deterministic approximation). Therefore, provided the diffusive approximation (mutation) and the continuous time approximation apply, then the system (6.6) describes exactly the expected densities in the two populations.

Unfortunately, it seems too hard to find analytic results on the solutions of the general problem (6.6). That is the reason why in this chapter we only focus on a symmetric case. In each habitats, the individuals mutate with the same rate $U$ (*i.e.*, $U := U_1 = U_2$), the growth rate at each optimum is equal (*i.e.*, $r_{\max} := r_{\max,1} = r_{\max,2}$), and the two densities $u_1$ and $u_2$ are initially symmetric according to the hyperplane $\{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n, \, x_1 = 0\}$:

$$\begin{cases} U := U_1 = U_2, \\ r_{\max} := r_{\max,1} = r_{\max,2} > 0, \\ \forall \mathbf{x} \in \mathbb{R}^n, \, u_1(0, \mathbf{x}) = u_2(0, \iota(\mathbf{x})), \end{cases} \tag{SH}$$

with:

$$\forall \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n, \, \iota(\mathbf{x}) = (-x_1, x_2, \ldots, x_n). \tag{6.8}$$

Unless otherwise specified, we make this assumption.

### Aim

In this chapter, we try to find analytic results regarding the differential equation (6.6). We only focus on the evolution of the mean fitness in each environment, defined by:

$$\overline{m}_1(t) = \int_{\mathbb{R}^n} m_1(\mathbf{x}) \frac{N_1(t,\mathbf{x})}{N_1(t)}\,\mathrm{d}\mathbf{x}, \quad \text{and} \quad \overline{m}_2(t) = \int_{\mathbb{R}^n} m_2(\mathbf{x}) \frac{N_2(t,\mathbf{x})}{N_2(t)}\,\mathrm{d}\mathbf{x}.$$

In Section 6.2, we enounce different mathematical properties on the differential equation satisfied by the Laplace transform of the fitness distribution. However, the differential system is hard to solve. Thus in Section 6.2.3, we develop some analytic approximation on the distribution. This approximation is checked against simulations in Section 6.3. The proof of all results are postponed in different appendices.

## 6.2. Results

In this section, we develop different analytical results about the mean fitness. Before this, we have to give a mathematical result, enounced in Chapter 5. We can prove that the Cauchy problem associated to (6.6) admits a unique solution, under some assumptions on the initial phenotype distribution $(u_1^0, u_2^0)$:

(H1) $u_1(0,\cdot), u_2(0,\cdot) \in C^{2+\alpha}(\mathbb{R}^n)$ for some $\alpha \in (0,1)$, that is:

$$\|u_1(0,\cdot)\|_{C^{2+\alpha}(\mathbb{R}^n)} < +\infty, \qquad \text{and} \qquad \|u_2(0,\cdot)\|_{C^{2+\alpha}(\mathbb{R}^n)} < +\infty;$$

(H2) $N_1(0) := \displaystyle\int_{\mathbb{R}^n} u_1(0,\mathbf{x})\,\mathrm{d}\mathbf{x} < +\infty$ and $N_2(0) := \displaystyle\int_{\mathbb{R}^n} u_2(0,\mathbf{x})\,\mathrm{d}\mathbf{x} < +\infty$;

(H3) there exists a non-increasing function $g : \mathbb{R}_+ \to \mathbb{R}$ (with $\mathbb{R}_+ = [0,+\infty)$) such that:

- $0 < u_1(0,\mathbf{x}) \leq g(\|\mathbf{x} - \mathbf{O}_1\|)$ and $0 < u_2(0,\mathbf{x}) \leq g(\|\mathbf{x} - \mathbf{O}_2\|)$ for all $\mathbf{x} \in \mathbb{R}^n$;
- $g \in L^1(\mathbb{R}_+, \mathbb{R}_+) \cap C^0(\mathbb{R}_+, \mathbb{R}_+)$;
- the functions $\mathbf{x} \mapsto m_1(\mathbf{x}) g(\|\mathbf{x} - \mathbf{O}_1\|)$ is bounded in $\mathbb{R}^n$;
- $\displaystyle\int_{\mathbb{R}^n} |m_1(\mathbf{x})| g(\|\mathbf{x} - \mathbf{O}_1\|)\,\mathrm{d}\mathbf{x}$ is finite.

Hereafter, we assume that these hypotheses hold.

### 6.2.1. Fitness distribution satisfying a degenerate parabolic system

As said in the Introduction, this chapter presents a framework based on the fitness distribution and its Laplace transform. This method has already been used

for the asexual population model (*e.g.*, Hamel; Lavigne; Martin; Roques 2020; Lavigne; Martin; Anciaux; Papaix, et al. 2020; Martin; Roques 2016).

While Hamel; Lavigne; Martin; Roques 2020 have studied the evolution of the distribution of one fitness, their approach must be extended here: after a migration event, the individuals have not the same fitness, and the two habitats composition evolves across the time. To address this problem, we study the evolution of the density of the individuals having the bivariate fitness set $(m_1, m_2)$, where $m_1$ (resp. $m_2$) is the fitness of the individual, if it lives in the first (resp. second) habitat: an individual with phenotype $\mathbf{x} \in \mathbb{R}^n$ has the fitness set $(m_1(\mathbf{x}), m_2(\mathbf{x}))$, with $m_1$ and $m_2$ given by (6.4)–(6.5). We can check that the fitness vector is in the subset:

$$\begin{aligned}\Gamma &= \left\{(m_1, m_2) \in \mathbb{R}^2_-,\ B\left(\mathbf{O}_1, \sqrt{2|m_1|}\right) \bigcap B\left(\mathbf{O}_2, \sqrt{2|m_2|}\right) \neq \emptyset\right\} \\ &= \left\{(m_1, m_2) \in \mathbb{R}^2_-,\ \beta^2 + m_1 + m_2 + \frac{(m_2 - m_1)^2}{4\beta^2} \leq 0\right\}, \quad (6.9)\end{aligned}$$

where $B(\mathbf{O}, r)$ is the closed ball with center at $\mathbf{O} \in \mathbb{R}^n$ and radius $r > 0$. These fitness densities, denoted by $p_1$ and $p_2$, are directly linked with the phenotype densities $u_1$ and $u_2$:

$$p_i(t, m_1, m_2) = \frac{1}{2\beta}\left(2|\widetilde{m}_2|\right)^{(n-1)/2-1} \int_{\mathbb{S}^{n-2}} u_i\left(t, \widetilde{m}_1, \sqrt{2|\widetilde{m}_2|}\,\sigma\right) \mathrm{d}\sigma, \qquad (6.10)$$

with:

$$\widetilde{m}_1 = \frac{m_2 - m_1}{2\beta}, \qquad \text{and} \qquad \widetilde{m}_2 = \frac{\beta^2}{2} + \frac{m_1 + m_2}{2} + \frac{(m_2 - m_1)^2}{8\beta}.$$

These two functions satisfy a complex degenerate parabolic partial differential equation, which admits no analytic results. The strategy is to define the Laplace transform of the functions $p_1$ and $p_2$.

## 6.2.2. Linear system for the Laplace transform

As in Hamel; Lavigne; Martin; Roques 2020, we define the Laplace transform of the functions $p_1$ and $p_2$:

$$M_1(t, \mathbf{z}) := \int_\Gamma e^{z_1 m_1 + z_2 m_2} p_1(t, m_1, m_2)\, \mathrm{d}m_1 \mathrm{d}m_2 = \int_{\mathbb{R}^n} e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} u_1(t, \mathbf{x})\, \mathrm{d}\mathbf{x},$$

and

$$M_2(t, \mathbf{z}) := \int_\Gamma e^{z_1 m_1 + z_2 m_2} p_2(t, m_1, m_2)\, \mathrm{d}m_1 \mathrm{d}m_2 = \int_{\mathbb{R}^n} e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} u_2(t, \mathbf{x})\, \mathrm{d}\mathbf{x}. \qquad (6.11)$$

for all $t \geq 0$ and $\mathbf{z} = (z_1, z_2) \in \mathbb{R}^2_+$. Remark first that if $M_1(t, \mathbf{z})$ and $M_2(t, \mathbf{z})$ are well-defined, then the population size of each habitat is given by:

$$N_1(t) = M_1(t, 0, 0), \qquad \text{and} \qquad N_2(t) = M_2(t, 0, 0), \tag{6.12}$$

and the mean fitness (while the habitats are not empty) by:

$$\overline{m}_1(t) = \frac{1}{N_1(t)}\, \partial_{z_1} M_1(t, 0, 0), \quad \text{and} \quad \overline{m}_2(t) = \frac{1}{N_2(t)}\, \partial_{z_2} M_2(t, 0, 0), \tag{6.13}$$

where $\partial_{z_1}$ (resp. $\partial_{z_2}$) denotes the first order partial derivative with respect to $z_1$ (resp. $z_2$).

**Remark 4.** *The bivariate Laplace transform also allows to compute the covariance in fitness between habitats, among individuals from a given habitat. For example the covariance in fitness between habitats, among individuals from habitat 1 is given by* $\text{cov}_1(m_1, m_2) = \partial_{z_1 z_2} M_1(t, 0, 0)/N_1(t)$.

*These functions also retrieve the marginal univariate distributions of fitness (evaluated in a given habitat) among individuals from a given habitat (not necessarily the same). For example, the distribution of the fitness, in habitat 1, among individuals from habitat 1, has Laplace transform* $M_1(t, z_1, 0)$ *while the distribution of the fitness, if evaluated in habitat 2, among the same individuals from habitat 1 has Laplace transform* $M_1(t, 0, z_2)$.

Let us to study the function $\mathbf{M} = (M_1, M_2)$, which satisfies a first order partial differential equation given by:

**Proposition 37.** *The functions* $M_1$ *and* $M_2$*, defined by* (6.11)*, are real valued. The vector function* $\mathbf{M} = (M_1, M_2)$ *is a classical* $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^2_+)$ *solution of:*

$$\forall t > 0,\ \forall (z_1, z_2) \in \mathbb{R}^2_+,\ \partial_t \mathbf{M}(t, \mathbf{z}) = S_1(\mathbf{z}) \partial_{z_1} \mathbf{M}(t, \mathbf{z}) + S_2(\mathbf{z}) \partial_{z_2} \mathbf{M}(t, \mathbf{z}) + A(\mathbf{z}) \mathbf{M}(t, \mathbf{z}), \tag{6.14}$$

*where the matrices are:*

$$S_1(\mathbf{z}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \mu^2 (z_1 + z_2) z_1\, I_2, \quad S_2(\mathbf{z}) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} - \mu^2 (z_1 + z_2) z_1\, I_2, \tag{6.15}$$

*and:*

$$A(\mathbf{z}) = \left[ r_{\max} - \mu^2 \left[ m_D z_1 z_2 + \frac{n}{2}(z_1 + z_2) \right] \right] I_2 + \begin{pmatrix} -\delta & \delta \\ \delta & -\delta \end{pmatrix}, \tag{6.16}$$

*the parameters* $\mu^2 = \lambda U$ *and* $m_D = 2\beta^2$*, and* $I_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

As for the differential system (6.6), the partial equation (6.14) has no general analytic solution. In the specific case of isolated islands, *i.e.*, $\delta = 0$, this model implies the same evolution of the mean fitness, and the same formula for the mutational load:

$$\text{Load}_{\text{mut}}(\mu, n) = \lim_{t\to\infty} \overline{m}(t) = -\frac{\mu n}{2}, \tag{6.17}$$

as in Hamel; Lavigne; Martin; Roques 2020. This formula can be clearly extended, even without the assumption (SH).

The simplification given by proposition 2.2 of Chapter 5 has consequences on the partial differential system (6.14):

**Corollary 9.**
*The solution* $(M_1, M_2)$ *of* (6.14) *given by Proposition 37 satisfies:*

$$\forall t > 0,\ \forall \mathbf{z} = (z_1, z_2) \in \mathbb{R}^2_+,\ M_1(t, z_1, z_2) = M_2(t, z_2, z_1),$$

*and* $M_1$ *is a classical solution of the nonlocal first order partial differential equation:*

$$\begin{aligned}\partial_t M_1(t,\mathbf{z}) &= [1-\mu^2(z_1+z_2)z_1]\partial_{z_1} M_1(t,\mathbf{z}) - \mu^2(z_1+z_2)z_2\partial_{z_2} M_1(t,\mathbf{z}) \\ &+ \left[r_{\max} - \mu^2\left(m_D z_1 z_2 + \frac{n}{2}(z_1+z_2)\right)\right] M_1(t,\mathbf{z}) + \delta[M_1(t,z_2,z_1) - M_1(t,\mathbf{z})],\end{aligned} \tag{6.18}$$

*for all* $t > 0$ *and* $\mathbf{z} = (z_1, z_2) \in \mathbb{R}^2_+$.

### 6.2.3. Solution of an approached problem

The partial differential equation (6.18) is not standard, due to the nonlocal term $M_1(t, z_2, z_1)$, and to our knowledge, there does not exist a general framework to study it analytically. However, this term can be approximated, using Taylor expansions, by:

$$M_1(t, z_2, z_1) \approx M_1(t, z_1, z_2) + (z_2 - z_1, z_1 - z_2) \cdot \nabla M_1(t, z_1, z_2),$$

This expansion applies when focusing on the region of the solution where $z_1 \approx z_2$, in particular near $z_1 = z_2 = 0$. An interpretation of this in terms of distributions is that the approximation should correctly describe the bulk of the bivariate fitness distribution but not necessarily its tail. To see this, consider, say, the marginal distributions of fitness in habitat 1, among individuals from this habitat. Its Laplace transform is $M_1(t, z_1, 0)$ , which is only well approximated by the Taylor expansion near $z_1 \approx z_2 = 0$. This means that the bulk of the marginal

distribution is well described but not necessarily its tail.

The Taylor expansion leads to the simpler PDE:

$$\begin{aligned}\partial_t M_1(t,\mathbf{z}) \approx [1+\delta(z_2-z_1)-\mu^2(z_1+z_2)z_1]\partial_{z_1}M_1(t,\mathbf{z})\\ +\left[\delta(z_1-z_2)-\mu^2(z_1+z_2)z_2\right]\partial_{z_2}M_1(t,\mathbf{z})\\ +\left[r_{\max}-\mu^2\left(m_D z_1 z_2+\frac{n}{2}(z_1+z_2)\right)\right]M_1(t,\mathbf{z}). \quad (6.19)\end{aligned}$$

Approximating the function $M_1$ by the solution of the previous partial differential equation, we find the same criterion of Explosion vs Extinction than in corollary 7 in Chapter 5, with an explicit approximation $\delta^*$ of the critic migration rate value.

**Theorem 38.**
*For $m_D \leq 4\left(r_{\max}-\frac{\mu n}{2}\right)$, the population explodes, while, if $m_D > 4\left(r_{\max}-\frac{\mu n}{2}\right)$, there exists a critical value $\delta^* > 0$, given by:*

$$\delta^* = \delta^*(r_{\max},\mu,m_D,n) := \frac{\mu}{2}\left[-1+\sqrt{\frac{m_D}{m_D-4\left(r_{\max}-\frac{\mu n}{2}\right)}}\right],$$

*such that:*

- *$\diamond$ for $\delta < \delta^*$, the population size explodes exponentially;*
- *$\diamond$ for $\delta > \delta^*$, the population goes to extinction.*

*The mean fitness converges to:*

$$\overline{m}(\infty) = -\frac{\mu n}{2}-\frac{m_D}{4}\left(1+\mu^2 \underset{\delta\to+\infty}{o}(1)\right), \quad (6.20)$$

*as $t\to+\infty$, with $\mathcal{L}(\mu,\delta)$ tending to 0, as $\delta\to+\infty$ (see the proof for a formula for $\mathcal{L}(\mu,\delta)$).*

The first remark of Theorem 38 is an explicit formula for the migration load:

$$\text{Load}_{\text{migr}}(\mu,m_D,\delta) = -\frac{m_D}{4}\left(1+\mu^2 \underset{\delta\to+\infty}{o}(1)\right).$$

Quite intuitively, the migration load is proportional to the stress harshness $m_D$: the more different the two environments, the more migration decreases the ultimate mean fitness. Moreover, for strong migration (*i.e.*, $\delta\to+\infty$), the population suffers to a migration load equal to $-m_D/4$. This constant corresponds to the fitness of the point $(0,\ldots,0)$ in each environment. Thus, when the migration is strong enough, the system is well mixed, and the population density in the whole

environment is unimodal: the individuals are generalists. On the other hand, as a sanity check we can also see (Section 6.5.2) that when $\delta = 0$ the migration load is zero, as should be. In between lies a continuum from local adaptation to generalism.

For all parameters $\delta$ and $m_D$, a measure of local adaptation is given by the following coefficient:

$$LA_1(r_{\max}, \mu, n, \delta, t) = \frac{\partial_{z_1} M_1(t,0,0)}{M_1(t,0,0)} - \frac{\partial_{z_2} M_1(t,0,0)}{M_1(t,0,0)},$$
$$\text{and} \quad LA_2(r_{\max}, \mu, n, \delta, t) = \frac{\partial_{z_2} M_2(t,0,0)}{M_2(t,0,0)} - \frac{\partial_{z_1} M_2(t,0,0)}{M_2(t,0,0)}, \quad (6.21)$$

for all time $t \geq 0$. In our case, as the habitats are symmetric, we can check that:

$$\forall t \geq 0,\ LA_1(r_{\max}, \mu, n, \delta, t) = LA_2(r_{\max}, \mu, n, \delta, t).$$

These coefficient represents the difference between the mean fitness of the population living in the first habitat, and the mean fitness of this population if it would live in the second habitat. The closer $LA_1$ gets to 0, the more generalist the population as it tends to have the same mean fitness in both habitats.

In the particular case, where the two environments are symmetric (with hypothesis SH, we can find an explicit formula for $LA_1$ (see Section 6.5.1.4), which converges, as $t \to +\infty$:

**Theorem 39.**
*The local adaptation coefficient $LA_1$ converges to:*

- $LA_1(r_{\max}, \mu, n, \delta, +\infty)$, *if* $\mu = 2\delta$, *with:*

$$LA_1(r_{\max}, \mu, n, \delta, +\infty) = \frac{m_D}{4}\left[1 - {}_2F_1(2,2;3,-1) + \frac{1}{2}\,{}_3F_2(2,2,2;3,3;-1)\right];$$

- $LA_1(r_{\max}, \mu, n, \delta, +\infty)$, *if* $\mu \notin \{0, 2\delta\}$, *with:*

$$\begin{aligned} LA_1(r_{\max}, \mu, n, \delta, +\infty) = {} & \frac{4\delta\mu^2 m_D}{(4\delta^2-\mu^2)(2\delta+\mu)}\,{}_2F_1\left[1, \frac{\delta}{\mu}+\frac{1}{2}; \frac{\delta}{\mu}+\frac{3}{2}; -1\right] \\ & - \frac{4\delta\mu^2 m_D}{(4\delta^2-\mu^2)(2\delta+\mu)}\,{}_2F_1\left[2, 2\frac{\delta}{\mu}+1; 2\frac{\delta}{\mu}+2; -1\right] \\ & - \frac{2\mu^3 m_D}{(4\delta^2-\mu^2)(2\delta+\mu)}\,{}_2F_1\left[2, \frac{\delta}{\mu}+\frac{1}{2}; \frac{\delta}{\mu}+\frac{3}{2}; -1\right] \\ & + \frac{2\mu^3 m_D}{(4\delta^2-\mu^2)(2\delta+3\mu)}\,{}_2F_1\left[2, \frac{\delta}{\mu}+\frac{3}{2}; \frac{\delta}{\mu}+\frac{5}{2}; -1\right]; \end{aligned}$$

*where ${}_pF_q$ is the Generalized Hypergeometric Function (see Appendix 6.5.1.3 for more details).*

This local adaptation coefficient has also a phenotypic interpretation:

$$\begin{aligned} \forall t \geq 0,\ LA_1(r_{\max}, \mu, n, \delta, t) &= \int_{\mathbb{R}^n} m_1(\mathbf{x})\frac{u_1(t,\mathbf{x})}{N(t)}\,\mathrm{d}\mathbf{x} - \int_{\mathbb{R}^n} m_2(\mathbf{x})\frac{u_1(t,\mathbf{x})}{N(t)}\,\mathrm{d}\mathbf{x}, \\ &= -2\beta\int_{\mathbb{R}^n} x_1\frac{u_1(t,\mathbf{x})}{N(t)}\,\mathrm{d}\mathbf{x}. \end{aligned}$$

Thus the $LA_1$ coefficient is linked with the mean first trait $\overline{\mathbf{x}}_1(t) := \int_{\mathbb{R}^n} x_1 \frac{u_1(t,\mathbf{x})}{N(t)}\,\mathrm{d}\mathbf{x}$ in the first habitat by:

$$\overline{\mathbf{x}}_1(t) = -\frac{1}{2\beta}\,LA_1(r_{\max}, \mu, n, \delta, t),$$

for all $t \geq 0$.

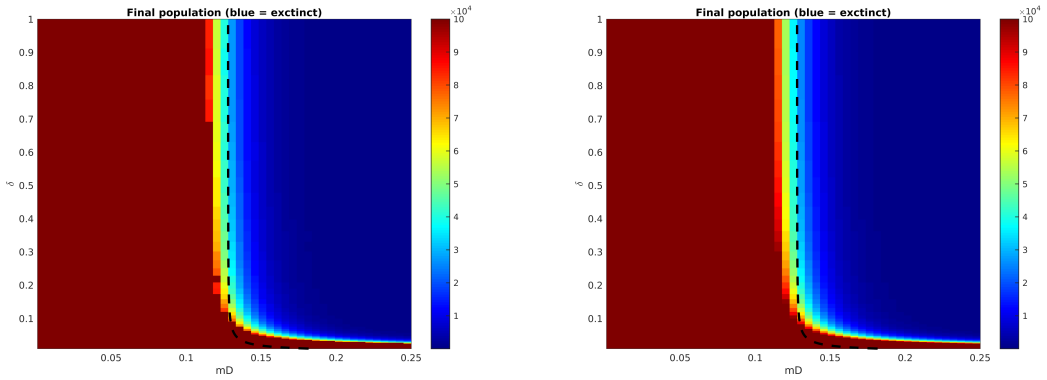
## 6.3. Simulations

In this section, although the previous results have been established with continuous initial distribution, we assume that the initial population is clonal distributed in each habitat with the same phenotype $(\mathbf{O}_1 + \mathbf{O}_2)/2 = 0 \in \mathbb{R}^n$. Formally, this initial condition can be seen as a continuous distribution which variance tends to

0. Thus the initial Laplace transforms are given by:

$$M_1(t, \mathbf{z}) = M_2(t, \mathbf{z}) = N_1^0 \exp[-m_D(z_1 + z_2)/4],$$

with $N_1^0 = 10^4$ individuals.




(a) (b)

Figure 6.1. – **Extinction vs explosion of the population, given by simulations induced by the stochastic model.** The heat map represents the population size $N_1(t)$ in the first habitat, with respect to the migration rate $\delta$ and the harshness of the stress $m_D$, at time $t = 300$, given by (a) the IBM (b) a numeric solution of Eq. (6.6). The parameters are $\lambda = 1/300$, $n = 2$, $U_1 = U_2 = 50U_c$ (with $U_c = n^2\lambda/4$). The dark dashed curve is the analytic formula for the critical value $\delta^*$ given by Theorem (38), with $r_{\max} = 50\,U_c/3$. Initially, each habitats has $N_0 = 10^4$ individuals, with same phenotype $\mathbf{x}_0 = (0, 0)$, in each habitat.

As seen in Chapter 5, the solution of the differential system (6.6) gives a good approximation of the evolution of the population size (see Fig.6.1), which shows that the continuous time deterministic approximation has some predictive power. At fixed $m_D$, there exists a threshold $\delta^*$ such that, for larger migration rate, *i.e.*, $\delta > \delta^*$, the population cannot adapt to the environments, and goes to extinction. For lower migration rates, *i.e.*, $\delta < \delta^*$, the population size seems to explode. As in Theorem 38, this critic value exists, only for $m_D > m_{D,crit}$. The dashed curve in Fig.6.1 represents the analytic formula of $\delta^*(m_D)$, which seems accurate.

The evolution of the mean fitness is ploted in Fig 6.2, for different values of the migration parameters $m_D$ and $\delta$. We can see that the mean value of $\overline{m}(t)$ over the different replicates (blue curve) is bounded by the mean fitness given by the differential system (6.6) (red curve) and our approximation, developed in Appendix 6.5.1.4 (dark curve). While our approximation of the mean fitness in

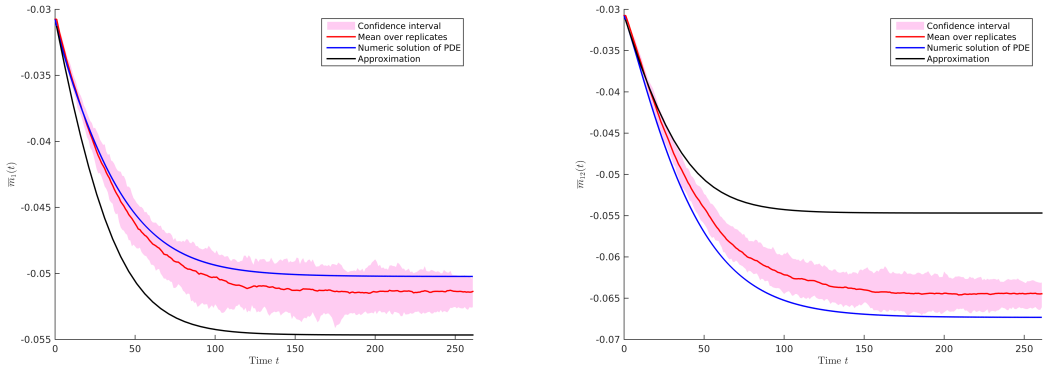
one environment seems be validate, our local adaptation coefficient $LA_1$ and so the mean phenotype $\overline{\mathbf{x}}_1$ does not approximate sufficiently well the mean value over the replicates (see Fig. 6.3). This phenomenon can be explained by Fig. 6.2. First, our approximation underestimate the real value of $\overline{m}_1(t)$. Last, the second mean fitness:

$$\overline{m}_{12}(t) = \int_{\mathbb{R}^n} m_2(\mathbf{x}) \frac{N_1(t,\mathbf{x})}{N_1(t)} \, \mathrm{d}\mathbf{x}, \tag{6.22}$$

is overestimated. As the local adaptation coefficient $LA_1$ satisfies:

$$LA_1(t) = \overline{m}_1(t) - \overline{m}_{12}(t),$$

our model underestimates the real value of the coefficient $LA_1$.




(a) (b)

Figure 6.2. – **Evolution of the mean fitness vector in the first habitat,** averaged over replicates, given by IBM (red curve), by a numeric solution of (6.6) (blue curve) and by our analytic approximation (dark curve). (a) Plot of the mean fitness $\overline{m}_1(t)$ (b) Plot of the mean fitness defined by (6.22). The parameters are $\lambda = 1/300$, $n = 2$, $U_1 = U_2 = 50U_c$ (with $U_c = n^2\lambda/4$), $m_D = 0.123$ and $\delta = 0.1$. Initially, each habitats has $N_0 = 10^4$ individuals, with same phenotype $\mathbf{x}_0 = (0,0)$, in each habitat.

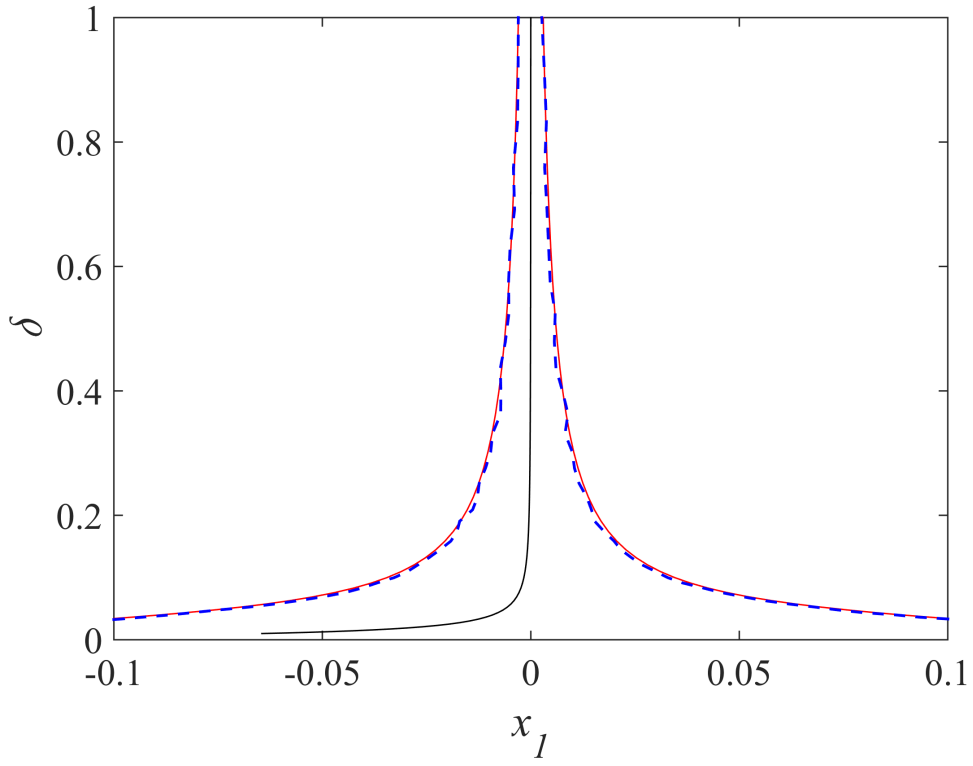



Figure 6.3. – **Large time behaviour of the mean phenotype $\overline{\mathbf{x}}_1$,** with respect to the migration rate $\delta$. The blue dashed curve is the mean value over the replicates given by the IBM, while the red plain curve is the mean phenotype $\overline{\mathbf{x}}_1$ of the numeric solution of (6.6). The dark plain curve represent the approximation of $\overline{\mathbf{x}}_1$, given by Theorem 39. The parameters are $\lambda = 1/300$, $n = 2$, $U_1 = U_2 = 50U_c$ (with $U_c = n^2\lambda/4$) and $m_D = 0.123$. Initially, each habitats has $N_0 = 10^4$ individuals, with same phenotype $\mathbf{x}_0 = (0, 0)$, in each habitat.

## 6.4. Discussion

In this chapter, we have developed a framework to approximate the dynamic of the mean fitness for the symmetric sinks problem. This formula lets us to explore the large time behaviour of the population size, which depends on the harshness $m_D$. When the two habitats are similar, *i.e.*, $m_D$ is low, the population size explodes, while when they are too different, *i.e.*, $m_D$ is strong, the population size can explode or go to extinction. More precisely, in this last regime, there exists a threshold $\delta_{crit}$, such that if $\delta \leq \delta_{crit}$, the population persists, and if $\delta > \delta_{crit}$, the population disappears in both habitats. This result is already known, without an explicit expression for $\delta_{crit}$ (*cf.* Corollary 7 in Chapter 5). The aim of this

chapter is to find so an approximation of this value $\delta_{crit}$ (see Theorem 38). This approximation was checked thanks to IBM simulations.

The last quantity of interest is the local adaptation coefficient, which can quantify the specialization (or the generalization) of the population. However, the error to approximate $LA$ is amplified: our framework seems to be not applied to study the faith of the phenotype distributions.

## Effect of the mutation rate $U$ – Seed bank problem

We are interested by the effect of the evolution and the adaptation of an asexual population, when the two habitats admit the same maximal growth rate $r_{\max}$, but the population mutate at different mutation rate $U_1$ and $U_2$. We believe this model could be relevant for the seed bank problem. Consider an asexual plant with a seed bank. Assume that the phenotypes with maximal survival of the seeds in the ground is distinct from the phenotypes with maximal fitness above ground. Migration here goes both ways (plants to seeds and seeds to plants). The mutation rate in the seed bank is zero (as there is no reproduction only selection on survival) while it is not zero above ground. We assume that the global population is clonal at the optimum $\mathbf{O}_1$, with the vector fitness $(0, -m_D)$.

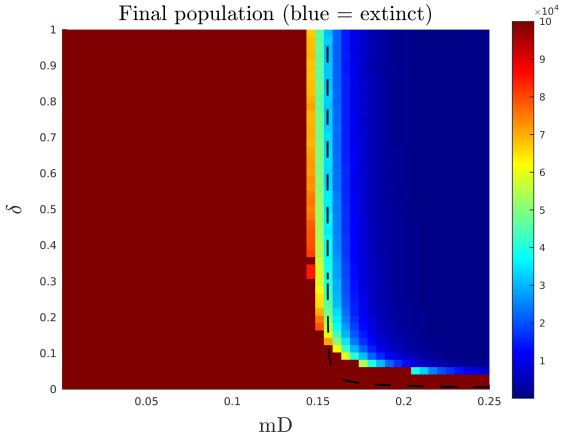



Figure 6.4. – **Extinction vs explosion of the population, given by simulations induced by the stochastic model.** The heat map represents the population size $N_1(t)$ in the first habitat, with respect to the migration rate $\delta$ and the harshness of the stress $m_D$, at time $t = 300$. The parameters are $\lambda = 1/300$, $n = 2$, $U_1 = 50U_c$ (with $U_c = n^2\lambda/4$), $U_2 = 0$, and $r_{\max} = 50U_c/3$. The dark dashed (resp. plain) curve is the analytic formula for the critical value $m_{D,crit}$ given by Theorem (38), with $U = (U_1 + U_2)/2 = 25U_c$. Initially, each habitats has $N_0 = 10^4$ individuals, with same phenotype $\mathbf{x}_0 = (0, 0)$, with in each habitat.

Comparing the heatmap of Fig. 6.4 with the analytic formula (dark dashed curve) for the symmetric habitats problem, with $U = 25U_c$, the population can easier adapt to two different habitats, when one is a seed bank, than when the two habitats are symmetric. However, for a strong migration rate $\delta$, the boundary extinction/explosion can be approximated by the critical value $m_{D,crit}(r_{\max}, \mu_1/\sqrt{2}, n\delta)$ (which corresponds to the case where $U$ is the mean value between $U_1 = 50U_c$ and $U_2 = 0$).

# 6.5. Appendix

## 6.5.1. Proofs

This section is devoted to the proofs of the different results presented in Section 6.2. The relation between the phenotype distributions and the fitness distribution is developed in Section 6.5.1.1. This relation will let us to define the Laplace transform of the fitness distributions (see Section 6.5.1.2. These functions are solutions of a coupled partial differential system, which has no simple explicit solution. We present the computations of the solution of the approximation (6.19) in Section 6.5.1.4.

### 6.5.1.1. The fitness distribution

First, we focus on the fitness vector distribution. Similarly as in Hamel; Lavigne; Martin; Roques 2020, we define a first fitness vector of a phenotype $\mathbf{x} \in \mathbb{R}^n$:

$$(\widetilde{m}_1, \widetilde{m}_2)(\mathbf{x}) = \left( x_1, -\sum_{i=2}^{n} \frac{x_i^2}{2} \right) \in \mathbb{R} \times \mathbb{R}_-.$$

As in Hamel; Lavigne; Martin; Roques 2020, the layer-cake formula yields that the density of individuals, having the fitness vector $(\widetilde{m}_1, \widetilde{m}_2)$ in habitat $i \in \{1, 2\}$, is given by:

$$\widetilde{p}_i(t, \widetilde{m}_1, \widetilde{m}_2) = (2|\widetilde{m}_2|)^{(n-1)/2-1} \int_{\mathbb{S}^{n-2}} u_i\Big(t, \widetilde{m}_1, \sqrt{2|\widetilde{m}_2|}\,\sigma\Big)\, \mathrm{d}\sigma, \tag{6.23}$$

for all $(t, \widetilde{m}_1, \widetilde{m}_2) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_-^*$, where $\mathbb{S}^{n-2}$ denotes the unit Euclidean sphere of $\mathbb{R}^{n-1}$. It is easy to see that $(\widetilde{p}_1, \widetilde{p}_2)$ is of class $C^{1,2}(\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_-^*)$. Since $(u_1, u_2)$ is a solution of (6.6), it is straighforward to check that $(\widetilde{p}_1, \widetilde{p}_2)$ is a classic solution

of:

$$\left\{\begin{array}{rl}\partial_t\widetilde{p}_1(t,\widetilde{m}_1,\widetilde{m}_2)= & \frac{\mu^2}{2}\partial_{\widetilde{m}_1\widetilde{m}_1}\widetilde{p}_1(t,\widetilde{m}_1,\widetilde{m}_2)-\mu^2\widetilde{m}_2\partial_{\widetilde{m}_2\widetilde{m}_2}\widetilde{p}_1(t,\widetilde{m}_1,\widetilde{m}_2)\\ & +\mu^2\frac{n-5}{2}\partial_{\widetilde{m}_1}\widetilde{p}_1(t,\widetilde{m}_1,\widetilde{m}_2)+\left[r_{\max}-\frac{\beta^2}{2}-\beta\widetilde{m}_1-\frac{\widetilde{m}_1^2}{2}+\widetilde{m}_2\right]\widetilde{p}_1(t,\widetilde{m}_1,\widetilde{m}_2)\\ & +\delta[\widetilde{p}_2(t,\widetilde{m}_1,\widetilde{m}_2)-\widetilde{p}_1(t,\widetilde{m}_1,\widetilde{m}_2)],\\ \partial_t\widetilde{p}_2(t,\widetilde{m}_1,\widetilde{m}_2)= & \frac{\mu^2}{2}\partial_{\widetilde{m}_1\widetilde{m}_1}\widetilde{p}_2(t,\widetilde{m}_1,\widetilde{m}_2)-\mu^2\widetilde{m}_2\partial_{\widetilde{m}_2\widetilde{m}_2}\widetilde{p}_2(t,\widetilde{m}_1,\widetilde{m}_2)\\ & +\mu^2\frac{n-5}{2}\partial_{\widetilde{m}_1}\widetilde{p}_2(t,\widetilde{m}_1,\widetilde{m}_2)+\left[r_{\max}-\frac{\beta^2}{2}+\beta\widetilde{m}_1-\frac{\widetilde{m}_1^2}{2}+\widetilde{m}_2\right]\widetilde{p}_2(t,\widetilde{m}_1,\widetilde{m}_2)\\ & +\delta[\widetilde{p}_1(t,\widetilde{m}_1,\widetilde{m}_2)-\widetilde{p}_2(t,\widetilde{m}_1,\widetilde{m}_2)],\end{array}\right.$$

for all $(t,\widetilde{m}_1,\widetilde{m}_2)\in\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_-^*$. This first fitness vector lets us to simplify the computation of the density $p_i(t,m_1,m_2)$ of individuals having the fitness vector $(m_1,m_2)$, in habitat $i$. The function:

$$(\widetilde{m}_1,\widetilde{m}_2)\mapsto(m_1,m_2)=\left(-\frac{\beta^2}{2}-\beta\widetilde{m}_1-\frac{\widetilde{m}_1^2}{2}+\widetilde{m}_2,-\frac{\beta^2}{2}+\beta\widetilde{m}_1-\frac{\widetilde{m}_1^2}{2}+\widetilde{m}_2\right),$$

implies that:

$$p_i(t,m_1,m_2)=\frac{1}{2\beta}\,\widetilde{p}_i\left(t,\frac{m_2-m_1}{2\beta},\frac{\beta^2}{2}+\frac{m_1+m_2}{2}+\frac{(m_2-m_1)^2}{8\beta}\right),\qquad(6.24)$$

for all $t\geq 0$, and for all $(m_1,m_2)\in\Gamma$. This relation yields the explicit relation (6.10) between $p_i$ and $u_i$. We can also find an explicit partial differential system satisfied by $(p_1,p_2)$:

$$\left\{\begin{array}{rl}\partial_t p_1(t,m_1,m_2)= & -\mu^2m_1\partial_{m_1m_1}p_1(t,m_1,m_2)-\mu^2\left[m_1+m_2+2\beta^2\right]\partial_{m_1m_2}p_1(t,m_1,m_2)\\ & -\mu^2m_2\partial_{m_2m_2}p_1(t,m_1,m_2)+\mu^2\left(\frac{n}{2}-3\right)\partial_{m_1}p_1(t,m_1,m_2)\\ & +\mu^2\left(\frac{n}{2}-3\right)\partial_{m_2}p_1(t,m_1,m_2)+[r_{\max}+m_1]\,p_1(t,m_1,m_2)\\ & +\delta[p_2(t,m_1,m_2)-p_1(t,m_1,m_2)],\\ \partial_t p_2(t,m_1,m_2)= & -\mu^2m_1\partial_{m_1m_1}p_2(t,m_1,m_2)-\mu^2\left[m_1+m_2+2\beta^2\right]\partial_{m_1m_2}p_2(t,m_1,m_2)\\ & -\mu^2m_2\partial_{m_2m_2}p_2(t,m_1,m_2)+\mu^2\left(\frac{n}{2}-3\right)\partial_{m_1}p_2(t,m_1,m_2)\\ & +\mu^2\left(\frac{n}{2}-3\right)\partial_{m_2}p_2(t,m_1,m_2)+[r_{\max}+m_2]\,p_2(t,m_1,m_2)\\ & +\delta[p_1(t,m_1,m_2)-p_2(t,m_1,m_2)],\end{array}\right.\qquad(6.25)$$

for all $t\geq 0$, and for all $(m_1,m_2)\in\Gamma$.

### 6.5.1.2. Laplace transform of the fitness distribution

The components $m_1$ and $m_2$ are both negative. As $p_1(t,\cdot),p_2(t,\cdot)\in L^1(\Gamma)\cap C(\Gamma)$, by Section 6.5.1.1, the functions $M_1$ and $M_2$ are well defined.

The relations (6.11), (6.23) and (6.24) yields that for all $t>0$, for all $\mathbf{z}=(z_1,z_2)\in\mathbb{R}_+^2$ and $i\in\{1,2\}$:

$$M_i(t,\mathbf{z})=e^{-\beta^2(z_1+z_2)/2}\int_{\mathbb{R}\times\mathbb{R}_-^*}\exp\left[\beta\widetilde{m}_1(z_2-z_1)+(z_1+z_2)\left(-\frac{\widetilde{m}_1^2}{2}+\widetilde{m}_2\right)\right](2|\widetilde{m}_2|)^{(n-1)/2-1}$$
$$\int_{\mathbb{S}^{n-2}}u_i\big(t,\widetilde{m}_1,\sqrt{2|\widetilde{m}_2|}\,\sigma\big)\,\mathrm{d}\sigma\,\mathrm{d}\widetilde{m}_1\,\mathrm{d}\widetilde{m}_2.\qquad(6.26)$$

Hence we have:

$$M_i(t,\mathbf{z}) = \int_{\mathbb{R}^n} \exp\left[z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})\right] u_i(t,\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{6.27}$$

**Continuity of $M_1$ and $M_2$.** Let us set for all $t>0$ and $\mathbf{x}\in\mathbb{R}^n$:

$$\left\{\begin{array}{l} h_1(t,\mathbf{x}) = e^{-\delta t}\left[K_t * u_1^0\right](\mathbf{x})\cosh(\delta t) + e^{-\delta t}\left[K_t * u_2^0\right](\mathbf{x})\sinh(\delta t), \\ h_2(t,\mathbf{x}) = e^{-\delta t}\left[K_t * u_1^0\right](\mathbf{x})\sinh(\delta t) + e^{-\delta t}\left[K_t * u_2^0\right](\mathbf{x})\cosh(\delta t), \end{array}\right. \tag{6.28}$$

with:

$$K_t(\mathbf{x}) = (2\pi\mu^2 t)^{-n/2}\ \exp\left(r_{\max}t - \frac{\|\mathbf{x}\|^2}{2\mu^2 t}\right),$$

and $(h_1(0,\mathbf{x}), h_2(0,\mathbf{x})) = (u_1^0(\mathbf{x}), u_2^0(\mathbf{x}))$. By Theorem 32 of Chapter 5, we know that $u_i(t,\mathbf{x}) \leq h_i(t,\mathbf{x})$, which implies that for $i\in\{1,2\}$, and for all $\mathbf{z}\in\mathbb{R}_+^2$, $T>0$, $t\in(0,T]$ and $\mathbf{x}\in\mathbb{R}^n$:

$$\begin{aligned}
0 &\leq \exp\left[z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})\right] u_i(t,\mathbf{x}) \leq \left[K_t * u_1^0\right](\mathbf{x}) + \left[K_t * u_2^0\right](\mathbf{x}), \\
&\leq (2\pi\mu^2 t)^{-n/2}\int_{\mathbb{R}^n} \exp\left(r_{\max}t - \frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\mu^2 t}\right)\left[u_1^0(\mathbf{y}) + u_2^0(\mathbf{y})\right]\mathrm{d}\mathbf{y}, \\
&\leq e^{r_{\max}T}\pi^{-n/2}\int_{\mathbb{R}^n} e^{-\|\mathbf{y}\|^2/2}\left[u_1^0\left(\mathbf{x}-\sqrt{\mu^2 t}\,\mathbf{y}\right) + u_2^0\left(\mathbf{x}-\sqrt{\mu^2 t}\,\mathbf{y}\right)\right]\mathrm{d}\mathbf{y}, \\
&\leq e^{r_{\max}T}\pi^{-n/2}\int_{\mathbb{R}^n} e^{-\|\mathbf{y}\|^2}\left[g\left(\left\|\mathbf{x}-\mathbf{O}_1-\sqrt{2\mu^2 t}\,\mathbf{y}\right\|\right) + g\left(\left\|\mathbf{x}-\mathbf{O}_2-\sqrt{2\mu_2^2 t}\,\mathbf{y}\right\|\right)\right]\mathrm{d}\mathbf{y},
\end{aligned}$$

thanks to hypothesis (H3). It follows from (H3) that:

$$\begin{aligned}
&\exp\left[z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})\right] u_i(t,\mathbf{x}) \\
&\qquad \leq e^{r_{\max}T}\pi^{-n/2}\left[\int_{\|\mathbf{y}'\|\leq\|\mathbf{x}-\mathbf{O}_1\|/(2\sqrt{2\mu^2 t})} e^{-\|\mathbf{y}'\|^2}\, g(\|\mathbf{x}-\mathbf{O}_1\|/2)\,\mathrm{d}\mathbf{y}'\right. \\
&\qquad\qquad \left. + g(0)\int_{\|\mathbf{y}'\|>\|\mathbf{x}-\mathbf{O}_1\|/(2\sqrt{2\mu^2 t})} e^{-\|\mathbf{y}'\|^2}\,\mathrm{d}\mathbf{y}'\right] \\
&\qquad + e^{r_{\max}T}\pi^{-n/2}\left[\int_{\|\mathbf{y}'\|\leq\|\mathbf{x}-\mathbf{O}_2\|/(2\sqrt{2\mu^2 t})} e^{-\|\mathbf{y}'\|^2}\, g(\|\mathbf{x}-\mathbf{O}_2\|/2)\,\mathrm{d}\mathbf{y}'\right. \\
&\qquad\qquad \left. + g(0)\int_{\|\mathbf{y}'\|>\|\mathbf{x}-\mathbf{O}_2\|/(2\sqrt{2\mu^2 t})} e^{-\|\mathbf{y}'\|^2}\,\mathrm{d}\mathbf{y}'\right], \\
&\qquad \leq e^{r_{\max}T} g(\|\mathbf{x}-\mathbf{O}_1\|/2) + e^{r_{\max}T} g(\|\mathbf{x}-\mathbf{O}_2\|/2) \\
&\qquad\qquad + g(0)e^{r_{\max}T}\pi^{-n/2} I(\mathbf{x}-\mathbf{O}_1) + g(0)e^{r_{\max}T}\pi^{-n/2} I(\mathbf{x}-\mathbf{O}_2),
\end{aligned}$$

where the function $I(\mathbf{x})$ is defined by:

$$I(\mathbf{x}) = \int_{\|\mathbf{y}'\|>\|\mathbf{x}\|/(2\sqrt{2\mu^2 T})} e^{-\|\mathbf{y}'\|^2}\,\mathrm{d}\mathbf{y}'.$$

By (H3), the right-hand side quantity of the last inequality belongs to $L^1(\mathbb{R}^n)$, and is independent of $\mathbf{z} \in \mathbb{R}^n$ and $t \in (0,T]$. Thus the equality (6.27) and Lebesgue's dominated convergence theorem imply that the functions $M_1$ and $M_2$ are continuous in $\mathbb{R}_+ \times \mathbb{R}_+^2$.

**First derivatives with respect to $z_1$ and $z_2$.** In this paragraph, we check that $M_1$ (and similarly $M_2$) is $C^{0,1}(\mathbb{R}_+ \times \mathbb{R}_+^2)$, *i.e.*, $\partial_{z_i} M_1$ exists and is continuous in $\mathbb{R}_+ \times \mathbb{R}_+^2$, for $i \in \{1,2\}$. As $p_1(t,\cdot)$ is a finite measure on $\Gamma$ for any time $t \geq 0$, by same arguments as in Hamel; Lavigne; Martin; Roques 2020, one infers that $\partial_{z_i} M_1(t,\mathbf{z})$ exists for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}_+^n$ and $i \in \{1,2\}$, with:

$$\partial_{z_i} M_i(t,\mathbf{z}) = \int_\Gamma m_i \, e^{z_1 m_1 + z_2 m_2} p_1(t,m_1,m_2) \, \mathrm{d}m_1 \mathrm{d}m_2, \tag{6.29}$$

which is equivalent, thanks to (6.11), (6.23) and (6.24):

$$\partial_{z_i} M_1(t,\mathbf{z}) = \int_{\mathbb{R}^n} m_i(\mathbf{x}) \exp\left[z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})\right] u_1(t,\mathbf{x}) \, \mathrm{d}\mathbf{x}. \tag{6.30}$$

The arguments for the continuity of $M_1$ and $M_2$ yield that for $i \in \{1,2\}$, for all $\mathbf{z} \in \mathbb{R}_+^2$, $T > 0$, $t \in (0,T]$ and $\mathbf{x} \in \mathbb{R}^n$:

$$\begin{aligned} 0 &\leq |m_i(\mathbf{x})| \, \exp\left[z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})\right] u_1(t,\mathbf{x}) \\ &\leq \frac{\|\mathbf{x} - \mathbf{O}_i\|^2}{2} e^{r_{\max} T} \left[g(\|\mathbf{x} - \mathbf{O}_1\|/2) + g(\|\mathbf{x} - \mathbf{O}_2\|/2)\right] \\ &\qquad + \frac{\|\mathbf{x} - \mathbf{O}_i\|^2}{2} g(0) e^{r_{\max} T} \pi^{-n/2} \left[I(\mathbf{x} - \mathbf{O}_1) + I(\mathbf{x} - \mathbf{O}_2)\right], \end{aligned}$$

where the function $I(\mathbf{x})$ is defined by:

$$I(\mathbf{x}) = \int_{\|\mathbf{y}'\| > \|\mathbf{x}\|/(2\sqrt{2\mu^2 T})} e^{-\|\mathbf{y}'\|^2} \, \mathrm{d}\mathbf{y}'.$$

The right-hand side quantity belongs in $L^1(\mathbb{R}^n)$ and is independent of $\mathbf{z} \in \mathbb{R}_+^2$ and of $t \in (0,T]$. By Lebesgue's dominated convergence theorem, we conclude that $\partial_{z_i} M_1$ is continuous in $\mathbb{R}_+ \times \mathbb{R}_+^2$.

**Differentiation of $M_1$ and $M_2$ with respect to $t$.** By (6.6), we already know that:

$$|\partial_t u_i(t,\mathbf{x})| \leq \frac{\mu^2}{2} |\Delta u_i(t,\mathbf{x})| + |m_i(\mathbf{x})| u_i(t,\mathbf{x}) + \delta \left[u_1(t,\mathbf{x}) + u_2(t,\mathbf{x})\right], \tag{6.31}$$

for all $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^n$. Fix $T > 0$. As $u_i(t,\mathbf{x}) \leq h_i(t,\mathbf{x})$, for all $\mathbf{x} \in \mathbb{R}^n$ and $t \in [0,T]$, Eq. 6.28 yields a constant $S_0 > 0$ such that:

$$u_1(t,\mathbf{x}), u_2(t,\mathbf{x}) \leq S_0 \left( \|u_1^0\|_{C^{2+\alpha}(\mathbb{R}^n)} + \|u_2^0\|_{C^{2+\alpha}(\mathbb{R}^n)} \right).$$

By the same arguments as in Hamel; Lavigne; Martin; Roques 2020, there exists so a constant $S \geq S_0$ for all $t \in [0,T]$ and $\mathbf{x} \in \mathbb{R}^n$, we have:

$$|\Delta u_i(t,\mathbf{x})| \leq S \left( \|u_1^0\|_{C^{2+\alpha}(\mathbb{R}^n)} + \|u_2^0\|_{C^{2+\alpha}(\mathbb{R}^n)} \right). \tag{6.32}$$

Let us bound the function $(t,\mathbf{x}) \mapsto |m_i(\mathbf{x})| u_i(t,\mathbf{x})$ in $[0,T] \times \mathbb{R}^n$. By continuity of this function, we have to show its boundedness in $(0,T] \times \mathbb{R}^n$. We easily check, thanks to the functions $h_1$ and $h_2$ defined in (6.28), that:

$$|m_i(\mathbf{x}) u_i(t,\mathbf{x})| \leq \int_{\mathbb{R}^n} \frac{\|\mathbf{x} - \mathbf{O}_i\|^2}{2\pi^{n/2}} \left[ u_1^0(\mathbf{x} - \sqrt{2\mu^2 t}\,\mathbf{y}) + u_2^0(\mathbf{x} - \sqrt{2\mu^2 t}\,\mathbf{y}) \right] e^{-\|y\|^2} \,\mathrm{d}\mathbf{y}, \tag{6.33}$$

for all $(t,\mathbf{x}) \in (0,T] \times \mathbb{R}^n$. This inequality yields that:

$$\begin{aligned}
|m_1(\mathbf{x}) u_1(t,\mathbf{x})| &\leq \pi^{-n/2} \int_{\mathbb{R}^n} \left[ \|\mathbf{x} - \mathbf{O}_1 - \sqrt{2\mu^2 t}\mathbf{y}\|^2 + 2\mu^2 t \|\mathbf{y}\|^2 \right] g\left( \|\mathbf{x} - \mathbf{O}_1 - \sqrt{2\mu^2 t}\mathbf{y}\| \right) e^{-\|y\|^2} \,\mathrm{d}\mathbf{y} \\
&\quad + \pi^{-n/2} \int_{\mathbb{R}^n} \left[ \|\mathbf{x} - \mathbf{O}_2 - \sqrt{2\mu^2 t}\mathbf{y}\|^2 + 8\beta^2 + 4\mu^2 t \|\mathbf{y}\|^2 \right] g\left( \|\mathbf{x} - \mathbf{O}_2 - \sqrt{2\mu^2 t}\mathbf{y}\| \right) e^{-\|y\|^2} \,\mathrm{d}\mathbf{y}, \\
&\leq 2 \left[ \|m_1 g(\|\cdot - \mathbf{O}_1\|)\|_{L^\infty(\mathbb{R}^n)} + \|m_2 g(\|\cdot - \mathbf{O}_2\|)\|_{L^\infty(\mathbb{R}^n)} + 4\beta^2 g(0) \right] + 8\pi^{-n/2} \mu^2 T g(0) \int_{\mathbb{R}^n} \|y\|^2 e^{-\|\mathbf{y}\|^2} \,\mathrm{d}\mathbf{y}.
\end{aligned}$$

We have proven that the functions $m_1 u_1$ (and similarly $m_2 u_2$) is bounded in $[0,T] \times \mathbb{R}^n$ for any time $T > 0$. Thus the function $\partial_t u_1$ (and $\partial_t N_2$) by (6.31)–(6.32) is also bounded. The Lebesgue's convergence theorem yields that the functions $M_1$ and $M_2$ are differentiable with respect to $t$ in $\mathbb{R}_+ \times \mathbb{R}_+^2$, with:

$$\begin{cases}
\partial_t M_1(t,\mathbf{z}) = \displaystyle\int_{\mathbb{R}^n} e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} \partial_t u_1(t,\mathbf{x}) \,\mathrm{d}\mathbf{x}, \\
\partial_t M_2(t,\mathbf{z}) = \displaystyle\int_{\mathbb{R}^n} e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} \partial_t u_2(t,\mathbf{x}) \,\mathrm{d}\mathbf{x},
\end{cases} \tag{6.34}$$

for all $(t,\mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}_+^2$.

**Partial differential equation satisfied by M.** Finding an equation satisfied by $(M_1, M_2)$ lets us to end the proof of Proposition 37. Fix $T > 0$, $t \in (0,T)$ and $\mathbf{z} \in (\mathbb{R}_+^*)^2$. Thanks to (6.6), we have:

$$\begin{aligned}
e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} \partial_t u_1(t,\mathbf{m}) &= -\frac{\mu^2}{2} e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} \Delta u_1(t,\mathbf{x}) \\
&\quad + \left[ r_{\max} + m_1(\mathbf{x}) \right] e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} u_1(t,\mathbf{x}) \\
&\quad + \delta e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} \left[ u_2(t,\mathbf{x}) - u_1(t,\mathbf{x}) \right],
\end{aligned} \tag{6.35}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Same arguments as in Hamel; Lavigne; Martin; Roques 2020 implies that the following integral converges:

$$\int_{\mathbb{R}^n} e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} \Delta u_1(t, \mathbf{x}) \, \mathrm{d}\mathbf{x} < \infty,$$

and that:

$$\begin{aligned}
&\int_{\mathbb{R}^n} e^{z_1 m_1(\mathbf{x}) + z_2 m_2(\mathbf{x})} \Delta u_1(t, \mathbf{x}) \, \mathrm{d}\mathbf{x} \\
&= -2z_1(z_1 + z_2)\partial_{z_1} M_1(t, \mathbf{z}) - 2z_2(z_1 + z_2)\partial_{z_2} M_1(t, \mathbf{z}) - 2\left[2\beta^2 z_1 z_2 + n(z_1 + z_2)\right] M_1(t, \mathbf{z}),
\end{aligned}$$

thanks to (6.30). Integrating (6.35) over $\mathbb{R}^n$, and using (6.30), (6.34) and the previous formula, the function $(M_1, M_2)$ satisfies the system given by Proposition 37.

#### 6.5.1.3. Lemma

For the sake of simplicity, we need some results about the Generalized Hypergeometric Function defined by:

$${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; x) = \sum_{n=0}^{+\infty} \frac{(a_1)_n \ldots (a_p)_n}{(b_1)_n \ldots (b_q)_n} \frac{x^n}{n!},$$

for $x \in [-1, 1[$ and $p, q \in \mathbb{N}$. We recall the Pocchamer symbol:

$$(x)_n = x(x + 1) \ldots (x + n - 1),$$

for $x \in \mathbb{R}$ and $n \in \mathbb{N}$.

**Lemma 7.** *Let $k \in \mathbb{N}$ and $(a,b) \in \mathbb{R}^2$, such that $b \geq 0$ and that:*

$$q := \frac{a+kb}{2b} \geq 0.$$

*Thus we have for all $t \geq 0$:*

$$I_k(a,b,t) := \int_0^t \frac{e^{-as}}{\cosh^k(bs)}\mathrm{d}s = \frac{2^k}{a+kb}\left[{}_2F_1\left(k,q;q+1;-1\right) - e^{-2bqt}\,{}_2F_1\left(k,q;q+1;-e^{-2bt}\right)\right],$$

$$\begin{aligned} J_k(a,b,t) &:= \int_0^t s\frac{e^{-as}}{\cosh^k(bs)}\mathrm{d}s, \\ &= \frac{2^k}{(a+kb)^2}\,{}_3F_2[k,q,q;q+1,q+1;-1] - \frac{2^k}{a+kb}\,te^{-2bqt}\,{}_2F_1\left[k,q;q+1;-e^{-2bt}\right] \\ &\quad - \frac{2^k}{(a+kb)^2}\,e^{-2bqt}\,{}_3F_2\left[k,q,q;q+1,q+1;-e^{-2bt}\right], \end{aligned}$$

*and:*

$$\begin{aligned} L_k(a,b,t) &:= \int_0^t s^2\frac{e^{-as}}{\cosh^k(bs)}\mathrm{d}s, \\ &= \frac{2^{k+1}}{(a+kb)^3}\,{}_4F_3[k,q,q,q;q+1,q+1,q+1;-1] \\ &\quad - \frac{2^{k+1}}{(a+kb)^3}\,e^{-2bqt}\,{}_4F_3\left[k,q,q,q;q+1,q+1,q+1;-e^{-2bt}\right] \\ &\quad - \frac{2^{k+1}}{(a+kb)^2}\,te^{-2bqt}\,{}_3F_2\left[k,q,q;q+1,q+1;-e^{-2bt}\right] \\ &\quad - \frac{2^k}{a+kb}\,t^2e^{-2bqt}\,{}_2F_1\left[k,q;q+1;-e^{-2bt}\right]. \end{aligned}$$

*Proof.* We begin to study the first integral $I_k(a,b,t)$. We have:

$$I_k(a,b,t) = \int_0^t \frac{e^{-as}}{\cosh^k(bs)}\mathrm{d}s = 2^k\int_0^t \frac{e^{-(a+kb)s}}{(1+e^{-2bs})^k}\mathrm{d}s = 2^k\sum_{n=0}^{+\infty}(-1)^n\frac{(k)_n}{n!}\int_0^t e^{-(a+kb+2bn)s}\mathrm{d}s.$$

Thus we get:

$$\begin{aligned} I_k(a,b,t) &= 2^k\sum_{n=0}^{+\infty}(-1)^n\frac{(k)_n}{a+kb+2bn}\frac{1-e^{-(a+kb+2bn)t}}{n!} = \frac{2^k}{2b}\sum_{n=0}^{+\infty}(-1)^n\frac{(k)_n}{q+n}\frac{1-e^{-(a+kb+2bn)t}}{n!}, \\ &= \frac{2^k}{a+kb}\sum_{n=0}^{+\infty}\frac{(k)_n(q)_n}{(q+1)_n}\frac{(-1)^n}{n!} - \frac{2^k}{a+kb}\,e^{-(a+kb)t}\sum_{n=0}^{+\infty}\frac{(k)_n(q)_n}{(q+1)_n}\frac{(-e^{-2bt})^n}{n!}, \end{aligned}$$

as $q(q+1)_n = (q+n)(q)_n$. By the definition of the Hypergeometric Function ${}_2F_1$, we have proven the formula for $I_k(a,b,t)$.

By the same kind of computations, we find the formula for $J_k(a, b, t)$ and $L_k(a, b, t)$, which ends the proof. □

#### 6.5.1.4. Approximation

In this section, we develop the computation of the formula of M, the mean fitness $\overline{m}_1$ (and so $\overline{m}_2$), but also the migration load.

**Formula of $\mathrm{M}(t, \mathbf{z})$ and migrational load.** In this section, we determine a formula for $M_1(t, \mathbf{z})$, thanks to the following theorem:

**Theorem 40.**
*Let $\mu > 0$ a real constant and $\widetilde{\beta}$ a given real function. We define the functions:*

$$\beta_1(\mathbf{z}) = -\mu^2(z_1+z_2)z_1, \quad \textit{and} \quad \beta_2(\mathbf{z}) = -\mu^2(z_1+z_2)z_2,$$

*for all $\mathbf{z} = (z_1, z_2) \in \mathbb{R}^2$. The nonnegative solution $F$ of the problem:*

$$\begin{aligned}\partial_t F(t,\mathbf{z}) = [1+\beta_1(\mathbf{z})+\delta(z_2-z_1)]\,\partial_{z_1}F(t,\mathbf{z}) \\ + [\beta_2(\mathbf{z})+\delta(z_1-z_2)]\,\partial_{z_2}F(t,\mathbf{z}) + \widetilde{\beta}(\mathbf{z})F(t,\mathbf{z}),\end{aligned} \tag{6.36}$$

*is given by:*

$$F(t,\mathbf{z}) = F\left[0,\mathbf{y}(\mu,t,\mathbf{z})\right]\,\exp\left[\int_0^t \widetilde{\beta}(\mathbf{y}(\mu,s,\mathbf{z}))\,ds\right], \tag{6.37}$$

*with $\mathbf{y}(\mu,t,\mathbf{z}) = (y_1(\mu,t,\mathbf{z}), y_2(\mu,t,\mathbf{z}))$, where:*

- *if $\mu = 2\delta$,*

$$\begin{aligned}y_1(\mu,t,\mathbf{z}) = \frac{\tanh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]}{2\mu} + \frac{(z_1-z_2)\cosh[\operatorname{atanh}(\mu(z_1+z_2))]}{2\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]}e^{-\mu t} \\ + \frac{\sinh[\mu t]\exp[\operatorname{atanh}(\mu(z_1+z_2))]}{4\mu\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]} + \frac{t\exp[-\mu t-\operatorname{atanh}(\mu(z_1+z_2))]}{4\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]},\end{aligned}$$

*and:*

$$\begin{aligned}y_2(\mu,t,\mathbf{z}) = \frac{\tanh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]}{2\mu} - \frac{(z_1-z_2)\cosh[\operatorname{atanh}(\mu(z_1+z_2))]}{2\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]}e^{-\mu t} \\ - \frac{\sinh[\mu t]\exp[\operatorname{atanh}(\mu(z_1+z_2))]}{4\mu\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]} - \frac{t\exp[-\mu t-\operatorname{atanh}(\mu(z_1+z_2))]}{4\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]};\end{aligned}$$

- *if $\mu \neq 2\delta$,*

$$\begin{aligned}y_1(\mu,t,\mathbf{z}) = \frac{\delta}{4\delta^2-\mu^2} + \frac{2\delta^2-\mu^2}{\mu(4\delta^2-\mu^2)}\tanh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))] \\ + \frac{2\delta^2(z_1-z_2)-\delta+\mu^2 z_2}{4\delta^2-\mu^2}\,\frac{\cosh[\operatorname{atanh}(\mu(z_1+z_2))]}{\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]}\,e^{-2\delta t},\end{aligned}$$

*and:*

$$\begin{aligned}y_2(\mu,t,\mathbf{z}) = -\frac{\delta}{4\delta^2-\mu^2} + \frac{2\delta^2}{\mu(4\delta^2-\mu^2)}\tanh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))] \\ - \frac{2\delta^2(z_1-z_2)-\delta+\mu^2 z_2}{4\delta^2-\mu^2}\,\frac{\cosh[\operatorname{atanh}(\mu(z_1+z_2))]}{\cosh[\mu t+\operatorname{atanh}(\mu(z_1+z_2))]}\,e^{-2\delta t}.\end{aligned}$$

*Proof.* It is immediate to check that the function $F$, given by Theorem (40), is a

classical $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^2)$ solution of (6.36). Now let $F_1$ and $F_2$ two nonnegative $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^2)$ solution of (6.36) and denote the $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^2)$ function $Q = F_1/F_2$, which satisfies:

$$\partial_t Q(t, \mathbf{z}) = [1 + \beta_1(\mathbf{z}) + \delta(z_2 - z_1)]\, \partial_{z_1} Q(t, \mathbf{z}) + [\beta_2(\mathbf{z}) + \delta(z_1 - z_2)]\, \partial_{z_2} Q(t, \mathbf{z}),$$

for all $t > 0$ and $\mathbf{z} \in \mathbb{R}_+^2$, and $Q(0, \mathbf{z}) = 1$, for all $\mathbf{z} \in \mathbb{R}_+^2$. Fix any time $t > 0$ and parameter $\mathbf{z} \in \mathbb{R}_+^2$. Consider now the auxiliary $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+^2)$ function $R(s) = Q(t - s, \mathbf{y}(\mu, s, \mathbf{z}))$, with $\mathbf{y}$ defined in Theorem 40. This function $R$ satisfies:

$$\begin{aligned} R'(s) &= -\partial_t Q(t - s, \mathbf{y}(\mu, s, \mathbf{z})) + \partial_t y_1(\mu, s, \mathbf{z}) \partial_{z_1} Q(t - s, \mathbf{y}(\mu, s, \mathbf{z})) \\ &\qquad + \partial_t y_2(\mu, s, \mathbf{z}) \partial_{z_2} Q(t - s, \mathbf{y}(\mu, s, \mathbf{z})), \\ &= -\partial_t Q(t - s, \mathbf{y}(\mu, s, \mathbf{z})) + [1 + \beta_1(\mathbf{y}(\mu, t, \mathbf{z}))] \partial_{z_1} Q(t - s, \mathbf{y}(\mu, s, \mathbf{z})) \\ &\qquad + \beta_2(\mathbf{y}(\mu, s, \mathbf{z})) \partial_{z_2} Q(t - s, \mathbf{y}(\mu, s, \mathbf{z})) = 0. \end{aligned}$$

Thus $R$ is constant and $R(t) = Q(0, \mathbf{y}(\mu, t, \mathbf{z})) = 1$. This yields that $R \equiv 1$ and so $F_1 = F_2$, which ends the proof. □

To have an explicit formula for $M_1$ (and so $M_2$), we can apply the previous theorem with the function:

$$\widetilde{\beta}(z_1, z_2) = r_{\max} - \mu^2 \left[ m_D z_1 z_2 + \frac{n}{2}(z_1 + z_2) \right].$$

However, we have to compute the integral $\int_0^t \widetilde{\beta}(\mathbf{y}(\mu, s, \mathbf{z}))\, \mathrm{d}s$. Thanks to this theorem, we will be able to study the evolution of the population.

**Explosion vs extinction.** Let us turn to compute the integral:

$$\begin{aligned}
\int_0^t \widetilde{\beta}(\mathbf{y}(\mu,s,\mathbf{z}))\,\mathrm{d}s &= r_{\max}t - \mu^2 m_D \int_0^t y_1(\mu,s,\mathbf{z})y_2(\mu,s,\mathbf{z})\mathrm{d}s - \frac{\mu^2 n}{2}\int_0^t [y_1(\mu,s,\mathbf{z})+y_2(\mu,s,\mathbf{z})]\,\mathrm{d}s,\\
&= r_{\max}t - \mu^2 m_D \int_0^t \frac{[y_1(\mu,s,\mathbf{z})+y_2(\mu,s,\mathbf{z})]^2 - [y_1(\mu,s,\mathbf{z})-y_2(\mu,s,\mathbf{z})]^2}{4}\,\mathrm{d}s\\
&\qquad - \frac{\mu n}{2}\int_0^t \tanh[\mu s + \operatorname{atanh}(\mu(z_1+z_2))]\,\mathrm{d}s,\\
&= r_{\max}t - \frac{m_D}{4}\int_0^t \tanh^2[\mu s + \operatorname{atanh}(\mu(z_1+z_2))]\,\mathrm{d}s\\
&\qquad + \frac{\mu^2 m_D}{4}\int_0^t [y_1(\mu,s,\mathbf{z})-y_2(\mu,s,\mathbf{z})]^2\mathrm{d}s\\
&\qquad - \frac{n}{2}\log\left[\frac{\cosh[\mu t + \operatorname{atanh}(\mu(z_1+z_2))]}{\cosh[\operatorname{atanh}(\mu(z_1+z_2))]}\right],\\
&= bt - \frac{m_D}{4}t + \frac{m_D}{4\mu}\left[\tanh[\mu t + \operatorname{atanh}(\mu(z_1+z_2))] - \mu(z_1+z_2)\right]\\
&\qquad + \frac{\mu^2 m_D}{4}\int_0^t [y_1(\mu,s,\mathbf{z})-y_2(\mu,s,\mathbf{z})]^2\mathrm{d}s\\
&\qquad - \frac{n}{2}\log\left[\frac{\cosh[\mu t + \operatorname{atanh}(\mu(z_1+z_2))]}{\cosh[\operatorname{atanh}(\mu(z_1+z_2))]}\right].
\end{aligned}$$

Thus we have:

$$\begin{aligned}
M_1(t,\mathbf{z}) = M_1[0,\mathbf{y}(\mu,t,\mathbf{z})]\left[\frac{\cosh[\operatorname{atanh}(\mu(z_1+z_2))]}{\cosh[\mu t + \operatorname{atanh}(\mu(z_1+z_2))]}\right]^{n/2}\exp\left[\left(r_{\max}-\frac{m_D}{4}\right)t - \frac{m_D}{4}(z_1+z_2)\right]\\
\exp\left[\frac{m_D}{4\mu}\tanh[\mu t + \operatorname{atanh}(\mu(z_1+z_2))] + \frac{\mu^2 m_D}{4}\,\mathcal{I}(\mu,t,\mathbf{z})\right],
\end{aligned}$$

with the integral:

$$\mathcal{I}(\mu,t,\mathbf{z}) := \int_0^t [y_1(\mu,s,\mathbf{z})-y_2(\mu,s,\mathbf{z})]^2\mathrm{d}s.$$

It seems impossible to explain $\mathcal{I}(\mu,t,\mathbf{z})$ with simple functions, for all values of $\mathbf{z}$, and that is the reason that we only focus on $\mathbf{z}=0$.

In the special case $\mu = 2\delta$, we can check that for all $t>0$:

$$\begin{aligned}
\mathcal{I}(\mu,t,0) = \frac{1}{16\mu^2}[I_2(2\mu,\mu,t)+I_2(-2\mu,\mu,t)] + \frac{1}{4}L_2(2\mu,\mu,t) - \frac{1}{8\mu^2}\tanh(\mu t)\\
- \frac{1}{4\mu}J_2(2\mu,\mu,t) + \frac{1}{4\mu}J_2(0,\mu,t),
\end{aligned}$$

with the integrals $I_k(a,b,t)$, $J_k(a,b,t)$ and $L_k(a,b,t)$ defined in Lemma 7. We remark that $I_2(-2\mu,\mu,t)$ is the only right-hand term which diverges as $t\to+\infty$, with $I_2(-2\mu,\mu,t) \underset{t\to+\infty}{\sim} 4t$. Thus Lemma 7 yields that the population size $N_1(t)$

satisfies for all $t > 0$:

$$N_1(t) \underset{t\to+\infty}{\sim} M_1[0, \mathbf{y}(\mu, t, 0)] \cosh^{-n/2}(\mu t) \exp\left[\left(r_{\max} - \frac{m_D}{4} + \frac{m_D}{16}\right) t + f(t)\right],$$

with $f$ a convergent function. Thus the population goes to extinction if and only if:

$$r_{\max} - \frac{3m_D}{16} - \frac{\mu n}{2} < 0,$$

which is equivalent to $m_D > m_{D,crit}(r_{\max}, 2\delta, n, \delta) = \frac{16}{3}\left(r_{\max} - \frac{\mu n}{2}\right)$.

Let us turn to the case when $\mu \neq 2\delta$. The reader can check that:

$$\begin{aligned}\mathcal{I}(\mu, t, 0) = \frac{4\delta^2}{(4\delta^2-\mu^2)^2} I_2(4\delta, \mu, t) + \frac{4\delta^2}{(4\delta^2-\mu^2)^2} t + \frac{\mu^2}{(4\delta^2-\mu^2)^2}\left[t - \frac{\tanh(\mu t)}{\mu}\right] \\ - \frac{8\delta^2}{(4\delta^2-\mu^2)^2} I_1(2\delta, \mu, t) + \frac{2\delta\mu}{(4\delta^2-\mu^2)^2}\left[I_2(2\delta-\mu, \mu, t) - I_2(2\delta+\mu, \mu, t)\right] \\ - \frac{4\delta}{(4\delta^2-\mu^2)^2} \log\left[\cosh(\mu t)\right],\end{aligned}$$

with $I_k(a, b, t)$ the integrals defined in Lemma 7, for all $t > 0$. This lemma implies that the population size $N_1(t)$ satisfies:

$$N_1(t) = M_1[0, \mathbf{y}(\mu, t, 0)] \cosh^{-\gamma}(\mu t) \exp\left[\left(r_{\max} - \frac{m_D}{4} + \frac{4\delta^2+\mu^2}{(4\delta^2-\mu^2)^2} \frac{\mu^2 m_D}{4}\right) t + f(t)\right],$$

with $f$ a convergent function and:

$$\gamma = \frac{n}{2} + \frac{\delta\mu^2 m_D}{(4\delta^2-\mu^2)^2}.$$

Thus $N_1$ diverges if and only if:

$$r_{\max} - \frac{m_D}{4} + \frac{4\delta^2+\mu^2}{(4\delta^2-\mu^2)^2} \frac{\mu^2 m_D}{4} - \gamma\mu > 0,$$

which is equivalent to:

$$m_D < \left[r_{\max} - \frac{\mu n}{2}\right] \frac{(2\delta+\mu)^2}{\delta(\delta+\mu)}.$$

This last inequality implies the existence of $\delta^*$.

**Evolution of the fitness.** In this paragraph, we study the evolution of the fitness with respect to the time $t \geq 0$. As $\mu > 0$, we get:

$$\overline{m}_1(t) = \partial_{z_1} y_1(\mu,t,0)\frac{\partial_{z_1} M_1[0,\mathbf{y}(\mu,t,0)]}{M_1[0,\mathbf{y}(\mu,t,0)]} + \partial_{z_1} y_2(\mu,t,0)\frac{\partial_{z_2} M_1[0,\mathbf{y}(\mu,t,0)]}{M_1[0,\mathbf{y}(\mu,t,0)]} - \frac{\mu n}{2}\tanh(\mu t) - \frac{m_D}{4} + \frac{m_D}{4\cosh^2(\mu t)} + \frac{\mu^2 m_D}{4}\partial_{z_1}\mathcal{I}(\mu,t,0). \quad (6.38)$$

If $\mu = 2\delta$, the definition of $\mathbf{y}$ given by Theorem 40 implies that:

$$\partial_{z_1} y_1(\mu,t,0) = \frac{3}{4}[1-\tanh^2(\mu t)] + \frac{e^{-\mu t}}{4\cosh(\mu t)} - \frac{\mu t e^{-\mu t}}{4\cosh(\mu t)}[1+\tanh(\mu t)],$$

and,

$$\partial_{z_1} y_2(\mu,t,0) = \frac{1}{4}[1-\tanh^2(\mu t)] - \frac{e^{-\mu t}}{4\cosh(\mu t)} + \frac{\mu t e^{-\mu t}}{4\cosh(\mu t)}[1+\tanh(\mu t)].$$

We fastly check that these two quantities converges to 0, as $t \to +\infty$. Lemma 7 let us to study the term $\partial_{z_1}\mathcal{I}$:

$$\begin{aligned}\partial_{z_1}\mathcal{I}(\mu,t,0) = &-\frac{3}{8\mu}I_2(2\mu,\mu,t) - \frac{1}{8\mu}\int_0^t \frac{e^{-2\mu s}\tanh(\mu s)}{\cosh^2(\mu s)}\mathrm{d}s - \frac{1}{8\mu^2}\tanh(\mu t)\\ &+ \frac{1}{16\mu^2}\left[1-\frac{1}{\cosh^2(\mu t)}\right] + \frac{1}{4}J_2(2\mu,\mu,t) + \frac{1}{4}\int_0^t \frac{s e^{-2\mu s}\tanh(\mu s)}{\cosh^2(\mu s)}\mathrm{d}s\\ &+ \frac{3}{8\mu^2}\tanh(\mu t) + \frac{1}{16\mu^2}\left[1-\frac{1}{\cosh^2(\mu t)}\right] + \frac{1}{8\mu}\int_0^t \frac{e^{2\mu s}[1-\tanh(\mu s)]}{\cosh^2(\mu s)}\mathrm{d}s\\ &- \frac{1}{4}\int_0^t \frac{s[1+\tanh(\mu s)]}{\cosh^2(\mu s)}\mathrm{d}s + \frac{3}{4}J_2(2\mu,\mu,t) + \frac{1}{4}\int_0^t \frac{s e^{-2\mu s}\tanh(\mu s)}{\cosh^2(\mu s)}\mathrm{d}s\\ &+ \frac{1}{4}\int_0^t \frac{s[1-\tanh(\mu s)]}{\cosh^2(\mu s)}\mathrm{d}s - \frac{\mu}{2}L_2(2\mu,\mu,t) - \frac{\mu}{2}\int_0^t \frac{s^2 e^{-2\mu s}\tanh(\mu s)}{\cosh^2(\mu s)}\mathrm{d}s.\end{aligned}$$

First, we remark that:

$$\int_0^t \frac{e^{2\mu s}[1-\tanh(\mu s)]}{\cosh^2(\mu s)}\mathrm{d}s = \int_0^t \frac{e^{2\mu s}[\cosh(\mu s)-\sinh(\mu s)]}{\cosh^3(\mu s)}\mathrm{d}s = \int_0^t \frac{e^{\mu s}}{\cosh^3(\mu s)}\mathrm{d}s = I_3(-\mu,\mu,t),$$

which converges as $t \to \infty$. Integrating by part, we notice that:

$$\int_0^t \frac{s}{\cosh^2(\mu s)}\mathrm{d}s = \frac{t\tanh(\mu t)}{\mu} - \frac{1}{\mu^2}\log\cosh(\mu t),$$

$$\text{and} \quad \int_0^t \frac{s\tanh(\mu s)}{\cosh^2(\mu s)}\mathrm{d}s = -\frac{t}{2\mu\cosh^2(\mu t)} + \frac{\tanh(\mu t)}{2\mu^2}.$$

Thus we have:

$$\partial_{z_1}\mathcal{I}(\mu,t,0) = -\frac{3}{8\mu}I_2(2\mu,\mu,t) + \frac{1}{16\mu}[I_3(3\mu,\mu,t) - I_3(\mu,\mu,t)] + \frac{1}{8\mu}I_3(-\mu,\mu,t)$$
$$+ J_2(2\mu,\mu,t) + \frac{1}{4}[J_3(\mu,\mu,t) - J_3(3\mu,\mu,t)] - \frac{\mu}{2}L_2(2\mu,\mu,t) + \frac{\mu}{4}[L_3(3\mu,\mu,t) - L_3(\mu,\mu,t)]$$
$$+ \frac{1}{8\mu^2}\left[1 - \frac{1}{\cosh^2(\mu t)}\right] + \frac{t}{4\mu\cosh^2(\mu t)}.$$

This expression, combined with Lemma 7 and (6.38), yields (6.20).

To end the proof of Theorem 38, we have to develop the computations, when $\mu \neq 2\delta$. First, we have, from Theorem 40:

$$\partial_{z_1} y_1(\mu,t,0) = \frac{2\delta^2-\mu^2}{(4\delta^2-\mu^2)\cosh^2(\mu t)} + \frac{2\delta}{4\delta^2-\mu^2}\frac{e^{-2\delta t}}{\cosh(\mu t)} + \frac{\delta}{4\delta^2-\mu^2}\frac{\tanh(\mu t)}{\cosh(\mu t)}e^{-2\delta t},$$

and:

$$\partial_{z_1} y_2(\mu,t,0) = \frac{2\delta^2}{(4\delta^2-\mu^2)\cosh^2(\mu t)} - \frac{2\delta}{4\delta^2-\mu^2}\frac{e^{-2\delta t}}{\cosh(\mu t)} - \frac{\delta}{4\delta^2-\mu^2}\frac{\tanh(\mu t)}{\cosh(\mu t)}e^{-2\delta t}.$$

Thus these two quantities converges to 0, as $t \to +\infty$. Moreover, we can check that:

$$\partial_{z_1}\mathcal{I}(\mu,t,0) = -\frac{16\delta^3}{(4\delta^2-\mu^2)^2}I_2(4\delta,\mu,t) - \frac{4\delta^2\mu}{(4\delta^2-\mu^2)^2}[I_3(4\delta-\mu,\mu,t) - I_3(4\delta+\mu,\mu,t)]$$
$$+ \frac{8\delta\mu^2}{(4\delta^2-\mu^2)^2}I_3(2\delta,\mu,t) + \frac{4\delta}{4\delta^2-\mu^2}I_1(2\delta,\mu,t)$$
$$- \frac{4\delta\mu\tanh(\mu t)}{(4\delta^2-\mu^2)^2} + \frac{\mu^2}{(4\delta^2-\mu^2)^2}\left[1 - \frac{1}{\cosh^2(\mu t)}\right],$$

with the integrals $I_k(a,b,t)$ defined in Lemma 7. This Lemma implies that as $t \to +\infty$, $\partial_{z_1}\mathcal{I}(\mu,t,0)$ converges.

**Local adaptation coefficient.** The local adaptation in a given habitat, *e.g.*,the firt habitat, is measured in this paper by the quantity:

$$LA_1(r_{\max},\mu,n,\delta,t) = \frac{\partial_{z_1}M_1(t,0)}{M_1(t,0)} - \frac{\partial_{z_2}M_1(t,0)}{M_1(t,0)}.$$

We remark that for all $t > 0$:

$$\begin{aligned} LA_1(r_{\max}, \mu, n, \delta, t) = [\partial_{z_1} y_1(\mu, t, 0) - \partial_{z_2} y_1(\mu, t, 0)] \frac{\partial_{z_1} M_1[0, \mathbf{y}(\mu, t, 0)]}{M_1[0, \mathbf{y}(\mu, t, 0)]} \\ + [\partial_{z_1} y_2(\mu, t, 0) - \partial_{z_2} y_2(\mu, t, 0)] \frac{\partial_{z_2} M_1[0, \mathbf{y}(\mu, t, 0)]}{M_1[0, \mathbf{y}(\mu, t, 0)]} \\ + \frac{\mu^2 m_D}{4} [\partial_{z_1} \mathcal{I}(\mu, t, 0) - \partial_{z_2} \mathcal{I}(\mu, t, 0)]. \end{aligned}$$

Thanks to the formulae given in Theorem 40, we have for all $\mu > 0$ and $t > 0$:

$$\partial_{z_1} y_1(\mu, t, 0) - \partial_{z_2} y_1(\mu, t, 0) = -[\partial_{z_1} y_2(\mu, t, 0) - \partial_{z_2} y_2(\mu, t, 0)] = \frac{e^{-2\delta t}}{\cosh(\mu t)}.$$

First,we assume that $\mu = 2\delta$. We have:

$$\begin{aligned} \partial_{z_1} \mathcal{I}(\mu, t, 0) - \partial_{z_2} \mathcal{I}(\mu, t, 0) &= 4 \int_0^t \frac{e^{-\mu s}}{\cosh(\mu s)} \left[ \frac{\tanh[\mu s]}{2\mu} + \frac{s e^{-\mu s}}{2 \cosh[\mu s]} \right] \mathrm{d}s, \\ &= \frac{\tanh[\mu t]}{\mu^2} - \frac{1}{\mu} I_2(2\mu, \mu, t) + 2 J_2(2\mu, \mu, t). \end{aligned}$$

Therefore, when $\mu = 2\delta$, the local adaptation coefficient is:

$$\begin{aligned} LA_1(r_{\max}, \mu, n, \delta, t) = \frac{e^{-2\delta t}}{\cosh(\mu t)} \left[ \frac{\partial_{z_1} M_1[0, \mathbf{y}(\mu, t, 0)]}{M_1[0, \mathbf{y}(\mu, t, 0)]} - \frac{\partial_{z_2} M_1[0, \mathbf{y}(\mu, t, 0)]}{M_1[0, \mathbf{y}(\mu, t, 0)]} \right] \\ + \frac{\mu^2 m_D}{4} \left[ \frac{\tanh[\mu t]}{\mu^2} - \frac{1}{\mu} I_2(2\mu, \mu, t) + 2 J_2(2\mu, \mu, t) \right]. \end{aligned}$$

for $t > 0$, and converges to:

$$LA_1(r_{\max}, \mu, n, \delta, +\infty) = \frac{m_D}{4} \left[ 1 - {}_2F_1(2, 2; 3, -1) + \frac{1}{2} \, {}_3F_2(2, 2, 2; 3, 3; -1) \right],$$

thanks to Lemma 7.

Now, let us turn to the general case $\mu \neq 2\delta$, for which we have for all $t > 0$:

$$\begin{aligned} \partial_{z_1} \mathcal{I}(\mu, t, 0) - \partial_{z_2} \mathcal{I}(\mu, t, 0) &= 4 \int_0^t \frac{e^{-2\delta s}}{\cosh(\mu s)} \left[ \frac{2\delta}{4\delta^2 - \mu^2} \left( 1 - \frac{e^{-2\delta s}}{\cosh[\mu s]} \right) - \frac{\mu}{4\delta^2 - \mu^2} \tanh[\mu s] \right] \mathrm{d}s, \\ &= \frac{8\delta}{4\delta^2 - \mu^2} I_1(2\delta, \mu, t) - \frac{8\delta}{4\delta^2 - \mu^2} I_2(4\delta, \mu, t) \\ &\quad - \frac{2\mu}{4\delta^2 - \mu^2} [I_2(2\delta - \mu, \mu, t) - I_2(2\delta + \mu, \mu, t)], \end{aligned}$$

by Lemma 7. Thus the local adaptation coefficient at time $t > 0$ is equal to:

$$LA_1(r_{\max}, \mu, n, \delta, t) = \frac{e^{-2\delta t}}{\cosh(\mu t)} \left[ \frac{\partial_{z_1} M_1[0, \mathbf{y}(\mu, t, 0)]}{M_1[0, \mathbf{y}(\mu, t, 0)]} - \frac{\partial_{z_2} M_1[0, \mathbf{y}(\mu, t, 0)]}{M_1[0, \mathbf{y}(\mu, t, 0)]} \right] + \frac{2\delta\mu^2 m_D}{4\delta^2 - \mu^2} I_1(2\delta, \mu, t) - \frac{2\delta\mu^2 m_D}{4\delta^2 - \mu^2} I_2(4\delta, \mu, t) - \frac{\mu^3 m_D}{4\delta^2 - \mu^2} \frac{I_2(2\delta - \mu, \mu, t) - I_2(2\delta + \mu, \mu, t)}{2},$$

which converges as $t \to +\infty$, thanks to Lemma 7.

### 6.5.2. Isolated islands

In this section, we develop the results associated to the model of isolated islands, using the fitness vector $(m_1, m_2)$. Taking $\delta = 0$ in (6.18) (or in (6.19)), we need the following lemma to find an explicit formula for $M_1$:

**Lemma 8.** *Let $\mu > 0$ a real constant and $\widetilde{\beta}$ a given real function. We define:*

$$\beta_1(\mathbf{z}) = -\mu^2(z_1 + z_2)z_1, \quad \text{and} \quad \beta_2(\mathbf{z}) = -\mu^2(z_1 + z_2)z_2.$$

*The nonnegative solution $F$ of the problem:*

$$\partial_t F(t, \mathbf{z}) = [1 + \beta_1(\mathbf{z})]\, \partial_{z_1} F(t, \mathbf{z}) + \beta_2(\mathbf{z})\, \partial_{z_2} F(t, \mathbf{z}) + \widetilde{\beta}(\mathbf{z})\, F(t, \mathbf{z}), \tag{6.39}$$

*is given by:*

$$F(t, \mathbf{z}) = F\left[0, \mathbf{y}(\mu, t, \mathbf{z})\right] \exp\left[ \int_0^t \widetilde{\beta}(\mathbf{y}(\mu, s, \mathbf{z}))\, \mathrm{d}s \right], \tag{6.40}$$

*with $\mathbf{y}(\mu, t, \mathbf{z}) = (y_1(\mu, t, \mathbf{z}), y_2(\mu, t, \mathbf{z}))$, where:*

$$y_1(\mu, t, \mathbf{z}) = -\frac{z_2 \cosh[\operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))]} + \frac{\tanh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))]}{\mu}, \tag{6.41}$$

*and:*

$$y_2(\mu, t, \mathbf{z}) = \frac{z_2 \cosh[\operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))]}. \tag{6.42}$$

*Proof.* The proof is similar than the proof of Theorem 40. □

The nonnegative function $M_1^*$ satisfies (6.39) with:

$$\widetilde{\beta}(\mathbf{z}) = r_{\max} - \mu^2 \left[ m_D z_1 z_2 + \frac{n}{2}(z_1 + z_2) \right].$$

Thus we can apply Lemma 8, which implies that:

$$M_1(t, \mathbf{z}) = M_1 \left[ 0, \mathbf{y}(\mu, t, \mathbf{z}) \right] \exp \left[ \int_0^t \widetilde{\beta}(\mathbf{y}(\mu, s, \mathbf{z}))\, \mathrm{d}s \right],$$

with $\mathbf{y}(\mu, t, \mathbf{z}) = (y_1(\mu, t, \mathbf{z}), y_2(\mu, t, \mathbf{z}))$ given by (6.41)–(6.42). Fix $t > 0$ and $\mathbf{z} \in \mathbb{R}_+^2$. To compute the integral $\int_0^t \widetilde{\beta}(\mathbf{y}(\mu, s, \mathbf{z}))\mathrm{d}s$, we need to compute the two integrals:

$$I_1(t, \mathbf{z}) := \int_0^t (y_1(\mu, s, \mathbf{z}) + y_2(\mu, s, \mathbf{z}))\mathrm{d}s, \quad \text{and} \quad I_2(t, \mathbf{z}) := \int_0^t y_1(\mu, s, \mathbf{z}) y_2(\mu, s, \mathbf{z})\mathrm{d}s.$$

We have:

$$I_1(t, \mathbf{z}) = \int_0^t \frac{\tanh[\mu s + \operatorname{atanh}(\mu(z_1 + z_2))]}{\mu} \mathrm{d}s = \frac{1}{\mu^2} \log \left[ \frac{\cosh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh[\operatorname{atanh}(\mu(z_1 + z_2))]} \right],$$

and:

$$\begin{aligned} I_2(t, \mathbf{z}) = {} & \frac{z_2 \cosh[\operatorname{atanh}(\mu(z_1 + z_2))]}{\mu} \int_0^t \frac{\sinh[\mu s + \operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh^2[\mu s + \operatorname{atanh}(\mu(z_1 + z_2))]} \mathrm{d}s \\ & - z_2^2 \int_0^t \frac{\cosh^2[\operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh^2[\mu s + \operatorname{atanh}(\mu(z_1 + z_2))]} \mathrm{d}s, \\ = {} & \frac{z_2}{\mu^2} \left[ 1 - \frac{\cosh[\operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))]} \right] \\ & - \frac{z_2^2 \cosh^2[\operatorname{atanh}(\mu(z_1 + z_2))]}{\mu} \left( \tanh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))] - \mu(z_1 + z_2) \right). \end{aligned}$$

This implies the explicit formula of $M_1$:

$$\begin{aligned} M_1(t, \mathbf{z}) = M_1(0, \mathbf{y}(\mu, t, \mathbf{z}))\, e^{r_{\max} t} & \left[ \frac{\cosh[\operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))]} \right]^{n/2} \\ & \exp \left[ m_D \mu z_2^2 \cosh^2[\operatorname{atanh}(\mu(z_1 + z_2))] \left( \tanh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))] - \mu(z_1 + z_2) \right) \right] \\ & \exp \left[ -m_D z_2 \left( 1 - \frac{\cosh[\operatorname{atanh}(\mu(z_1 + z_2))]}{\cosh[\mu t + \operatorname{atanh}(\mu(z_1 + z_2))]} \right) \right], \end{aligned} \tag{6.43}$$

Using the fact that $N_1(t) = M_1(t, 0, 0)$, and $\overline{m}_1(t) = \partial_{z_1} \log M_1(t, 0, 0)$, the population size in the first habitat at time $t \geq 0$ is equal to:

$$N_1(t) = M_1 \left[ 0, \tanh(\mu t)/\mu, 0 \right] \cosh^{-n/2}(\mu t) \exp(r_{\max} t).$$

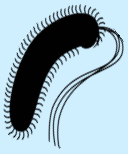
This explicit expression is equivalent to the formula given by Martin; Roques 2016. Similarly, the mean fitness is:

$$\overline{m}_1(t) = \frac{\partial_{z_1} M_1 \left[ 0, \tanh(\mu t)/\mu, 0 \right]}{M_1 \left[ 0, \tanh(\mu t)/\mu, 0 \right]} - \frac{\mu n}{2} \tanh(\mu t).$$

# III

# Invasion into a continuous space

# 7



# Invasion of an asexual population with evolving dispersion

We consider a reaction-diffusion-reproduction equation, modeling a population which is spatially heterogeneous. The dispersion of each individual is influenced by its phenotype. In the literature, the asymptotic propagation speed of an asexual population has already been rigorously determined. The topic of this chapter is to sum up these results, derived from the differential model:

$$\partial_t f(t,x,\theta) = r_1 f(t,x,\theta)\left[1 - K^{-1}\varrho(t,x)\right] + \theta\Delta_x f(t,x,\theta) + \mu\,\Delta_\theta f(t,x,\theta),$$

where $f(t,x,\theta)$ is the density of individuals presenting the trait $\theta \in (\theta_{\min}, +\infty)$ at location $x \in \mathbb{R}$ at time $t > 0$. The quantity $\varrho(t,x) := \int_{\theta_{\min}}^{\infty} f(t,x,\theta)\mathrm{d}\theta$ corresponds to the number of the individuals at position $x \in \mathbb{R}$ at time $t$. In this model, we assume that each individual can invade a new environment according to its phenotype $\theta$. Moreover, when a new individual is born, he can mutate: the mutations are characterized by a positive constant $\mu > 0$. At each position $x \in \mathbb{R}$, the population suffers from competition (*e.g.* light, nutriments, *etc.*), with a constant rate $r_1 > 0$ and a carrying capacity $K > 0$. Different biological details on these parameters will be given in the next section.

This chapter will be followed by a formal study of the invasion of a sexual population (see Chapter 8). The sexual reproduction, modeled by a non-local term, brakes on the speed of the invasion.

In collaboration with V. Calvez, J. Crevat, L. Deckens, B. Fabrèges, F. Kuczma and G. Raoul at the summer school CEMRACS (2018 session), we have made some numerical and formal researches about this biological problem, which yield to a proceeding (Calvez; Crevat; Dekens; Fabrèges, et al. 2019). This chapter and the next one are both based on this paper and on an other one in progress.

# Sommaire

# 7.1. Introduction

In this chapter, we study the time evolution and the spatial and phenotypical propagation of population encowed with an asexual reproduction mode. More precisely, the population is distributed according to its location $x \in \mathbb{R}$ and its phenotypic trait $\theta \in (\theta_{\min}, +\infty)$, with $\theta_{\min} > 0$, and is modeled with the following reaction-diffusion equation for all $t > 0$:

$$\partial_t f(t,x,\theta) = r_1 f(t,x,\theta)\,[1 - K^{-1}\varrho(t,x)] + \theta\Delta_x f(t,x,\theta) + B[f](t,x,\theta), \tag{7.1}$$

where $r_1 > 0$ and $K > 0$ are fixed constants, $f(t,x,\theta)$ is the density of individuals presenting the trait $\theta$ at location $x \in \mathbb{R}$ at time $t \geq 0$, and $\rho(t,x) := \int_{\theta_{\min}}^{\infty} f(t,x,\theta)\,\mathrm{d}\theta$ is the population size at $x \in \mathbb{R}$ and time $t > 0$. The reaction term $B[f]$ will be detailed later. We also assume that initially, the density support is compact. Let us discuss the modeling interest of each term.

Firstly, the term $r_1\, f\, \left[1 - K^{-1}\, \varrho\right]$ stands for an adaptation force called *selection*. At point $x \in \mathbb{R}$ and at time $t \geq 0$, any individual is in competition with the others for resources: when the density $\varrho(t,x)$ at $x$ is less than a threshold $K$, named the *carrying capacity*, there is enough resources for everyone so the population can grow; whereas, if $\varrho(t,x) > K$, then competition is involved, and the population is decreasing. The constant $r_1 > 0$ is therefore called *competition rate*.

Then, the term $\theta\,\Delta_x f$ models the *migration* phenomenon. Individuals are assumed to diffuse through space at each time $t$, at a rate given by the phenotypic trait $\theta$. For example, those with long legs or bigger wings can be faster to go into a new environment, and so help the invasion. In fact, the equation (7.1) without the last reaction term $B[f]$ is the generalized non-local Fisher-KPP equation, where the diffusion depends on the phenotypic traits, which can be seen as a given constant. In the case that $\varrho$ is a convolution term with $f$, the previous equation has been studied in some different ways (Hamel; Ryzhik 2014b; Fang; Zhao 2011; Genieys; Volpert; Auger 2006a; Gourley 2000).

Finally, the last reaction term $B[f]$ models the *reproduction* event. Let us turn to explain this new term. For asexual populations, reproduction is clonal: the offspring receives at birth the same trait as its sole parent, but mutations change the value of the trait. Assuming the variance of the mutation effects is very small, we can model the phenomenon by the following evolutionary equation, as in Benichou; Calvez; Meunier; Voituriez 2012:

$$\partial_t f(t,x,\theta) = r_1 f(t,x,\theta)\,[1 - K^{-1}\varrho(t,x)] + \theta\Delta_x f(t,x,\theta) + \mu\,\Delta_\theta f(t,x,\theta), \tag{7.2}$$

where $\mu > 0$ is a constant depending on the mutation rate and on the variance of the mutation effects. Replacing $f$ by $\theta_{\min} K^{-1}\, f(\theta_{\min}^2 t/\mu, \sqrt{\theta_{\min}^3/\mu}\, x, \theta_{\min}\theta)$ yields us to simplify the initial problem and to study:

$$\partial_t f(t,x,\theta) = r f(t,x,\theta)\,[1 - \varrho(t,x)] + \theta\Delta_x f(t,x,\theta) + \Delta_\theta f(t,x,\theta), \tag{7.3}$$

with $\theta \in (1, +\infty)$, $r > 0$ and $\varrho(t,x) = \int_1^{\infty} f(t,x,\theta)\mathrm{d}\theta$. In this case, the reproduction term is just $Bf(t,x,\theta) = \Delta_\theta f(t,x,\theta)$.

**Remark 5.** *The variation deduced by the mutation event is an approximation of an infinitesimal model (Fisher 1930; Kimura 1965), taking the limit of the equation when the variance of the mutation effect $\lambda^2 > 0$ tends to 0. In this infinitesimal model, this variation follows a centred normally distributed random variable. Without selection and migration, the mutations (Kimura 1965) can be modeled by the evolutionary equation:*

$$\partial_t f(t,x,\theta) = r_2 \left[ J * f(t,x,\theta) - f(t,x,\theta) \right],$$

*where $r_2 > 0$ is the mutation rate, $J$ is the gaussian distribution of $\mathcal{N}(0,\lambda^2)$ and the "convolution term" $J * f(t,x,\theta) = \int_{\theta_{\min}}^{\infty} J(\theta-\theta') f(t,x,\theta') \mathrm{d}\theta'$. Linearising the previous equation, when $\lambda$ is very small, we retrieve the same term as in (7.3), with $\mu = r_2 \lambda$.*

**Boundary condition.** By definition of phenotypic trait, no individual can have a phenotype in $(-\infty, 1)$ or generate a descendent with such phenotype: there is no phenotypical exchange at $\theta = 1$, which means that:

$$\forall t \geq 0,\ \forall x \in \mathbb{R},\ \partial_\theta f(t,x,1) = 0.$$

## 7.2. Analytic results

Without the competition, the PDE, that we obtained in the asexual reproduction case, becomes linear and therefore calls for exponential solutions. Aligning thereby with others studies about diffusion of individuals and speed of the invasion (see for instance Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018), where the common method is to define the function $u$ such that:

$$f(t,x,\theta) = \exp\left[ -t\, u\left( s(t), t^{-\alpha} x, t^{-\beta} \theta \right) \right], \tag{7.4}$$

where $s$ is a time parametrization, and $\alpha$ and $\beta$ are introduced to re-scale the spatial and phenotypic variables.

Finding $(\alpha, \beta)$ crystallizes the main issue tackled by the formal analysis, as it is the crucial key to get asymptotically a non trivial PDE satisfied by $u$ (that is taking into account all the biological forces, namely selection, migration and reproduction), as it is shown in Fig. 7.1. Thanks to (7.3), we can check that, for all $t \geq 0$, $y \geq t^{-\alpha}$ and $\eta \geq t^{-\beta}$:

$$\begin{aligned} &- u(s(t),y,\eta) - t s'(t)\, \partial_s u(s(t),y,\eta) + \alpha y \partial_y u(s(t),y,\eta) + \beta \eta \partial_\eta u(s(t),y,\eta) \\ &\qquad = -\eta t^{1-2\alpha+\beta} \Delta_y u(s(t),y,\eta) + \eta t^{2-2\alpha+\beta} [\partial_y u(s(t),y,\eta)]^2 \\ &\qquad\qquad - t^{1-2\beta} \Delta_\eta u(s(t),y,\eta) + t^{2-2\beta} [\partial_\eta u(s(t),y,\eta)]^2 + r(1-\varrho), \end{aligned}$$

with $\varrho = t^\beta \int_{t^{-\beta}}^{\infty} \exp(-t u(s(t),y,\eta)) \mathrm{d}\eta$. We choose the parametrization $s = \log(t)$ so that $t s'(t) = 1$. In this case, we obtain a viscous Hamilton-Jacobi PDE :

$$\begin{aligned} -u - \partial_s u + \alpha y \partial_y u + \beta \eta \partial_\eta u &= -\eta e^{(1-2\alpha+\beta)s} \Delta_y u + \eta e^{(2-2\alpha+\beta)s} (\partial_y u)^2 \\ &\quad - e^{(1-2\beta)s} \Delta_\eta u + e^{(2-2\beta)s} (\partial_\eta u)^2 + r(1-\varrho) \\ &= e^{(2-2\alpha+\beta)s} \left( \eta (\partial_y u)^2 - \eta e^{-s} \Delta_y u \right) \\ &\quad + e^{(2-2\beta)s} \left( (\partial_\eta u)^2 - e^{-s} \Delta_\eta u \right) + r(1-\varrho). \end{aligned} \tag{7.5}$$

**Remark 6.** *In large time $t$, which is equivalent to taking the limit as $s$ tends to $+\infty$, the factor $\exp(-\text{ Cste } s)$ in front of the laplacians terms will make the second order terms vanish. Nevertheless, the transformation (7.4) enables to capture the long time asymptotics. In fact, it transforms the nature of the problem, from a second order parabolic equation to a nonlinear viscous Hamilton-Jacobi, for which characteristic lines can be analytically computed. However, reviewing the rigorous derivation of the viscous Hamilton-Jacobi equation relies on the theory of viscosity solutions, which is beyond our scope, and we refer to Calvez; Henderson; Mirrahimi; Turanova, et al. 2018 for details.*

Formally, our purpose is to determine the asymptotic behaviour of (7.5) in the limit $s \to +\infty$, that is to derive a PDE in which all biological phenomena intervene, especially:

- ◇ the reproduction term: as the factor $e^{(2-2\beta)s}(\partial_\eta u)^2$ tends to $+\infty$ or $0$ depending on the sign of $(2-2\beta)$, the reproduction term will either annihilate the contribution of the others or vanish. Thus the only relevant choice derives from $2-2\beta=0$.
- ◇ the diffusion term: the same considerations lead us to $2-2\alpha+\beta=0$.

We thereby get the following system for $(\alpha,\beta)$:

$$\begin{cases} 2-2\beta=0, \\ 2-2\alpha+\beta=0, \end{cases}$$

which leads to:

$$(\alpha,\beta)=\left(\frac{3}{2},1\right).$$

In fact, some authors have already rigorously proved the spatial propagation speed, for different values of $r$ (Berestycki; Mouhot; Raoul 2015; Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018). These results are assumed to be summarized by the following (not already proved) general theorem:

**Theorem 41. (SPEED OF INVASION FOR AN ASEXUAL POPULATION)**
*Let $f$ the solution of (7.3), with initial data $f_0$ compactly supported. Then there exist a critical value $c^* = c^*(r) > 0$ such that:*

$$\liminf_{t\to+\infty} \left( \frac{\inf\{x \in \mathbb{R},\ \varrho(t,x) \leq 1/2\}}{t^{3/2}} \right) \leq c^*.$$

*Moreover, for all $\varepsilon > 0$, there exists $\delta > 0$ such that:*

$$\limsup_{t\to+\infty} \left( \frac{\sup\{x \in \mathbb{R},\ \varrho(t,x) \geq \delta\}}{t^{3/2}} \right) \geq c^* - \varepsilon.$$

*Thus the front position $X(t)$ at time $t \geq 0$ is given by:*

$$X(t) = c^*\, t^{3/2}.$$

For $r = 1$, just some approximations have been made (Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018), thanks to a PDE approach. We already know that $c^*(1) \leq 4/3$, and:

$$c^*(1) \approx 1.315135.$$

## 7.3. Numerical results

In this section, we present a numerical scheme to approximate the diffusion with asexual reproduction.

We consider $x_{\max} \geq 0$ and $\theta_{\max} \geq 1$ so that we work with couples $(x, \theta)$ in the bounded domain $[0, x_{\max}] \times [1, \theta_{\max}]$, discretized with the meshes $(x_i)_{1\leq i\leq N_x}$ and $(\theta_j)_{1\leq j\leq N_\theta}$, respectively of step $\Delta x > 0$ and $\Delta\theta > 0$. As for the time discretization, let $\Delta t > 0$ be a time step, and let us define for all $n \in \mathbb{N}$, $t_n := n\,\Delta t$.

In the following, we denote by $A_x^N$ the matrix of the discrete Laplace operator in $x$ of size $N_x$ respectively with Neumann boundary conditions:

$$A_x^N = \frac{1}{\Delta x^2} \begin{pmatrix} -1 & 1 & & & (0) \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ (0) & & & 1 & -1 \end{pmatrix}.$$

We consider $A_\theta^N$ the matrix of the discrete Laplace operator in $\theta$ of size $N_\theta$ respectively with Neumann boundary conditions with similar definitions. In the following, we choose to work with the fixed parameter $r = 1$.

For all $n \in \mathbb{N}$, we approximate $(f(t_n, x_i, \theta_j))_{1\leq i\leq N_x, 1\leq j\leq N_\theta}$ with a matrix:

$$F^n \;=\; \left(F^n_{ij}\right)_{1\leq i\leq N_x, 1\leq j\leq N_\theta}.$$

Then, for all $n \in \mathbb{N}$, we approach $(\varrho(t_n, x_i))_{1\leq i\leq N_x}$ by a vector $(\widetilde{\varrho}^n_i)_{1\leq i\leq N_x}$ computed with an approximation of the integral in $\theta$ using the trapezoidal rule. We have chosen an initial truncated Gaussian distribution:

$$f(0, x, \theta) = \frac{2}{\sqrt{\pi}} \exp\left[-\frac{x^2 + (1-\theta)^2}{2}\right] \mathbb{1}_{\theta\geq 1}. \tag{7.6}$$

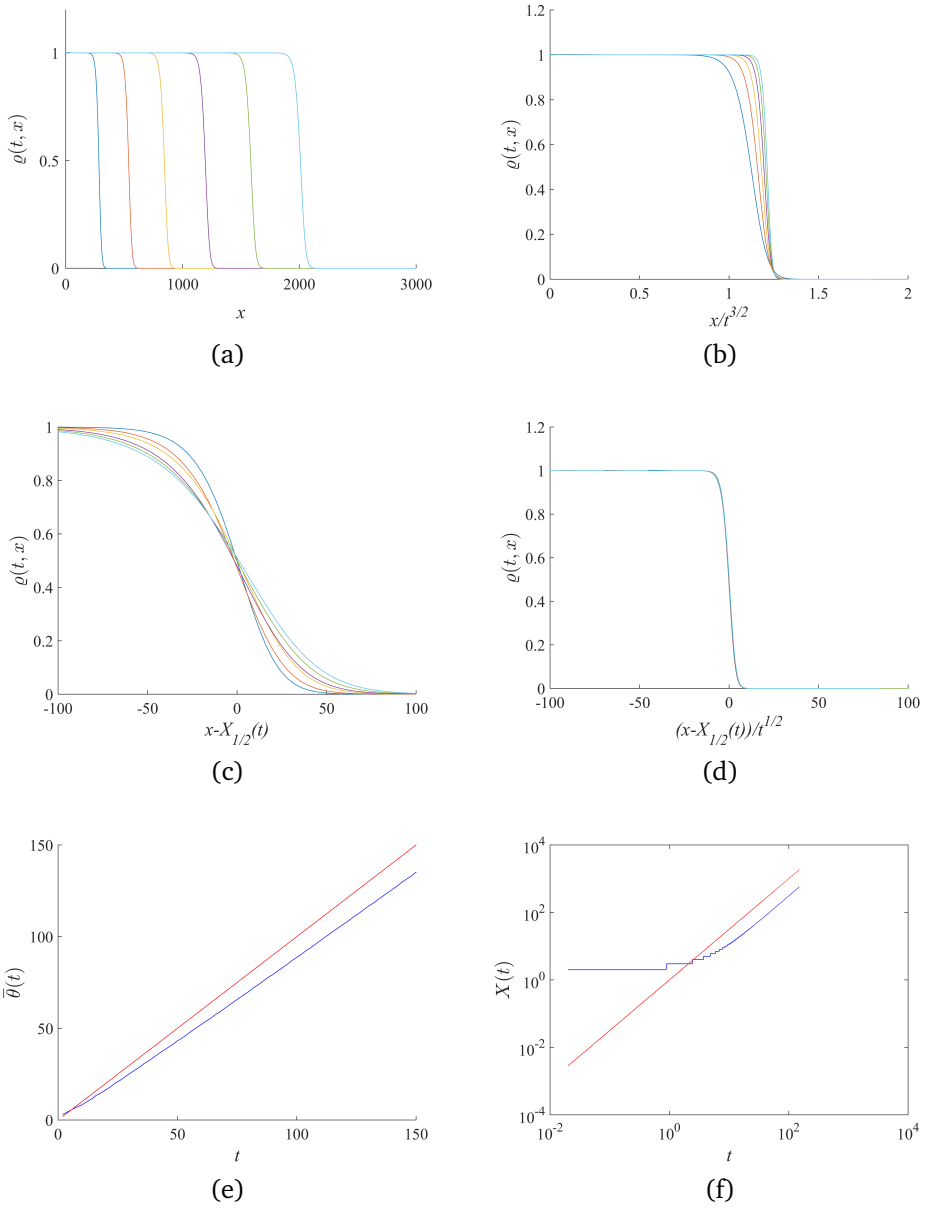



Figure 7.1. – **Simulations of the invasion of an asexual population**, associated to the equation (7.3) with parameters $\Delta t = 0.02$, $\Delta x = 4$, $\Delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. Plots of the density of population $\varrho(t,\cdot)$ for successive fixed times at regular intervals from $t = 20$ to $t = 150$, with respect to (a) the position $x$, (b) to the auto-similar variable $xt^{-3/2}$, (c) to the re-centered variable $x - X_{1/2}(t)$, and (d) to the re-scaled variable $(x - X_{1/2}(t))t^{-1/2}$, where for all $t \geq 0$, $\varrho(t, X_{1/2}(t)) = 1/2$. (e) Plot of the mean phenotype $\overline{\theta}(t)$ (see (7.9)) at the front position with respect to time (blue curve) and of the function $t \to t$ (red curve), in $\log - \log$ scale. (f) Plot of the position of the wave front $\log(X(t))$ with respect to time (blue curve) and of the function $t \mapsto 3/2\log(t)$ (red curve).

First of all, we approximate the reaction-diffusion equation (7.3). We start by discretizing the diffusion terms in $x$ and $\theta$ with the matrix of discrete Laplace operator with Neumann boundary conditions $A_x^N$ and $A_\theta^N$. Then, we approximate in time using an explicit Euler scheme that is for all $n \in \mathbb{N}$, we compute the new value of $F^{n+1}$ by:

$$F^{n+1} \;=\; F^n + \Delta t \,\left[ A_x^N \times F^n \times D_\theta + r\,(F^n - D_\varrho^n \times F^n) + F^n \times A_\theta^N \right], \tag{7.7}$$

where $D_\varrho^n := \mathrm{diag}\left((\widetilde{\varrho}_i^n)_{1\leq i\leq N_x}\right)$, and $D_\theta := \mathrm{diag}\left((\theta_j)_{1\leq j\leq N_\theta}\right)$.

In the following, the steps $\Delta t$, $\Delta x$ and $\Delta\theta$ have been chosen to satisfied the stability condition:

$$2\,\Delta t \left[\frac{\theta_{\max}}{\Delta x^{\,2}} + \frac{1}{\Delta\theta^{\,2}}\right] \leq 1.$$

In Fig. 7.1 (a), we plot the density of population $\varrho(t,\cdot)$ at different fixed times at regular intervals from the instant $t = 20$ to $t = 150$. As expected, the function $\varrho$ seems to quickly converge towards an invasion front propagating towards right connecting $0$ to $1$. Moreover, we observe that the front of propagation in space seems to accelerate. Theoretically, according to Bouin; Henderson; Ryzhik 2017; Berestycki; Mouhot; Raoul 2015, we should observe a propagation of the front of order $t^{3/2}$ in space and of order $t$ in phenotype.

On Fig. 7.1 (b), we check the order of the spatial acceleration plotting the same curves with respect to the auto-similar variable $y = xt^{-3/2}$. We observe that the spatial densities $\varrho$ converge towards an Heaviside distribution in $y$, centered at a constant $y_c$ which is approximately equal to $1.25$. This formula is also consistent with Fig 7.1 (f), comparing the trajectory of the logarithm of position of the wave front ($\log(X(t))$), and the theoretical function $3/2\log(t)$. Theoretically, according to Bouin; Henderson; Ryzhik 2017, this constant is approximately $y_c \approx 1.315$.

Then, we focus on the shape of the transition front. We plot in Fig. 7.1 (c) the spatial distribution $\varrho$ with respect to a re-centered scale in $X_{1/2}(t)$, where $\varrho(t, X_{1/2}(t)) \;=\; 1/2$ for all $t \geq 0$. We can observe that the shape of the front flattens as time goes to infinity. In order to study this evolution, we display in Fig. 7.1 (d) the same curves with respect to the re-scaled variable $\left(x - X_{1/2}(t)\right)\, t^{-1/2}$. We can see that all curves are superposed. Therefore, the shape of the front seems to flatten at order $t^{1/2}$.

Finally, in order to numerically study the accelerations in phenotype of the solution of the equation (7.3), we define for all $t \geq 0$ the position of the front at time $t$ with:

$$X(t) \;:=\; \underset{x\in[0,x_{\max}]}{\mathrm{argmin}}\; |\varrho(t,x) - 0.01|\,, \tag{7.8}$$

and the main phenotype at the head of the front $\bar{\theta}(t)$ with:

$$\bar{\theta}(t) \;:=\; \frac{\int_{\mathbb{R}} \theta\, f(t, X(t), \theta)\,\mathrm{d}\theta}{\varrho(t, X(t))}. \tag{7.9}$$

In Fig. 7.1 (e), we show in log-log scale the evolution of the numerical and formal results of $\bar{\theta}$ with respect to time, and the identity function $t \mapsto t$, in order to compare their slopes. We can observe that for high times, these two curves seem to be parallel. Using a linear regression without taking into account the values where $t$ is too small in order not to consider an eventual transitory state, we find that our numerical approximation of the main phenotype $\bar{\theta}(t)$ is of order $t^{1.01}$. Therefore, this explicit Euler scheme seems to be efficient to reproduce the expected acceleration phenomenon in the asexual case.

8

# Front propagation of a sexual population with evolution of dispersion: a formal analysis

**L. DEKENS [a], F. LAVIGNE [b,c,d]**

[a] Institut Camille Jordan, Université Claude Bernard Lyon 1, France
[b] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
[c] BioSP, INRA, 84914, Avignon, France
[d] ISEM (UMR 5554), CNRS, 34095, Montpellier, France

**Abstract**

The adaptation of biological species to their environment depends on their traits. When various biological processes occur (survival, reproduction, migration, *etc.*), the trait distribution may change with respect to time and space. In the context of invasions, when considering the evolution of a heritable trait that encodes the dispersive ability of individuals, the trait distribution develops a particular spatial structure that leads to the acceleration of the front propagation. That phenomenon is known as spatial sorting. Many biological examples can be cited like the bush cricket in Britain, the cane toad invasion in Australia or the common myna one in South Africa.

Adopting this framework, recent mathematical studies have led to highlight the influence of the reproductive mode on the front propagation. Asexual populations have been shown to spread with an asymptotic rate of $t^{3/2}$ in a minimal reaction-diffusion model, whereas the analogous rate for sexual populations is of $t^{5/4}$ (where $t$ denotes the time). However, the precise description of the behaviour of the front propagation in the sexual case is still an open question.

The aim of this paper is to give precise approximations for large times of its position, as well as some features of the local trait distribution at the front. To do so, we solve explicitly the asymptotic problem derived formally. Numerical simulations are shown to confirm these calculations.

# Sommaire

## 8.1. Introduction

Individuals can be more or less adapted to their environment, depending on their traits. Various processes shape the trait distributions: some of them intervene locally, like survival and reproduction, and others highly depend on the spatial structure of the environment, like migration. Biological invasions are an example of a process where the role of space is structuring. As the combination of locally limited amount of resources and large available inhabited space tends to drive individuals further away, the ability to explore can be selected upon. Morphological features can therefore evolve to increase dispersion: closer to the front of the invasion, cane toads in Australia tends to develop longer legs Phillips; Brown; Webb; Shine 2006, common myna birds in South Africa and *conocephalus discolor* bush cricket in Britain, larger wings Berthouly-Salazar; Rensburg; Le Roux; Vuuren, et al. 2012; Thomas; Bodsworth; Wilson; Simmons, et al. 2001, and seeds of *Lodgepole pine* in western North America, better dispersive habilities (Cwynar; MacDonald 1987).

However, that process is not homogeneous in space: individuals with higher dispersal ability are typically located at the range expansion front. This phenomenon is called *spatial sorting*. Its relationship with the evolution of dispersion has been studied by biologists for the past two decades Birzu; Hallatschek; Korolev 2017; Shine; Brown; Phillips 2011; Thomas; Bodsworth; Wilson; Simmons, et al. 2001; Travis; Dytham 2002; Travis; Mustin; Benton; Dytham 2009. More recently, mathematical studies have been quantifying its influence on the asymptotic speed of the invasion. The model equation that we consider to translate effects of evolution of dispersion traits $\boldsymbol{\theta} > 1$ along reproduction and competition through time ($\boldsymbol{t} \geq 0$) and space ($\boldsymbol{x} \in \mathbb{R}$) on the trait density $\boldsymbol{f}$ is the following:

$$\partial_t \boldsymbol{f}(\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta}) = \boldsymbol{r}\,[\,\underbrace{\boldsymbol{B}[\boldsymbol{f}](\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta})}_{\text{reproduction}} - \underbrace{\boldsymbol{K}^{-1}\boldsymbol{\varrho}(\boldsymbol{t},\boldsymbol{x})\boldsymbol{f}(\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta})}_{\text{competition}}\,] + \underbrace{\boldsymbol{\theta}\Delta_{\boldsymbol{x}}\boldsymbol{f}(\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta})}_{\text{dispersion}}, \quad (8.1)$$

When the dispersal rate is possibly unbounded, the relationship between the front propagation and sustained spatial sorting leads to an acceleration of front propagation Berestycki; Mouhot; Raoul 2015; Bouin; Calvez; Meunier; Mirrahimi, et al. 2012; Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018, contrary to the case of constant dispersion for which it is well established that the front expands asymptotically at constant speed Aronson; Weinberger 1978; Berestycki; Hamel; Nadin 2008; Fang; Zhao 2011; Genieys; Volpert; Auger 2006a; Gourley 2000; Hamel; Ryzhik 2014b; Mirrahimi; Raoul 2013.

To our knowledge, analytical results describing the asymptotic accelerating rate of propagation exist only for asexual (clonal) populations (*e.g.*, see Berestycki; Mouhot; Raoul 2015; Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018), for whom the reproduction operator in (8.1)

is:

$$\boldsymbol{B}[\boldsymbol{f}] = \boldsymbol{f} + \boldsymbol{\sigma^2 \Delta_\theta}\, \boldsymbol{f},$$

for some constant $\boldsymbol{\sigma^2} \geq 0$ depending on the mutation variance and mutation rate. In this case, the position of the population range asymptotically expands as $t^{3/2}$ (see Berestycki; Mouhot; Raoul 2015; Bouin; Calvez; Meunier; Mirrahimi, et al. 2012; Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018 for more details). Furthermore, the precise asymptotic position of the front has been derived in Calvez; Henderson; Mirrahimi; Turanova, et al. 2018, by specifying the prefactor term. The value of this prefactor is sensitive to how the competition is modelled : when it is local in trait, it has been shown to be equal to a larger value Berestycki; Mouhot; Raoul 2015; Bouin; Calvez; Meunier; Mirrahimi, et al. 2012; Bouin; Henderson; Ryzhik 2017.

However, as reproductive mode is thought to potentially significantly influence the rate of propagation Williams; Hufbauer; Miller 2019, we take interest into invasions of sexually reproducing populations. An analogous model as for asexual populations can be built using Fisher's infinitesimal model, a model of allelic segregation that has been studied and used for a century in quantitative genetics, a branch of evolutionary biology Barton; Etheridge; Véber 2017; Bulmer 1972; Fisher 1919; Lange 1978; Tufto 2000; Turelli 2017; Turelli; Barton 1994. This model has also been used to model sexually reproducing populations in several integro-differential studies Bouin; Bourgeron; Calvez; Corro, et al. 2018; Calvez; Garnier; Patout 2019; Mirrahimi; Raoul 2013; Raoul 2017, with the following reproduction operator in (8.1):

$$\boldsymbol{B}[\boldsymbol{f}](\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\theta}) = \iint_{(\boldsymbol{\theta_{\min}}, \infty)^2} \mathcal{G}_{\boldsymbol{\lambda}} \left[ \boldsymbol{\theta} - \frac{\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2}{2} \right] \frac{\boldsymbol{f}(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\theta}_1)\, \boldsymbol{f}(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{\theta}_2)}{\boldsymbol{\varrho}(\boldsymbol{t}, \boldsymbol{x})} \mathrm{d}\boldsymbol{\theta}_1\, \mathrm{d}\boldsymbol{\theta}_2.$$

It assumes that the trait of the offspring is given by the mean parental trait up to a random normal deviation given by $\mathcal{G}_{\boldsymbol{\lambda}}$ with constant segregational variance $\boldsymbol{\lambda^2}$. Using this model, the authors of the report Calvez; Crevat; Dekens; Fabrèges, et al. 2019 predicted and numerically confirmed an asymptotic invasion rate of $t^{5/4}$ for sexually reproducing populations.

However, to understand the complexity of the interplay between ecology and evolution in the dynamics of an invasion, the relationship between the propagation and the trait distribution has to be untangled. That requires to describe precisely the trait distribution and the effect of spatial sorting at the front of the invasion, which is the goal of this paper. First, we present our model and the explicit formula that we derive to approximate the position of the front propagation and its local trait distribution at large times (Section 8.2). Next, we present numerical simulations that confirm this formula (Section 8.3). Finally, we derive formally the limit problem for large times and find an explicit solution to it (Section 8.4).

## 8.2. Deterministic model

In this section, we present the integro - differential model that we use and state our formal result as an approximation of the solutions of the resulting equation. The population is described according to its location $x \in \mathbb{R}$ and its dispersive trait $\theta \in (\theta_{\min}, +\infty)$, with $\theta_{\min} > 0$. Here we are interested by the evolution of the density $f(t, x, \theta)$ of individuals being at time $t \geq 0$ at the location $x \in \mathbb{R}$, presenting the trait $\theta$. We also assume that, initially, the density is compactly supported.

**Our model.** The evolution of the density $f(t, x, \theta)$ can be modeled with the following reaction - diffusion equation for all $t > 0$, $x \in \mathbb{R}$ and $\theta > \theta_{\min}$:

$$\partial_t f(t, x, \theta) = r \left[ B[f](t, x, \theta) - K^{-1} \varrho(t, x) f(t, x, \theta) \right] + \theta \Delta_x f(t, x, \theta), \quad (8.2)$$

where $r > 0$ and $K > 0$ are fixed constants, and $\varrho(t, x) := \int_{\theta_{\min}}^{\infty} f(t, x, \theta) \, \mathrm{d}\theta$ is the population size at $x \in \mathbb{R}$ and time $t > 0$. We will detail the reaction term $B[f]$ later. At first, let us discuss the modelling motivation of each term.

First, the term $r \left[ B[f](t, x, \theta) - K^{-1} \varrho(t, x) f(t, x, \theta) \right]$ is analogous to a logistic growth term that models *reproduction* and *competition*. More precisely, the reproduction term $B[f](t, x, \theta)$ represents the number of new individuals that are born with the trait $\theta$ at time $t \geq 0$ and position $x \in \mathbb{R}$ and we will detail the modelling of the segregational process later. Moreover, at point $x \in \mathbb{R}$ and at time $t \geq 0$, there is a competition between individuals for resources, related to the parameter $K$ which is a measure of the *carrying capacity* of the environment. When the local population size at $x$ is relatively small – $\varrho(t, x) \ll K$ – the local population disposes of enough resources to allow an exponential - like growth, while, if $\varrho(t, x) \gg K$, then competition between individuals is strong, and consequently the local population size decreases. The constant $r > 0$ is therefore called *growth rate at low density*.

Then, the diffusion term $\theta \, \Delta_x f$ models the *dispersion* phenomenon. Individuals are assumed to diffuse through space at each time $t$, at a rate given by the dispersive trait $\theta \geq \theta_{\min}$. When $\theta$ gets larger, it models situations like having longer legs or bigger wings, which potentially give an advantage to explore a new environment faster.

Finally, let us come back to the reproduction operator $B[f]$. We consider a monoecious population in which the individuals breed randomly and only with those at the same location $x \in \mathbb{R}$. At time $t$, an individual with trait $\theta_1$ finds a mate with trait $\theta_2$ with the probability density equal to the trait frequency at position $x$ : $f(t, x, \theta_2)/\varrho(t, x)$. To model the segregation, we use Fisher's infinitesimal model, which classically states that the offspring trait differs from the mean parental trait $(\theta_1 + \theta_2)/2$ according to a normal distribution with a segregrational variance $\lambda^2 > 0$ assumed to be constant and independent of the parental trait values. These assumptions imply the following formulation of the

reproduction term:

$$\boldsymbol{B}[\boldsymbol{f}](\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta}) = \iint_{(\boldsymbol{\theta_{\min}},\infty)^2} \mathcal{G}_{\boldsymbol{\lambda}}\left[\boldsymbol{\theta} - \frac{\boldsymbol{\theta}_1+\boldsymbol{\theta}_2}{2}\right] \frac{\boldsymbol{f}(\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta}_1)\,\boldsymbol{f}(\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta}_2)}{\boldsymbol{\varrho}(\boldsymbol{t},\boldsymbol{x})}\mathrm{d}\boldsymbol{\theta}_1\,\mathrm{d}\boldsymbol{\theta}_2.$$

The term $\mathcal{G}_{\boldsymbol{\lambda}}[\boldsymbol{\theta} - (\boldsymbol{\theta}_1+\boldsymbol{\theta}_2)/2]$, symbolizing the stochasticity of the segregation process, is defined as a normalized Gaussian density with variance $\boldsymbol{\lambda}^2 > 0$, that is:

$$\mathcal{G}_{\boldsymbol{\lambda}}(\boldsymbol{\theta}) := \frac{1}{\sqrt{2\pi\boldsymbol{\lambda}^2}}\exp\left[-\frac{\boldsymbol{\theta}^2}{2\boldsymbol{\lambda}^2}\right]. \tag{8.3}$$

Let us rescale the equation by setting :

$$t = \boldsymbol{r}\boldsymbol{t}, \qquad x = \sqrt{\frac{\boldsymbol{r}}{\boldsymbol{\theta_{\min}}}}\,\boldsymbol{x}, \qquad \theta = \frac{\boldsymbol{\theta}}{\boldsymbol{\theta_{\min}}}, \quad \text{and} \quad f(t,x,\theta) = \frac{\boldsymbol{\theta_{\min}}}{\boldsymbol{K}}\boldsymbol{f}(\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta}).$$

Then, we can simplify the previous PDE into:

$$\partial_t f(t,x,\theta) = B[f](t,x,\theta) - \varrho(t,x)f(t,x,\theta) + \theta\Delta_x f(t,x,\theta), \tag{8.4}$$

with the rescaled population size:

$$\varrho(t,x) = \int_1^\infty f(t,x,\theta)\,\mathrm{d}\theta.$$

By this simplification, the reproduction term is:

$$B[f](t,x,\theta) = \iint_{(1,\infty)^2} \mathcal{G}_{\lambda}\left[\theta - \frac{\theta_1+\theta_2}{2}\right] f(t,x,\theta_1)\,\frac{f(t,x,\theta_2)}{\varrho(t,x)}\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2, \tag{8.5}$$

where $\mathcal{G}_\lambda$ is given by (8.3), and $\lambda = \boldsymbol{\lambda}/\boldsymbol{\theta_{\min}}$. One can notice the truncature at the bottom level $\theta_{\min} = 1$, chosen for the sake of simplicity (note that $\theta_{\min}$ can only take positive values), which does not influence the long time asymptotics in the subsequent analysis as $\theta$ is expected to take large values at the front.

**Main result.** In this paper, we denote by $x \cdot J$, for some $x \in \mathbb{R}$ and $J = [a,b]$, the interval $[xa, xb]$ and $|J|$ the length of the interval $J$. As some computations are only formal, we state our main result as a conjecture:

**Conjecture 42.** *Let the constant:*

$$y_c = 4\left(\frac{\lambda}{3}\right)^{1/2}. \tag{8.6}$$

*There exists an interval $J_0$ centered in 1 such that, for all $J \subset J_0$ open interval centered in 1, the density $f$ at large time $t \geq 0$ can be approximated by:*

$$f(t,x,\theta) = \begin{cases} \exp\left[-\frac{1}{4\lambda^2}\left[\theta - \lambda^{4/5}(6x^2)^{1/5}\right]^2 + \underset{t\to\infty}{\mathcal{O}}(|J|^2)\right], \\ \qquad\qquad \text{for } x \leq y_c\, t^{5/4},\ \theta \in \lambda^{4/5}(6x^2)^{1/5}\cdot J, \\ \exp\left[t - \left(\frac{9x^4}{256\,\lambda^2 t^2}\right)^{1/3}\right]\exp\left[-\frac{1}{4\lambda^2}\left[\theta - \left(\frac{3\lambda^2x^2}{2t}\right)^{1/3}\right]^2 + \underset{t\to\infty}{\mathcal{O}}\left(|J|^2\frac{x^{8/3}}{t^{10/3}}\right)\right], \\ \qquad\qquad \text{for } x \geq y_c\, t^{5/4},\ \theta \in \left(\frac{3\lambda^2x^2}{2t}\right)^{1/3}\cdot J. \end{cases}$$

The justification of this conjecture is postponed to Section 8.4.

Conjecture 42 yields that at each time $t \geq 0$ large enough, the front propagation is at the position:

$$X(t) \approx y_c\, t^{5/4} = 4\left(\frac{\lambda}{3}\right)^{1/2} t^{5/4}. \tag{8.7}$$

Additionally, at large time $t$ and all space position $x \in \mathbb{R}$, the dispersive traits are normally distributed, of variance $2\lambda^2$. Behind the front, *i.e.*, at all position $x \ll X(t)$, the mean dispersive trait $\overline{\theta}$ can be approximated by the value:

$$\overline{\theta}(x) \approx \lambda^{4/5}(6x^2)^{1/5}, \tag{8.8}$$

while ahead of the front, *i.e.*, at all position $x \gg X(t)$, it can be approximated by:

$$\overline{\theta}(t,x) \approx \left(\frac{3\lambda^2x^2}{2t}\right)^{1/3}. \tag{8.9}$$

## 8.3. Simulations and validation

In this section, we display numerical simulations, in order to validate the approximation of the solution of the equation (8.4) provided by Conjecture 42. The initial distribution used for simulation is assumed to be a truncated Gaussian distribution:

$$f(0,x,\theta) = \sqrt{\frac{2}{\pi}}\exp\left[-\frac{x^2+(1-\theta)^2}{2}\right]\mathbb{1}_{\theta\geq 1}, \tag{8.10}$$

with $\mathbb{1}_{\theta\geq 1}$ the characteristic function of $\{\theta \geq 1\}$. The segregational variance $\lambda^2$ is taken equal to $1/2$. The discretization of the sexual reproduction term $B[f]$ represents the biggest challenge for the simulations, in comparison to the asexual case (see Calvez; Crevat; Dekens; Fabrèges, et al. 2019).

### 8.3.1. Scheme

We consider $x_{\max} \geq 0$ and $\theta_{\max} \geq 1$ so that we work with couples $(x, \theta)$ in the bounded domain $[0, x_{\max}] \times [1, \theta_{\max}]$, discretized with the meshes $(x_i)_{1\leq i\leq N_x}$ and $(\theta_j)_{1\leq j\leq N_\theta}$, respectively of step $\delta x > 0$ and $\delta\theta > 0$. As for the time discretization, let $\delta t > 0$ be a time step, and let us define for all $n \in \mathbb{N}$, $t_n := n\,\delta t$. We denote by $A_x^N$ the matrix of the discrete Laplace operator in $x$ of size $N_x$ with Neumann boundary condition at $x = 0$ and Dirichlet boundary condition at $x = x_{\max}$:

$$A_x^N = \frac{1}{\delta x^2}\begin{pmatrix} -1 & 1 & & & (0) \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ (0) & & & 1 & -2 \end{pmatrix} \in \mathcal{M}_{N_x}(\mathbb{R}),$$

and the diagonal matrix:

$$D_\theta = \begin{pmatrix} \theta_1 & & (0) \\ & \ddots & \\ (0) & & \theta_{N_\theta} \end{pmatrix} \in \mathcal{M}_{N_\theta}(\mathbb{R}).$$

Futhermore, we introduce a 3D hypermatrix $G_\theta \in M_{N_\theta,N_\theta,N_\theta}(\mathbb{R})$ such that:

$$\forall 1 \leq i, j, k \leq N_\theta,\ G_\theta(i, j, k) = \mathcal{G}_\lambda\left[\theta_k - \frac{\theta_i + \theta_j}{2}\right],$$

representing the discretization of the segregation kernel ($\mathcal{G}_\lambda$ given by (8.3)).

For all $n \in \mathbb{N}$, we approximate $(f(t_n, x_i, \theta_j))_{1\leq i\leq N_x,1\leq j\leq N_\theta}$ with a matrix:

$$F^n = \left(F_{ij}^n\right)_{1\leq i\leq N_x,1\leq j\leq N_\theta} \in \mathcal{M}_{N_x,N_\theta}(\mathbb{R}),$$

and the population size $(\varrho(t_n, x_i))_{1\leq i\leq N_x}$ with the vector:

$$\widetilde{\varrho}_i^n := \sum_{k=1}^{N_\theta} F_{i,k}^n\,\delta\theta \approx \varrho(t_n, x_i),$$

using the following scheme. At each time iteration $n$,

1. For all index $1 \leq k \leq N_\theta$, we compute the vector $V_{k,l}^n$ defined by:

$$\forall 1 \leq l \leq N_x,\ V_{k,l}^n := \delta\theta^2 \left[F^n\, G_\theta(\cdot,\cdot,k)\, (F^n)^T\right]_{ll}.$$

We can check that $V_{k,l}^n$ is the discretization of the reproduction integral term:

$$\begin{aligned} V_{k,l}^n &= \delta\theta^2 \sum_{i,j=1}^{N_\theta} F_{l,i}^n G_\theta(i,j,k) F_{l,j}^n, \\ &\approx \delta\theta^2 \sum_{i,j=1}^{N_\theta} f(t_n, x_l, \theta_i)\mathcal{G}_\lambda\left[\theta_k - \frac{\theta_i+\theta_j}{2}\right] f(t_n, x_l, \theta_j), \\ &\approx \iint_{(1,\infty)^2} f(t_n, x_l, \theta_1)\mathcal{G}_\lambda\left[\theta_k - \frac{\theta_1+\theta_2}{2}\right] f(t_n, x_l, \theta_2)\,\mathrm{d}\theta_1\mathrm{d}\theta_2. \end{aligned}$$

Now to compute the reproduction matrix $\mathrm{Mat}_{\text{Reprod}} \in \mathcal{M}_{N_x,N_\theta}(\mathbb{R})$, we need to divide the previous quantities by the corresponding $\widetilde{\varrho}_i^n$. To be consistent, we set:

$$\forall 1 \leq i \leq N_x,\ 1 \leq k \leq N_\theta,\ \mathrm{Mat}_{\text{Reprod}}^n(i,k) = \begin{cases} V_{k,i}^n/\widetilde{\varrho}_i^n, & \text{if } \widetilde{\varrho}_i^n > 0, \\ 0, & \text{else.} \end{cases}$$

2. We define the diagonal matrix $D_\varrho^n := \mathrm{diag}\left((\widetilde{\varrho}_i^n)_{1\leq i\leq N_x}\right) \in \mathcal{M}_{N_x}(\mathbb{R})$.
3. We approximate in time using an explicit Euler scheme that is for all $n \in \mathbb{N}$:

$$F^{n+1} := F^n + \delta t\left[A_x^N \times F^n \times D_\theta + r\left(\mathrm{Mat}_{\text{Reprod}}^n - K^{-1} \times D_\varrho^n \times F^n\right)\right]. \tag{8.11}$$

In this section, the parameters $r$ and $K$ are equal to 1. The general scheme 8.11 is used in supplementary materials, to show the effects of differents parameters on the invasion.

To be sure that this scheme gives a good approximation of the solution of the PDE (8.4), the spatial step $\delta x$ is taken large enough.

### 8.3.2. Numerical results

We show our results of simulations of the solution of the Eq. (8.4) in two figures Fig. 8.1 and Fig. 8.2. In the first one, we display different features of the front, whereas in the second one, we compare the numerical trait distribution behind the front with the approximation formally obtained in Conjecture 42.

In the top subfigure Fig. 8.1 (a), the population size $\varrho(t,x)$ is displayed at multiple time regularly spaced between $t = 20$ and $t = 200$ for different scaled position $x$. As expected, thanks to Fig. 8.1 (a), we can see that this front accelerates: there

exists a constant $y_c^{num}$ such that the front at time $t$ is at position:

$$X^{num}(t) = y_c^{num}\, t^{5/4} \approx 1.7\, t^{5/4},$$

where the numerical front position $X^{num}(t)$ at time $t \geq 0$ is defined by:

$$X^{num}(t_n) = x_{i^{num}(t_n)}, \quad \text{with} \quad i^{num}(t_n) := \underset{1 \leq i \leq N_x}{\operatorname{argmin}} \left| \widetilde{\varrho}_i^n - 0.1 \right|. \tag{8.12}$$

More precisely, thanks to a linear regression, the constant $y_c^{num}$ can be approximated by 1.7002, while the time power by 1.2624 (with $R^2 = 1$ and $p$-value $< 10^{-4}$). These numerical results are consistent with (8.7), which numerically gives:

$$X(t) = 4 \left( \frac{1}{2 \times 9} \right)^{1/4} t^{5/4} \approx 1.94\, t^{5/4}.$$

With Fig. 8.1 (b), we confirm that the mean dispersive trait at the front that we get from the numerical simulations is consistent with the approximation given by Conjecture 42. Precisely, let us define the mean dispersive trait $\bar{\theta}^{num}(t)$ at the front position $X^{num}(t)$, given by:

$$\bar{\theta}^{num}(t) := \frac{\int_{\mathbb{R}} \theta\, f(t, X^{num}(t), \theta)\, \mathrm{d}\theta}{\varrho(t, X^{num}(t))}. \tag{8.13}$$

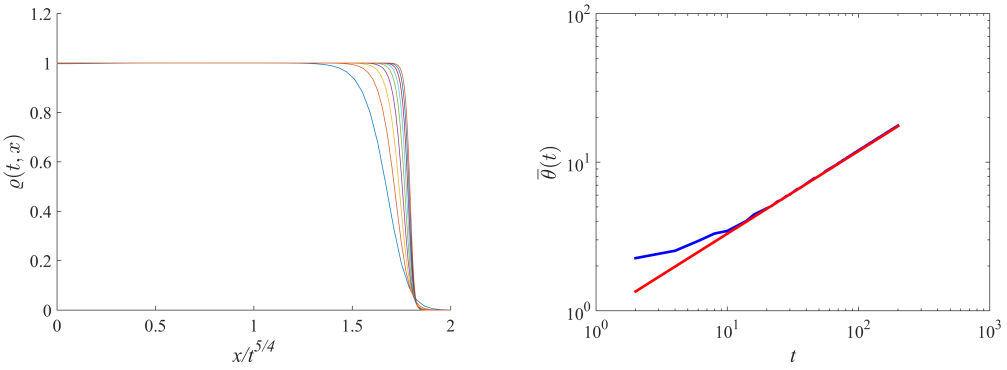
Using a linear regression on the values for $t \in [60, 200]$ (illustrated in Fig. 8.1 (b)), the mean dispersive trait $\bar{\theta}^{num}$ can be approximated by:

$$\bar{\theta}^{num}(t) \approx 0.9654\, t^{0.5588}, \quad (R^2 = 0.9999,\ p\text{-value} < 10^{-14}).$$

We can compare this relationship with the mean dispersive trait $\bar{\theta}(t)$ at the front $X(t)$, given respectively by (8.8) and (8.7):

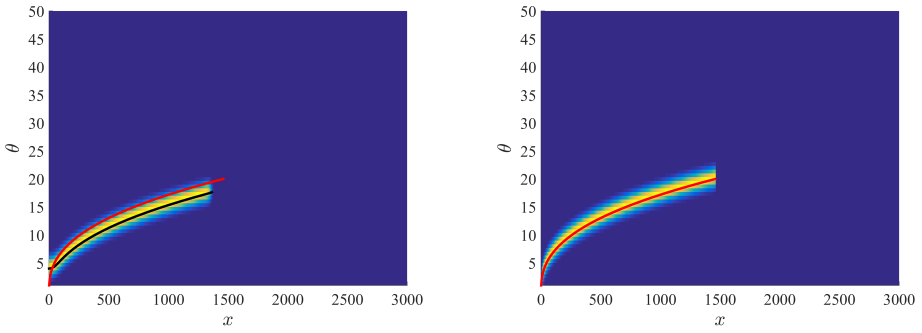
$$\bar{\theta}(t) = \lambda^{4/5} (6X(t)^2)^{1/5} = 2\lambda\sqrt{t} = \sqrt{2t}.$$

Let us turn to the description of the trait distribution behind the front. In Fig. 8.2, we display the contour lines of the trait distribution at time $t = 200$: subfigure (a) is the trait distribution given by the simulations, while (b) is the formal trait distribution (behind the front only) given by Conjecture 42. Our approximation appears to fit the numerical results. More precisely, the red curve, representing the mean dispersive trait at each position behind the front given by (8.8), yields a good approximation of the numerical mean dispersive trait. Moreover, if we represent the numerical trait distribution behind the front at multiple times (see Fig. 8.3), we can see that it seems to remain stationary, which is consistent with the fact that the expression of the approximation behind the front given by Conjecture 42 is independent of the time.



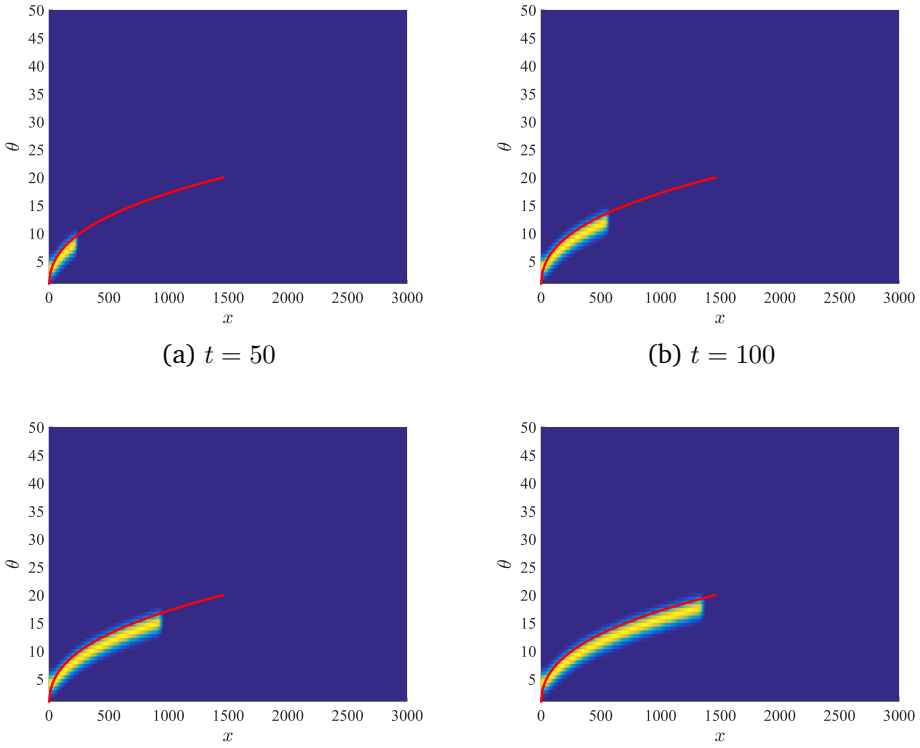

(a) (b)

Figure 8.1. – **Simulations of the invasion of a sexual population**, associated to the Eq.(8.4) with parameters $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. (a) Plot of the density of population $\varrho(t,\cdot)$ for successive fixed times at regular intervals from $t = 20$ to $t = 200$, with respect to the auto - similar variable $xt^{-5/4}$. (b) Plot of the mean dispersive trait $\overline{\theta}^{num}(t)$ (see (8.13)) at the front position with respect to time (blue curve) and of the function $t \to 0.9654t^{0.5588}$ (red curve), in $\log - \log$ scale.




(a) (b)

Figure 8.2. – **Contour lines of the trait distribution of a sexual population**, associated to the Eq. (8.4) with parameters $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. (a) Trait distribution given by the numerical simulations, at $t = 200$. (b) Trait distribution behind the front propagation given by Conjecture 42, at $t = 200$. The red line represents the approximation of the mean trait behind the front propagation given by (8.8), and is common to both subfigures, while the dark line is the mean trait behind the front propagation given by the simulations.




(a) $t = 50$ (b) $t = 100$

(c) $t = 150$ (d) $t = 200$

Figure 8.3. – **Contour lines of the trait distribution during the invasion of a sexual population**, given by simulations, at (a) $t = 50$ (b) $t = 100$ (c) $t = 150$ (d) $t = 200$. The red line represents the approximation of the mean trait behind the front propagation given by (8.8), at time $t = 200$. The parameters are $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$.

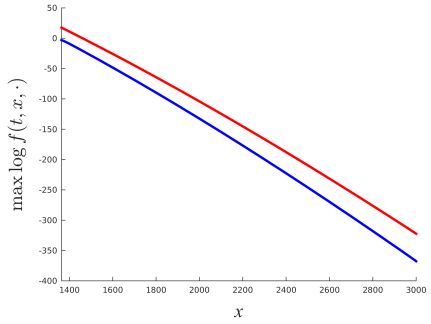



Figure 8.4. – **Plot of logarithm of the amplitude of the distribution ahead of the front at time $t = 200$.** The blue curve represents the log of the maximum of the distribution $f(t, x, \cdot)$ of the numerical approximation given by the scheme, for different positions $x$ located beyond the numerical value of the front position $(X^{\text{num}}(200) \approx 1400)$. The red curve represents the prefactor of the trait distribution ahead of the front given by Conjecture 42. The parameters are $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$.

Fig. 8.4 shows the evolution of the amplitude of the trait distribution $f(t, x, \cdot)$ ahead of the front, in blue curve (log scale). We can see that it can be approximated by the red curve, which displays the prefactor of the trait distribution ahead of the front given by Conjecture 42, and that this approximation holds even at very low density. The difference is due to the other terms of higher power, which are neglected.

## 8.4. Formal proof of the results

This section is devoted to the formal proof of Section 42. In Section 8.4.1, we set the self-similar variables framework suitable to capture the asymptotic invasion acceleration process. Then in Section 8.4.2, we formally derive an asymptotic equation that will allow us in Section 8.4.3 to determine the position of the front and to derive an approximation of the trait distribution $f(t, x, \theta)$ by finding a solution to the limit problem.

### 8.4.1. Preliminaries

According to the same methodology used in previous studies that model the evolution of dispersion (see for instance Bouin; Henderson; Ryzhik 2017; Calvez; Crevat; Dekens; Fabrèges, et al. 2019; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018), we define the function $u$ such that:

$$f(t, x, \theta) = \exp\left[-t\, u\left(s(t), t^{-5/4}x, t^{-1/2}\theta\right)\right], \tag{8.14}$$

where $s(t) = \log(t)$ is a time parametrization (chosen so that $ts'(t) = 1$). We also scale the spatial variable: $y = t^{-5/4}x$; and trait variable: $\eta = t^{-1/2}\theta$; according to the formal arguments of Calvez; Crevat; Dekens; Fabrèges, et al. 2019, which leads to the rate at which the spatial invasion occurs accelerates proportionally to $t^{5/4}$ (see Calvez; Crevat; Dekens; Fabrèges, et al. 2019 for details). Like in the latter, we recall that the power exponents are chosen so that the all biological forces (particularly, migration and reproduction) contribute in a balanced way in the following PDE on $u$, satisfied for all $t \geq 0$, for all $y \in \mathbb{R}$ and for all $\eta \geq e^{-s/2}$:

$$\begin{aligned} -u(s,y,\eta) - \partial_s u(s,y,\eta) + \frac{5}{4} y \partial_y u(s,y,\eta) + \frac{\eta}{2} \partial_\eta u(s,y,\eta) \\ = \eta \left[ (\partial_y u(s,y,\eta))^2 - e^{-s} \Delta_y u(s,y,\eta) \right] \\ + \left( I[u](s,y,\eta) - \varrho_u(s,y) \right), \end{aligned} \tag{8.15}$$

where:

$$\varrho_u(s,y) = e^{s/2} \int_{e^{-s/2}}^{\infty} \exp\left[ -e^s u(s,y,\eta) \right] \mathrm{d}\eta, \tag{8.16}$$

and:

$$\begin{aligned} I[u](s,y,\eta) = \frac{e^s}{\sqrt{2\pi\lambda^2}\varrho_u(s,y)} \\ \iint_{(e^{-s/2},\infty)^2} \exp\left[ e^s \left( -\frac{\left(\eta - \frac{\eta_1+\eta_2}{2}\right)^2}{2\lambda^2} + \left[ u(s,y,\eta) - u(s,y,\eta_1) - u(s,y,\eta_2) \right] \right) \right] \mathrm{d}\eta_1 \, \mathrm{d}\eta_2. \end{aligned} \tag{8.17}$$

Henceforth, we note for the sake of clarity: $\alpha = 5/4$, $\beta = 1/2$ and $o_{\epsilon \to 0}(\epsilon^p)$ any function $r_\epsilon$ such that $\epsilon^{-p}\, r_\epsilon$ converges uniformly relatively to $y$ and $\eta$ towards $0$, as $\epsilon \to 0$.

Our formal aim is to determine the large time behaviour of the solution of (8.15), as $s \to \infty$ (which is equivalent to take $t \to \infty$).

## 8.4.2. Formal asymptotic equation

In this subsection, we will derive from (8.15) an asymptotic equation in the limit $s \to \infty$ that will explicit the interplay between spatial sorting and trait distribution at the front of the solution. The main idea is to perform a Taylor expansion of $u$. For that purpose, let us define the variation $\epsilon = e^{-s/2}$. In the line of Calvez; Garnier; Patout 2019, we make the following ansatz:

$$u(s,y,\eta) = u_0(y,\eta) + \epsilon^2 u_1(y,\eta) + o_{\epsilon \to 0}(\epsilon^2). \tag{8.18}$$

In the next paragraph, we justify the following separation of trait and space

variable in $u_0$, where:

$$u_0(y,\eta) = b(y) + \frac{(\eta - a(y))^2}{4\lambda^2}. \tag{8.19}$$

where $a$ and $b$ are continuous and piecewise differentiable functions of the space variable. Let us interpret them.

Using the ansatz (8.18) and (8.19) in (8.14) yields (we recall that $\epsilon = e^{-s/2}$):

$$f(s,y,\eta) = \exp\left[-\frac{b(y)}{\epsilon^2}\right]\exp\left[-\frac{(\eta - a(y))^2}{4\lambda^2\epsilon^2}\right]\exp\left[-u_1(y,\eta) + \underset{\epsilon\to 0}{o}(\epsilon^2)\right]. \tag{8.20}$$

Hence, when $s \to \infty$, the leading term of the trait distribution $\eta \mapsto f(s,y,\cdot)$ is Gaussian, and the correction is brought by a term determined by $u_1$. The space dependent functions $a$ and $b$ crystallize the main effect of spatial sorting on the trait distribution:

- $\diamond$ $a(y)$ gives the mean rescaled dispersal trait $\eta > 0$ at position $y$. It is therefore positive and satisfies the relation:

$$u_0(y, a(y))) = \min\{u_0(y,\eta), \text{ with } \eta \in (0,\infty)\};$$

- $\diamond$ $b(y)$ determines the prefactor of this distribution: formally, we will see that if $b(y) > 0$, $\varrho_u(s,\cdot)$ vanishes when $s$ tends to $\infty$. On the contrary, the set $\{b(y) = 0\}$ is associated to that area where $\varrho_u$ is asymptotically non-zero. In the context of a spatial invasion, it corresponds to the spatial area that has already been invaded. Hence, we are searching $b$ such that there exists a constant $y_c$ such that $\{b(y) = 0\} = \{y \le y_c\}$. We can interpret $y_c$ as the rescaled position of the front.

Finally, the space dependent functions $a$ and $b$ are linked to the corrector term $u_1$ by an asymptotic equation that we deduce from (8.15) (see below for the details). For $y$ where $a$ and $b$ are differentiable:

$$\begin{aligned} -b(y) - \frac{(\eta - a(y))^2}{4\lambda^2} + \alpha y\left[b'(y) - a'(y)\frac{\eta - a(y)}{2\lambda^2}\right] + \beta\eta\frac{\eta - a(y)}{2\lambda^2} - \eta\left[b'(y) - a'(y)\frac{\eta - a(y)}{2\lambda^2}\right]^2 \\ = \exp\left[u_1(y,\eta) + u_1(y,a(y)) - 2u_1\left(y, \frac{\eta + a(y)}{2}\right)\right] - \mathbf{1}_{\{y\le y_c\}}. \end{aligned} \tag{8.21}$$

In the next section, we find an explicit solution to (8.21), which encodes the intertwined relationship between spatial sorting and trait distribution.

**Explanation for the decomposition of $u_0$ (8.19).** We will recall the fundamental steps, more extensively detailed formally in Bouin; Bourgeron; Calvez; Corro, et al. 2018 and rigorously in Calvez; Garnier; Patout 2019 (for a model without any spatial structure). From the Taylor expansion of $u$ given in (8.18),

we get the following expression for $I[u]$:

$$I[u](s,\eta,y) = \frac{1}{\epsilon\sqrt{2\pi\lambda^2}} \iint_{(\epsilon,\infty)^2} \frac{\exp\left[\frac{1}{\epsilon^2}A^0_{y,\eta}(\eta_1,\eta_2)\right]\exp\left[A^1_{y,\eta}(\eta_1,\eta_2)\right]\exp\left[\underset{\epsilon\to 0}{o}(1)\right]\mathrm{d}\eta_1\mathrm{d}\eta_2}{\int_\epsilon^\infty \exp\left[-\frac{u_0(y,\eta')}{\epsilon^2} - u_1(y,\eta')\right]\mathrm{d}\eta'},$$

where:

$$\begin{cases} A^0_{y,\eta}(\eta_1,\eta_2) = -\frac{1}{2\lambda^2}\left[\eta - \frac{\eta_1+\eta_2}{2}\right]^2 + u_0(y,\eta) - u_0(y,\eta_1) - u_0(y,\eta_2), \\ \\ A^1_{y,\eta}(\eta_1,\eta_2) = u_1(y,\eta) - u_1(y,\eta_1) - u_1(y,\eta_2). \end{cases}$$

Then, we have several considerations to make. First, if we assume that $u_0$ reaches its minimum at a non degenerated-point, then the following modified expression of the denominator:

$$\int_\epsilon^\infty \exp\left[-\frac{1}{\epsilon^2}\left[u_0(y,\eta') - \min u_0(y,.)\right] - u_1(y,\eta')\right]\mathrm{d}\eta',$$

will concentrate, as $\epsilon$ goes to 0, around the minimum of $u_0(y,\cdot)$ and have a finite limit. Therefore it is relevant to introduce it both at the numerator and the denominator:

$$\frac{1}{\left[\epsilon\sqrt{2\pi\lambda^2}\right]^2} \frac{\iint_{(\epsilon,\infty)^2}\exp\left[\frac{1}{\epsilon^2}\left(A^0_{y,\eta}(\eta_1,\eta_2) + \min u_0(y,.)\right)\right]\exp\left[A^1_{y,\eta}(\eta_1,\eta_2) + \underset{\epsilon\to 0}{o}(1)\right]d\eta_1 d\eta_2}{\frac{1}{\epsilon\sqrt{2\pi\lambda^2}}\int_\epsilon^\infty \exp\left[-\frac{1}{\epsilon^2}\left[u_0(y,\eta') - \min u_0(y,.)\right] - u_1(y,\eta')\right]d\eta'}.$$

As we want consequently the numerator not to diverge as $\epsilon \to 0$, we need that:

$$\forall \eta \in \mathbb{R},\ \max_{(\eta_1,\eta_2)}\left[-\frac{1}{2\lambda^2}\left(\eta - \frac{\eta_1+\eta_2}{2}\right)^2 + u_0(y,\eta) - u_0(y,\eta_1) - u_0(y,\eta_2) + \min u_0(y,.)\right] = 0. \tag{8.22}$$

As shown in Bouin; Bourgeron; Calvez; Corro, et al. 2018, thanks to some convexity arguments, this leads necessarily to choose $u_0(y,\cdot)$ as a quadratic function in $\eta$ with variance $\lambda^2$, hence (8.19).

**Deriving the asymptotic Eq. (8.21) verified by $u_1(\eta,y), a(y)$ and $b(y)$.** To get an asymptotic equation from (8.15), we still need to establish (formally) the limit of $I[u](s,y,\eta)$ as $s = -2\log(\epsilon)$ goes to $\infty$, by incorporating the quadratic expression (8.19) of $u_0$ in $I[u]$. We will separate the cases of the numerator and the denominator for the sake of clarity.

According to Laplace's method, as we expect the denominator to concentrate around the minimum of $u_0$, namely at $a(y)$, one can perform the change of

variable $z := \frac{\eta' - a(y)}{\epsilon}$:

$$\frac{1}{\epsilon\sqrt{2\pi\lambda^2}} \int_\epsilon^\infty \exp\left[-\frac{1}{\epsilon^2}\left[u_0(y,\eta') - \min u_0(y,.)\right] - u_1(y,\eta')\right] d\eta'$$
$$= \frac{1}{\sqrt{2\pi\lambda^2}} \int_{1-a(y)/\epsilon}^\infty \exp\left[-\frac{z^2}{4\lambda^2}\right] \exp\left[-u_1\left[y, a(y)+\epsilon z\right]\right] dz \underset{\epsilon\to 0}{\to} \sqrt{2}\exp\left[-u_1\left[y,a(y)\right]\right].$$

Similarly, following the analysis of the authors of Bouin; Bourgeron; Calvez; Corro, et al. 2018 and Calvez; Garnier; Patout 2019 on (8.22), we get that the numerator concentrates around the point $(\overline{\eta},\overline{\eta})$, with $\overline{\eta} = \frac{\eta + a(y)}{2} > 0$, realizing its minimum. One can thus perform the change of variables $(\eta_1,\eta_2) = (\overline{\eta}+\epsilon z_1, \overline{\eta}+\epsilon z_2)$, so that a straightforward computation following the quadratic expression (8.19) of $u_0$ leads to:

$$-\frac{1}{\epsilon^2}\left[-\frac{1}{2\lambda^2}\left[\eta - \frac{\eta_1+\eta_2}{2}\right]^2 + u_0(y,\eta) - u_0(y,\eta_1) - u_0(y,\eta_2) + \min u_0(y,.)\right]$$
$$= \frac{1}{4\lambda^2} z_1 z_2 + \frac{3}{8\lambda^2}(z_1^2+z_2^2), \quad (8.23)$$

and therefore:

$$\frac{1}{\left[\epsilon\sqrt{2\pi\lambda^2}\right]^2} \iint_{(\epsilon,\infty)^2} \exp\left[\frac{1}{\epsilon^2}\left(A^0_{y,\eta}(\eta_1,\eta_2) + \min u_0(y,.)\right)\right] \exp\left[A^1_{y,\eta}(\eta_1,\eta_2) + o_{\epsilon\to 0}(1)\right] d\eta_1 d\eta_2,$$
$$= \iint_{(1-\overline{\eta}/\epsilon,\infty)^2} \frac{\exp\left[-\frac{z_1 z_2}{4\lambda^2} - \frac{3}{8\lambda^2}(z_1^2+z_2^2)\right]}{[\sqrt{2\pi\lambda^2}]^2} \exp\left[u_1(y,\eta) - u_1(y,\overline{\eta}+\epsilon z_1) - u_1(y,\overline{\eta}+\epsilon z_2)\right] dz_1 dz_2,$$
$$\underset{\epsilon\to 0}{\to} \sqrt{2}\exp\left[u_1(y,\eta) - 2u_1(y,\overline{\eta})\right].$$

We can thereby obtain the formal limit of $I[u]$:

$$I[u](s,y,\eta) \underset{s\to\infty}{\to} \exp\left[u_1(y,\eta) + u_1(y,a(y)) - 2u_1\left(y, \frac{\eta+a(y)}{2}\right)\right].$$

Moreover, we need the formal limit of $\varrho_u(s,y)$ as $s = -2\log(\epsilon)$ tends to $\infty$:

$$\varrho_u(-2\log(\epsilon), y) = \frac{1}{\epsilon}\int_\epsilon^\infty \exp\left[-\frac{u(-2\log(\varepsilon,y,\eta)}{\epsilon^2}\right] \mathrm{d}\eta,$$
$$= \exp\left[-\frac{b(y)}{\epsilon^2}\right] \frac{1}{\epsilon}\int_\epsilon^\infty \exp\left[-\frac{(\eta-a(y))^2}{4\lambda^2\epsilon^2}\right] \exp\left[-u_1(y,\eta) + \underset{\varepsilon\to 0}{o}(1)\right] \mathrm{d}\eta,$$
$$= \exp\left[-\frac{b(y)}{\epsilon^2}\right] \int_{1-\frac{a(y)}{\varepsilon}}^\infty \exp\left[-\frac{z^2}{4\lambda^2}\right] \exp\left[-u_1(y,a(y)+\epsilon z) + \underset{\varepsilon\to 0}{o}(1)\right] \mathrm{d}z.$$

Hence, formally, we get:

$$\varrho_u(-2\log(\epsilon), y) \underset{\epsilon\to 0}{\longrightarrow} \mathbf{1}_{\{b(y)=0\}} 2\sqrt{\pi}\lambda \exp\left[-u_1(y,a(y))\right].$$

By integrating all these formal computations in (8.15), we formally obtain an asymptotic equation satisfied by $a$, $b$ and $u_1$, where $a$ and $b$ are differentiable:

$$\begin{aligned} -b(y) - \frac{(\eta - a(y))^2}{4\lambda^2} + \alpha y \left[ b'(y) - a'(y) \frac{\eta - a(y)}{2\lambda^2} \right] + \beta\eta \frac{\eta - a(y)}{2\lambda^2} - \eta \left[ b'(y) - a'(y) \frac{\eta - a(y)}{2\lambda^2} \right]^2 \\ = \exp\left[ u_1(y,\eta) + u_1(y, a(y)) - 2u_1\left( y, \frac{\eta + a(y)}{2} \right) \right] \\ - \mathbf{1}_{\{b(y)=0\}} 2\sqrt{\pi}\lambda \exp\left[ -u_1(y, a(y)) \right]. \end{aligned}$$

As we are describing a front propagation, we are looking for $a$ and $b$ continuous on $\mathbb{R}$ and differentiable everywhere but not necessarily at the front position (to be determined):

$$y_c = \sup\{y,\ b(y) = 0\}.$$

For such functions $a$ and $b$, we have by evaluating the latter at $\eta = a(y)$ for $y < y_c$:

$$2\sqrt{\pi}\lambda \exp\left[ -u_1(y, a(y)) \right] = 1.$$

Hence, for $y \neq y_c$ and $\eta \in J_y$ (subset of $\mathbb{R}_+^*$ to be determined), we consider the asymptotic Eq. (8.21):

$$\begin{aligned} -b(y) - \frac{(\eta - a(y))^2}{4\lambda^2} + \alpha y \left[ b'(y) - a'(y) \frac{\eta - a(y)}{2\lambda^2} \right] + \beta\eta \frac{\eta - a(y)}{2\lambda^2} - \eta \left[ b'(y) - a'(y) \frac{\eta - a(y)}{2\lambda^2} \right]^2 \\ = \exp\left[ u_1(y,\eta) + u_1(y, a(y)) - 2u_1\left( y, \frac{\eta + a(y)}{2} \right) \right] - \mathbf{1}_{\{y<y_c\}}. \end{aligned}$$

### 8.4.3. Resolution of the asymptotic Eq. (8.21)

Let us define for $y \neq y_c$, $\eta > 0$:

$$\begin{aligned} g(y,\eta) := -b(y) - \frac{(\eta - a(y))^2}{4\lambda^2} + \alpha y \left[ b'(y) - a'(y) \frac{\eta - a(y)}{2\lambda^2} \right] \\ + \beta\eta \frac{\eta - a(y)}{2\lambda^2} - \eta \left[ b'(y) - a'(y) \frac{\eta - a(y)}{2\lambda^2} \right]^2 + \mathbf{1}_{\{y<y_c\}}. \end{aligned}$$

Let us fix $y \neq y_c$. For $\eta > 0$ such that $g(y,\eta) > 0$, we can reformulate (8.21) as:

$$T_y(\eta) = L_y(u_1)(\eta), \tag{8.24}$$

where:

$$T_y(\eta) = \log\left[ g(y,\eta) \right],$$

and:

$$L_y(u_1) : \eta \mapsto u_1(y,\eta) + u_1(y,a(y)) - 2u_1\left(y, \frac{\eta + a(y)}{2}\right).$$

Eq. (8.24) suggests that $a$, $b$ and $y_c$ are to be chosen so that $T_y$ lies in the image of the linear operator $L_y$. One can notice that the kernel of $L_y$ is composed of the linear functions, hence:

$$\dim \ker (L_y) = 2.$$

Heuristically, the image of $L_y$ is orthogonal to a two dimensional space, which is generated by $\delta_{a(y)}$ and $\delta'_{a(y)}$. More precisely, following Calvez; Garnier; Patout 2019, one can show that if $T_y$ verifies:

$$\begin{cases} T_y\left(a(y)\right) = 0, \\ T'_y\left(a(y)\right) = 0, \end{cases} \tag{8.25}$$

then the following sum converges:

$$u_y : \eta \mapsto \sum_{k=0}^{\infty} 2^k T_y\left[a(y) + (\eta - a(y))\ 2^{-k}\right], \tag{8.26}$$

and $L_y(u_y) = T_y$.

Hence, we first need to solve (8.25), that is to find $y_c > 0$, $(a,b) \in C^0(\mathbb{R}) \cap C^1(\mathbb{R}\backslash\{y_c\})$, such that:

$$\forall y \neq y_c, \quad \begin{cases} -b(y) + \alpha y b'(y) - a(y)(b'(y))^2 + \mathbf{1}_{\{y<y_c\}} = 1, \\ -\alpha y a'(y) + \beta a(y) - 2\lambda^2 (b'(y))^2 + 2a(y)b'(y)a'(y) = 0. \end{cases} \tag{8.27}$$

Here, we present an explicit solution to (8.27):

**Proposition 43.** *Let us define:*

$$y_c = 4\sqrt{\frac{\lambda}{3}}, \qquad a : y \mapsto \begin{cases} \lambda^{4/5}\, 6^{1/5}\, y^{2/5}, & \text{if } y \leq y_c, \\ \left(\frac{3\lambda^2}{2}\right)^{1/3} y^{2/3}, & \text{if } y > y_c, \end{cases}$$

*and:*

$$b : y \mapsto \begin{cases} 0, & \text{if } y \leq y_c, \\ \left(\frac{3}{\lambda 2^4}\right)^{2/3} y^{4/3} - 1, & \text{if } y > y_c. \end{cases}$$

*Then $a, b \in C^0(\mathbb{R}) \cap C^1(\mathbb{R}\backslash\{y_c\})$ and $y_c$, $a$ and $b$ are solutions of (8.27).*

**Remark 7.** *The functions $a$, $b$ and $y_c$ given in the previous proposition are the only solutions of (8.27) of the form : $a(y) = Cy^m$, $b(y) = Ky^n - 1$ that are positive for $y > y_c$ and continuous in $y_c$.*

To derive a solution for (8.21) from Proposition 43, one still has to define $T_y(\eta)$, which requires $g(y, \eta) > 0$. As $g(y, \cdot)$ is a three order polynomial in $\eta$ with a negative leading coefficient, it is not positive as $\eta$ becomes large so we can not define $T_y$ on $\mathbb{R}_+^*$. However, $a, b$ and $y_c$ are solutions of (8.27), which is equivalent to:

$$g(y, a(y)) = 1, \quad \partial_\eta g(y, a(y)) = 0.$$

We aim therefore at solving (8.21) locally in $\eta$ around $a(y)$:

**Proposition 44.** *Let $a, b$ and $y_c$ be as in Proposition 43. Then, there exists $J_0$ an interval centered in 1 such that, for all $y \neq y_c$, $\eta > 0$ such that $\frac{\eta}{a(y)} \in J_0$, we have $g(y, \eta) > 0$. Moreover, for $y \neq y_c$, $T_y = \log(g(y, \cdot))$ is well defined on $a(y) \cdot J_0$ and for all $J \subset J_0$ open interval centered in 1:*

- *for $y < y_c$ and $\eta \in a(y) \cdot J$, the series defined in (8.26) converges and is bounded uniformly with regard to $\eta$ and $y$ and the bound is of the form $A\,|J|^2$.*
- *for $y > y_c$ and $\eta \in a(y) \cdot J$, the series defined in (8.26) converges and is bounded uniformly with regard to $\eta$, and the bound is of the form: $B|J|^2 y^{8/3}$.*

*Proof.* Since $g(y, \cdot)$ is a polynomial of order three in $\eta$ such that:

$$g(y, a(y)) = 1, \qquad \partial_\eta g(y, a'(y)) = 0,$$

we can define $P_y$ polynomial of order three such that:

$$\forall \eta > 0, \quad g(y, \eta) = 1 - P_y\left(\frac{\eta}{a(y)}\right).$$

As $P_y(1) = P_y'(1) = 0$, we get:

$$P_y(X) = (X-1)^2\left[\gamma X + P_y(0)\right],$$

where $\gamma > 0$ is the leading coefficient of $P_y$.

We next compute, for $y \neq y_c$ (by continuity for $P_y(0)$):

$$\gamma = \frac{a'(y)^2 a(y)^3}{4\lambda^4}, \quad P_y(0) = b(y) + \frac{a(y)^2}{4\lambda^2} - \alpha y\left[b'(y) + a'(y)\,\frac{a(y)}{2\lambda^2}\right] + \mathbf{1}_{\{y>y_c\}}.$$

Hence (adopting the notations $K_{a^-}$,$K_{a^+}$ and $K_b$ such that for $y < y_c$, $a(y) = K_{a^-} y^{2/5}$ and for $y > y_c$, $a(y) = K_{a^+} y^{2/3}$, $b(y) = K_b y^{4/3} - 1$ – see the previous proposition):

- ◇ for $y < y_c$, $\gamma = \frac{a(y)^5}{25y^2\lambda^4} = \frac{{K_{a^-}}^5}{25\lambda^4}$ and:

$$P_y(0) = \frac{a^2}{4\lambda^2} - \frac{\alpha a'(y) y a(y)}{2\lambda^2} = \frac{a^2}{4\lambda^2} - \frac{5}{4} \cdot \frac{2a^2}{10\lambda^2} = 0.$$

  So, in that case, $P_y = \frac{{K_{a^-}}^5}{25\lambda^4}(X-1)^2 X := P(X)$ does not depend on $y$. As $P(1) = 0$, there exists $\delta \in (0,1)$ such that for all $y < y_c$ and $\eta \in ]a(y)(1-\delta), a(y)(1+\delta)[$, $P\left(\frac{\eta}{a(y)}\right) < 1$, hence $g(y,\eta) > 0$.

- ◇ for $y > y_c$, $\gamma = \frac{4}{9}\frac{a(y)^5}{4y^2\lambda^4} = \frac{K_{a^+}^5}{9\lambda^4} y^{4/3}$ and:

$$\begin{aligned} P_y(0) &= (b(y)+1) - \frac{5}{3}(1+b(y)) + \frac{a(y)^2}{4\lambda^2} - \frac{5a(y)^2}{12\lambda^2}, \\ &= -y^{4/3}\left[\frac{K_{a^+}^2}{6\lambda^2} + \frac{2K_b}{3}\right] = -\gamma y^{4/3}\left[\frac{3\lambda^2}{2K_{a^+}^3} + \frac{6K_b\lambda^4}{K_{a^+}^5}\right], \\ &= -\gamma y^{4/3}\left[1 + 6 \times \frac{3^{2/3}\lambda^4 2^{5/3}}{2^{8/3}\lambda^{12/3}3^{5/3}}\right] = -2\gamma y^{4/3}. \end{aligned}$$

  Hence: $P_y(X) = \gamma y^{4/3}(X-1)^2(X-2)$, thus: $\forall y > y_c, \forall \eta \in ]0, 2a(y)[, g(y,\eta) > 1 > 0$.

That proves the first part of the proposition. Let us call $J_0$ a closed interval centered in 1 on which, for all $y \neq y_c$, $g(y,.)$ is positive, and on which $T_y$ is therefore well-defined.

Let us now consider $J \subset J_0$ an open interval centered in 1. For $y \neq y_c$, $\eta \in a(y) \cdot J$, let us define, for $k \in \mathbb{N}$:

$$\eta_k := a(y) + \frac{\eta - a(y)}{2^k}.$$

Next, as $T_y(a(y)) = T'(a(y)) = 0$, we get the following:

$$2^k T_y(\eta_k) = 2^k \int_{a(y)}^{\eta_k} T_y''(t) \frac{\eta_k - t}{2} \mathrm{d}t.$$

With the change of variables $s = 2^k\,(t - a(y))$, we get:

$$2^k T_y(\eta_k) = \int_0^{\eta - a(y)} T_y''\left(a(y) + s2^{-k}\right) \frac{\eta - a(y) - s}{2^k} \mathrm{d}s. \tag{8.28}$$

$T_y''$ is continuous on $a(y) \cdot J$, so the latter ensures that $\sum_{k\geq 0} 2^k T\left(\eta_k\right)$ converges for all $\eta \in a(y) \cdot J$.

Finally, for $y \neq y_c$, we need to uniformly bound $\sum_{k\geq 0} 2^k T\left(\eta_k\right)$ with regard to $\eta \in a(y) \cdot J$. For $y < y_c$, from the first part of the proof, we have:

$$\forall \eta \in a(y) \cdot J, \quad T_y(\eta) = \log\left(1 - P\left(\frac{\eta}{a(y)}\right)\right),$$

with $P(X) = \gamma X(X-1)^2$ and $\gamma$ independent of $y$ and $\eta$. Setting:

$$\begin{array}{rcccl} F & : & J & \to & \mathbb{R}, \\ & & x & \mapsto & \log\left(1 - P(x)\right), \end{array}$$

we dispose of a smooth function, independent from $y$ and $\eta$, such that:

$$\forall \eta \in a(y) \cdot J, \quad T_y(\eta) = F\left(\frac{\eta}{a(y)}\right),$$

and therefore $T_y''(\eta) = F''\left(\eta/a(y)\right)/a(y)^2$. Following (8.28), we get (writing $|J|$ as the length of $J$):

$$\forall y < y_c, \eta \in a(y) \cdot J, \quad \sum_{k \geq 0} |2^k T_y(\eta_k)| \leq \sum_{k \geq 0} 2^{-(k+1)} \|F''\|_{\infty, J} \frac{(\eta - a(y))^2}{a(y)^2} \leq |J|^2 \|F''\|_{\infty, J_0}.$$

For $y > y_c$, we have from above:

$$\forall \eta \in a(y) \cdot J, \quad T_y(\eta) = \log\left(1 - y^{4/3} Q\left(\frac{\eta}{a(y)}\right)\right),$$

with $Q(X) = \gamma_Q (X-1)^2 (X-2)$ ($\gamma_Q$ a constant independent of $y$ and $\eta$). A straightforward calculus leads to:

$$T_y''(\eta) = -\frac{y^{4/3}}{a(y)^2} \left[ \frac{Q''\left(\frac{\eta}{a(y)}\right)}{1 - y^{4/3} Q\left(\frac{\eta}{a(y)}\right)} + y^{4/3} \frac{Q'\left(\frac{\eta}{a(y)}\right)^2}{\left(1 - y^{4/3} Q\left(\frac{\eta}{a(y)}\right)\right)^2} \right].$$

We recall that, additionally, for $y > y_c$ and $\eta \in a(y) \cdot J$, we have: $1 - y^{4/3} Q\left(\frac{\eta}{a(y)}\right) > 1$. Hence, from (8.28), we get:

$$\begin{aligned} \forall y > y_c, \forall \eta \in a(y) \cdot J, \quad \sum_{k \geq 0} |2^k T_y(\eta_k)| &\leq y^{4/3} |J|^2 \left[ \|Q''\|_{\infty, J} + y^{4/3} \|Q'^2\|_{\infty, J} \right] \\ &\leq y^{8/3} |J|^2 \left[ \frac{\|Q''\|_{\infty, J_0}}{{y_c}^{4/3}} + \|Q'^2\|_{\infty, J_0} \right]. \end{aligned}$$

□

The last proposition allows us to complete our solution for (8.21) for $y \neq y_c$ and $\eta \in a(y) \cdot J$, by defining:

$$u_1 : (y, \eta) \mapsto \sum_{k \geq 0} 2^k T_y\left(a(y) + (\eta - a(y))\, 2^{-k}\right).$$

It also highlights the fact that this solution is local in trait around the mean trait $a(y)$. Finally, we use it in Conjecture 42 to specify the magnitude of the error terms in our approximation at large times.

## 8.5. Discussion

**Contributions** In this paper, we have developed a different framework than the one used for the study of asexual populations (Berestycki; Mouhot; Raoul 2015; Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018) by using a mixing operator to analyze the behaviour of the propagation front for sexual population. We have formally found an explicit approximation of the trait distribution during the invasion by finding a solution to the limit problem at large times. These formal computations have been numerically compared to the solution of (8.4) and thus confirmed. All the computations have been made after having rescaled the partial differential Eq. (8.2). By a variable change, we have that, for all growth rate at low density $\boldsymbol{r} > 0$, carrying capacity $\boldsymbol{K} > 0$ and segregational variance $\boldsymbol{\lambda}^2 > 0$, for a population with dispersive traits $\boldsymbol{\theta} \geq \boldsymbol{\theta_{\min}} > 0$, the density $\boldsymbol{f}$ can be approximated at large time $\boldsymbol{t} > 0$ by:

$$\boldsymbol{f}(\boldsymbol{t},\boldsymbol{x},\boldsymbol{\theta}) \approx \frac{\boldsymbol{K}}{\boldsymbol{\theta_{\min}}} \begin{cases} \exp\left[-\frac{1}{4\boldsymbol{\lambda}^2}\left[\boldsymbol{\theta} - \boldsymbol{\lambda}^{4/5}\left(6\boldsymbol{r}\boldsymbol{x}^2\right)^{1/5}\right]^2\right], \text{ for } \boldsymbol{x} \leq \boldsymbol{y_c}\,\boldsymbol{t}^{5/4}, \\ \exp\left[\boldsymbol{r}\boldsymbol{t} - \left(\frac{9\boldsymbol{x}^4}{256\,\boldsymbol{\lambda}^2\boldsymbol{t}^2}\right)^{1/3}\right] \exp\left[-\frac{1}{4\boldsymbol{\lambda}^2}\left[\boldsymbol{\theta} - \left(\frac{3\boldsymbol{\lambda}^2\boldsymbol{x}^2}{2\boldsymbol{t}}\right)^{1/3}\right]^2\right], \\ \qquad\qquad \text{for } \boldsymbol{x} \geq \boldsymbol{y_c}\,\boldsymbol{t}^{5/4}. \end{cases}$$

with:

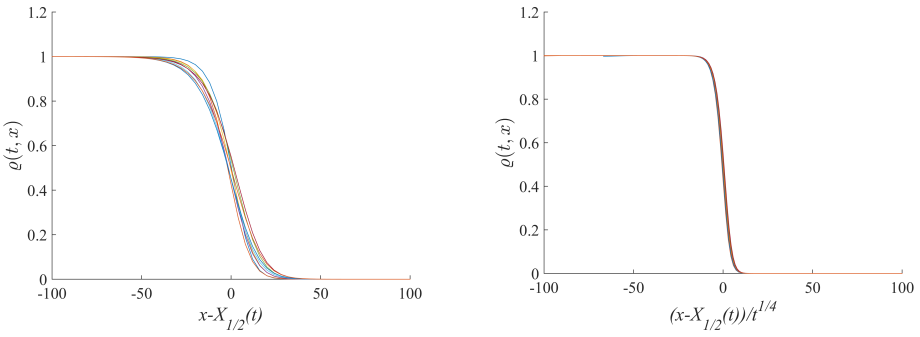
$$\boldsymbol{y_c} = y_c\sqrt{\frac{\boldsymbol{\theta_{\min}}}{\boldsymbol{r}}}\,\boldsymbol{r}^{5/4} = 4\left[\frac{\boldsymbol{\lambda}}{3}\right]^{1/2}\sqrt{\boldsymbol{\theta_{\min}}}\,\boldsymbol{r}^{3/4} = 4\left[\frac{\boldsymbol{\lambda}}{3}\right]^{1/2}\boldsymbol{r}^{3/4}.$$

**Difference in acceleration rate between asexual and sexual invasive populations** Our study shows that the effect of spatial sorting only, through the evolution of dispersion, accelerates the speed at which a sexual population invades. The rate of this acceleration, of $t^{5/4}$, is lower than when considering the influence of the same phenomenon on asexual populations ($t^{3/2}$, see Berestycki; Mouhot; Raoul 2015; Bouin; Henderson; Ryzhik 2017; Calvez; Henderson; Mirrahimi; Turanova, et al. 2018). Mathematically, the blending inheritance property of the infinitesimal model operator reduces the effect of the spatial sorting by crossing extremely dispersive individuals with less dispersive ones, which does not happen for individuals reproducing clonally.

**Extension: Shape of the front** However, there are still structural questions to answer on the asymptotic behaviour of the front that we can observe numerically. For instance, the additional Fig. 8.5 allows us to study the deformation of the front propagation, more precisely the shape of the transition front. In Fig. 8.5 (a), the spatial distribution $\varrho$ is displayed with respect to a re-centered scale in:

$$X_{1/2}(t) = \sup\{x \in \mathbb{R}, \varrho(t,x) = 1/2\}. \tag{8.29}$$

We can observe a flattening of the front shape, as $t \to +\infty$. More precisely, Fig 8.5 (b), displaying $\varrho$ with respect to the re-scaled variable $\left(x - X_{1/2}(t)\right)\, t^{-1/4}$, shows that the shape of the front seems to flatten at order $t^{1/4}$, as the different curves overlap.




(a) (b)

Figure 8.5. – **Plots of the density $\varrho(t, \cdot)$ of a sexual population, with respect to re-centered variables.** The two plots show the evolution of the population density, associated to (8.4), for successive times at regular intervals from $t = 20$ to $t = 200$, with respect to (a) the re-centered variable $x - X_{1/2}(t)$, and (b) to the re-scaled variable $(x - X_{1/2}(t))t^{-1/4}$, with $X_{1/2}$ defined by (8.29). The parameters are $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$.

**Expansion load** Here, we consider only a trait linked to the dispersive ability, thus isolating the sole effect of spatial sorting in range expansions, for which there existed no previous precise results. By doing so, our model does not account for any process of selection by adaptation to the local environment. However, in cases of fast range expansion, a phenomenon called the *expansion load* can occur Peischl; Dupanloup; Kirkpatrick; Excoffier 2013. As the density of individuals at the front is low, the effective strength of natural selection is reduced allowing deleterious mutations to accumulate at the front. That would eventually undermine the invasion process by reducing the fitness of leading individuals (see Burton; Phillips; Travis 2010), with the potential effect of slowing down the speed of the front in comparison to the asymptotic formal result of our study. Nevertheless, the clear relationship between the effect of spatial sorting and expansion load is yet to be explored, as a recent analysis using a discrete space framework seems to indicate that the evolution of dispersal rate can prevent expansion load in certain cases (see Peischl; Gilbert 2020). By isolating the effect of spatial sorting, our study can therefore constitute a first step in understanding the intricate relationship between the evolution of dispersion and of life history traits, ultimately providing tools to analyse the source of variability in range

expansions (see Williams; Hufbauer; Miller 2019).

Because of the formal computations leading to disregard the effect of competition ahead of the front, even though the simulations seems to validate our results, this paper has to be seen as a premise for a consistent and rigorous proof for this problem.

## 8.6. Supplementary materials

**Simulations with a different initial distribution.** We explore the robustness of the numerical confirmation of our approximation drawn in Section 8.2, with regard to the initial conditions. We hereby present the simulations given by the explicit Euler scheme, with the same model parameters as for Fig. 8.1. However, we assume that the initial distribution is a Dirac distribution at $x = 0$ and $\theta = 1$.

The same conclusions seem to hold: at large time $t$, this invasion seems to follow the same evolution as when being initially normally distributed (*cf.* Section 8.3). More precisely, we can see that the propagation front accelerates (see Fig. 8.6 (a) and (b)) and that the acceleration in space can be quantified similarly: $X(t) = y_c t^{5/4}$, with $y_c > 0$.

Fig. 8.7 displays the numerical solution to our differential equation with a Dirac initial distribution for successive times. It is compared to the mean dispersive trait (red line) derived from the asymptotic approximation stated in Conjecture 42. Behind the front, the distribution seems to be stationary at large time.

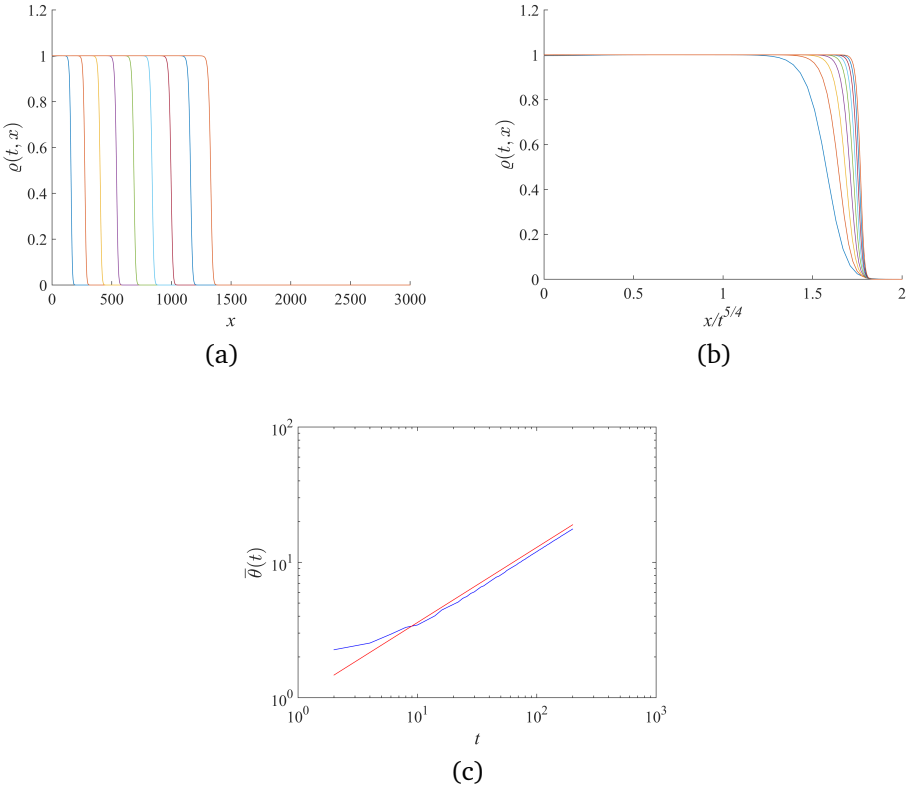
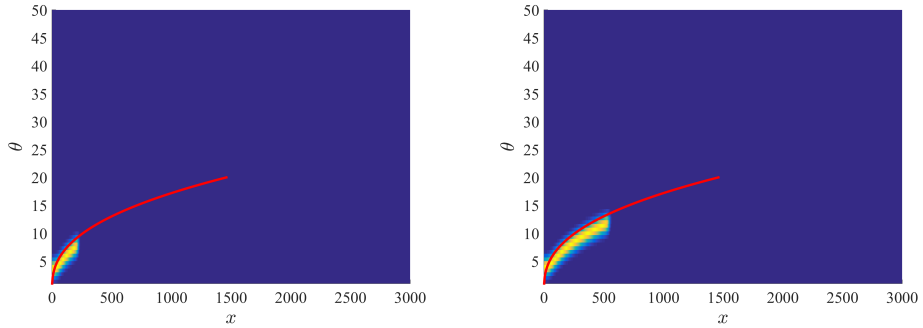
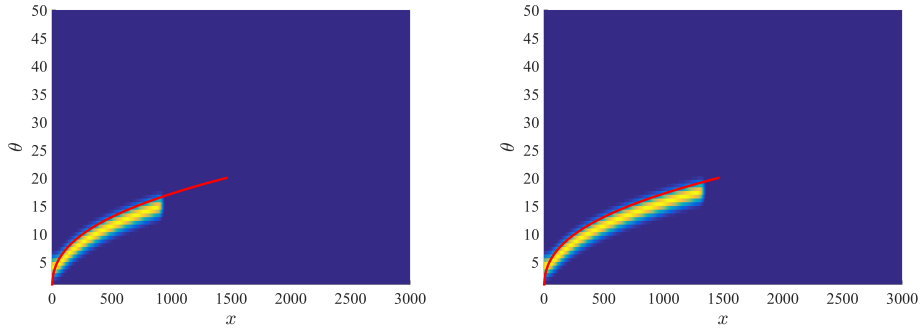



Figure 8.6. – **Simulations of the invasion of a sexual population, initially distributed according to a Dirac distribution**, with parameters $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. Plots of the density of population $\varrho(t,\cdot)$ for successive fixed times at regular intervals from $t = 20$ to $t = 200$, with respect to (a) the position $x$ and (b) the auto - similar variable $xt^{-5/4}$. (c) Plot of the mean dispersive trait $\overline{\theta}^{num}(t)$ at the front position with respect to time (blue curve) and of the function $t \to t^{0.5553}$ (red curve), in $\log - \log$ scale.




(a) $t = 50$ (b) $t = 100$




(c) $t = 200$

Figure 8.7. – **Contour lines of the trait distribution during the invasion of a sexual population, initially distributed according to a Dirac distribution**, given by simulations, at (a) $t = 50$ (b) $t = 100$ (c) $t = 150$ (d) $t = 200$. The parameters are $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. The red line represents the approximation of the mean trait behind the front propagation.

**Simulations with a different value for $r$.** This supplementary material shows the simulations of the invasion of a sexual population, initially gaussian distributed. We change the value of the growth rate at low density: we take $r = 0.1$.

As for the other cases, the front propagation accelerates (Fig. 8.8). The front position is given by:

$$X^{num}(t) = y_c(r = 0.1)t^{5/4}, \qquad \text{for } y_c^{num}(r = 0.1) > 0.$$

As seen in 8.5, the constant $y_c^{num} \approx 0.3$ depends on the growth rate at low density ($r$). We have already discussed about the general case (with a general coefficient $r$). We have guessed that the front at time $t > 0$ is at the position:

$$X(t) = 4\left(\frac{\boldsymbol{\lambda}}{3}\right)^{1/2} \boldsymbol{r}^{3/4}\, t^{5/4} = \frac{2}{\sqrt{3}\, 5^{3/4}}\, t^{5/4} \approx 0.35\, t^{5/4},$$

which is consistent with our simulations. Thus the growth rate at low density $r$ is stronger, the invasion is faster: when individuals have more children, the population can invade faster areas. We present also the evolution of the invasion in Fig. 8.7.

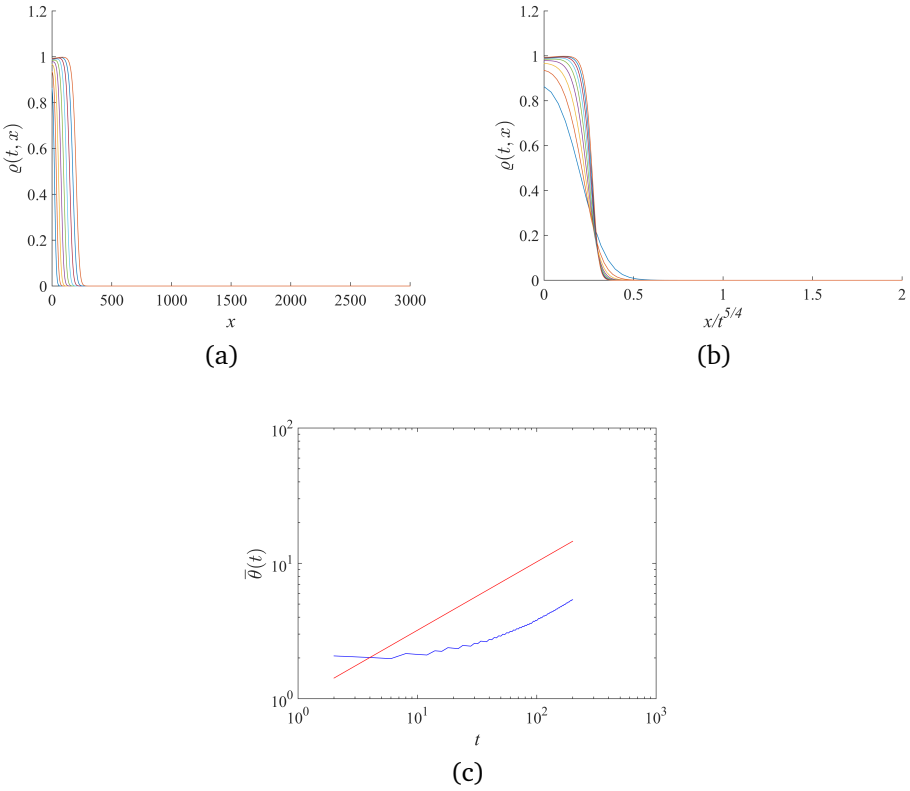
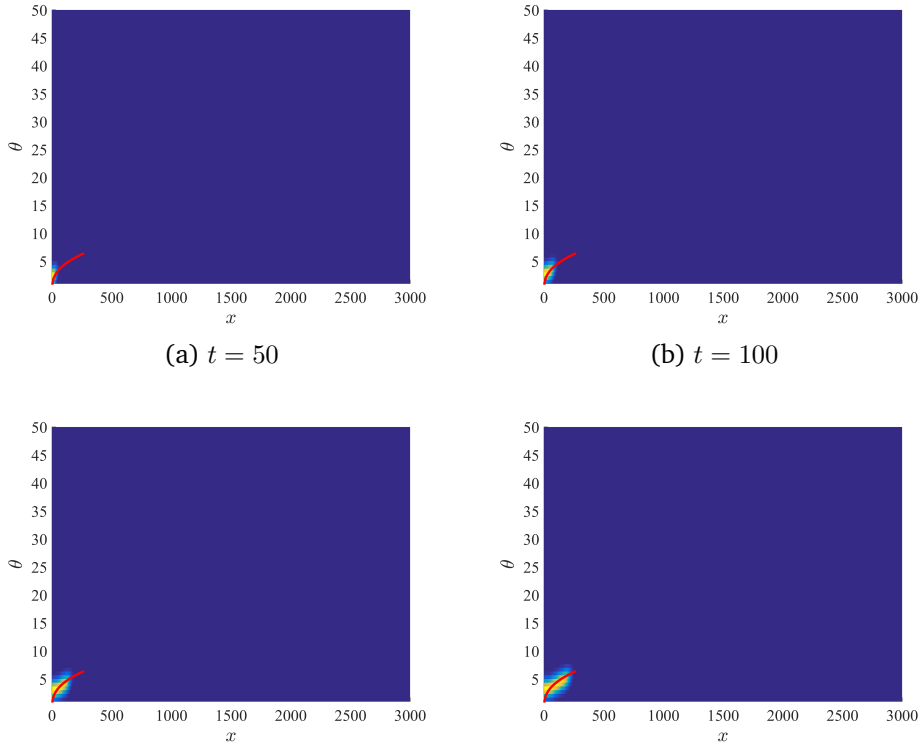



Figure 8.8. – **Simulations of the invasion of a sexual population, initially gaussian distributed, with a different growth rate at low density ($r = 0.1$)**, with parameters $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. Plots of the density of population $\varrho(t,\cdot)$ for successive fixed times at regular intervals from $t = 20$ to $t = 200$, with respect to (a) the position $x$ and (b) the auto - similar variable $xt^{-5/4}$. (c) Plot of the mean dispersive trait $\overline{\theta}^{num}(t)$ at the front position with respect to time (blue curve) and of the function $t \to t^{0.5056}$ (red curve), in $\log - \log$ scale.




(a) $t = 50$ (b) $t = 100$

(c) $t = 150$ (d) $t = 200$

Figure 8.9. – **Contour lines of the trait distribution during the invasion of a sexual population, initially gaussian distributed, with a different growth rate at low density ($r = 0.1$)**, given by simulations, at (a) $t = 50$ (b) $t = 100$ (c) $t = 150$ (d) $t = 200$. The parameters are $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. The red line represents the approximation of the mean trait behind the front propagation.

**Simulations with a different value for $\lambda$.** This last supplementary material is devoted to simulations of the invasion of a sexual population, initially gaussian distributed. We change the value of the segregrational variance: we take $\lambda = 1$.

With these parameters, we have approximated the position of the front propagation at time $t \geq 0$ by:

$$X(t) = 4\left(\frac{\boldsymbol{\lambda}}{3}\right)^{1/2} \boldsymbol{r}^{3/4}\, t^{5/4} = \frac{4}{\sqrt{3}}\, t^{5/4} \approx 2.31\, t^{5/4}.$$

This seems again consistent with the simulations (see Fig. 8.10). This shows that when the segregational variance $\lambda^2 > 0$ is strong, the individuals can have individuals with better dispersive trait, which accelerate the invasion. In Fig. 8.11,

we see that the approximation given by the associated article is accurate.

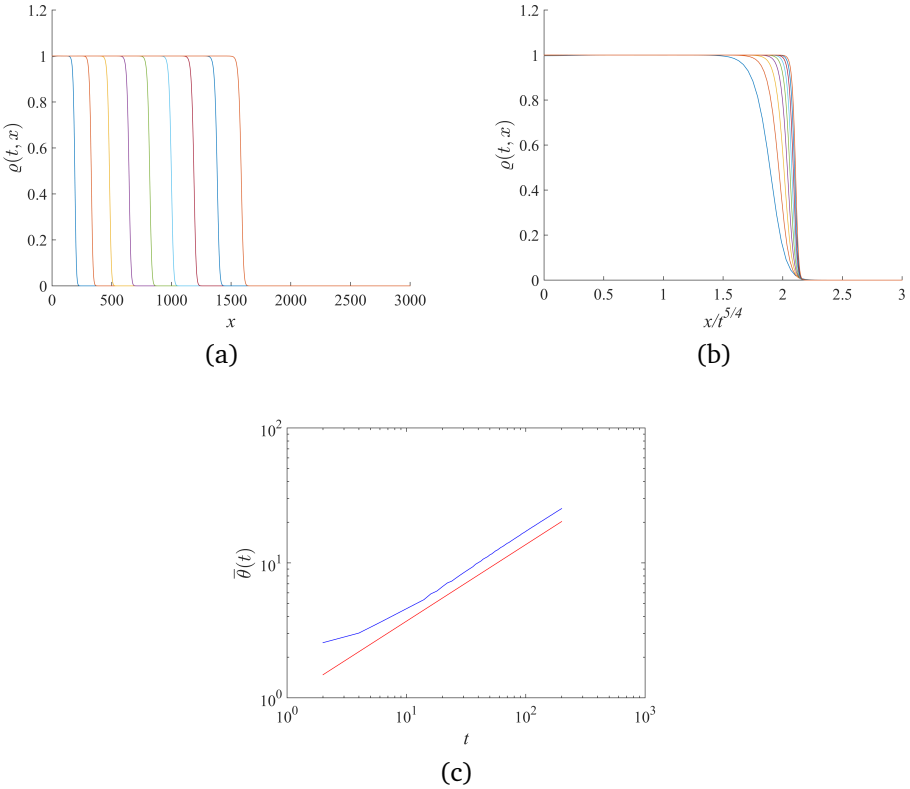
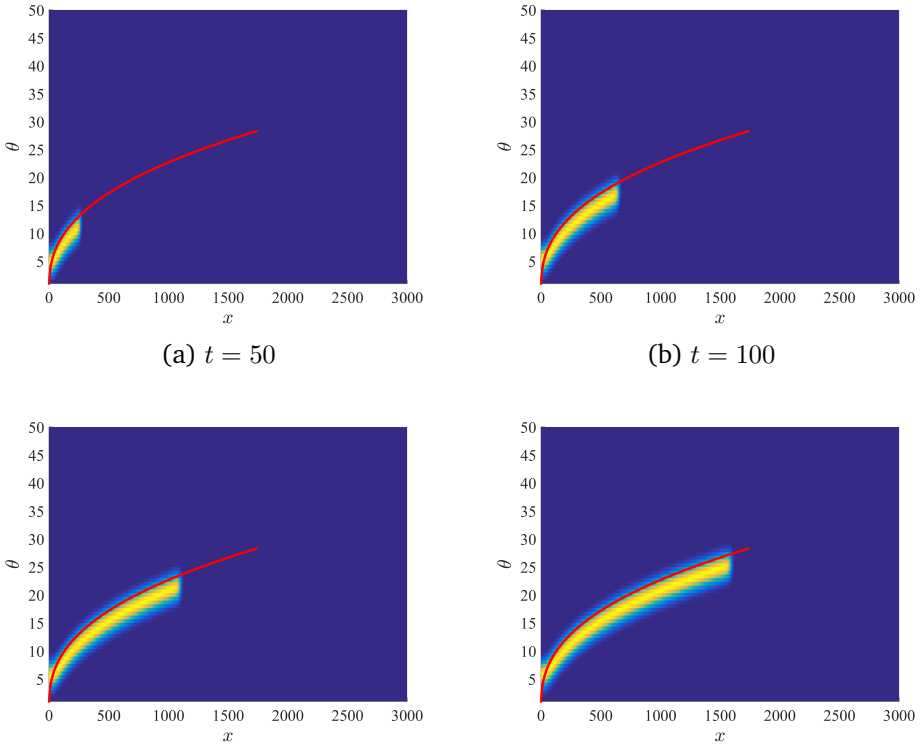



Figure 8.10. – **Simulations of the invasion of a sexual population, initially gaussian distributed, with a different segregational variance ($\lambda^2 = 1$)**, with parameters $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. Plots of the density of population $\varrho(t,\cdot)$ for successive fixed times at regular intervals from $t = 20$ to $t = 200$, with respect to (a) the position $x$ and (b) the auto - similar variable $xt^{-5/4}$. (c) Plot of the mean dispersive trait $\overline{\theta}^{num}(t)$ at the front position with respect to time (blue curve) and of the function $t \to t^{0.5678}$ (red curve), in $\log - \log$ scale.




(a) $t = 50$ (b) $t = 100$

(c) $t = 150$ (d) $t = 200$

Figure 8.11. – **Contour lines of the trait distribution during the invasion of a sexual population, initially gaussian distributed, with a different segragional variance ($\lambda^2 = 1$)**, given by simulations, at (a) $t = 50$ (b) $t = 100$ (c) $t = 150$ (d) $t = 200$. The parameters are $\delta t = 0.02$, $\delta x = 4$, $\delta\theta = 2/3$, $x_{\max} = 3000$ and $\theta_{\max} = 201$. The red line represents the approximation of the mean trait behind the front propagation.

# Bibliographie


ALEXANDER, H K; MARTIN, G; MARTIN, O Y; BONHOEFFER, S, 2014. Evolutionary rescue: linking theory for conservation and medicine. *Evolutionary applications*. Vol. 7, no. 10, p. 1161–1179 (cit. on p. 163).

ALFARO, M; BERESTYCKI, H; RAOUL, G, 2017. The effect of climate shift on a species submitted to dispersion, evolution, growth, and nonlocal competition. *SIAM Journal on Mathematical Analysis*. Vol. 49, no. 1, p. 562–596 (cit. on p. 149).

ALFARO, M; CARLES, R, 2014a. Explicit Solutions for Replicator-Mutator Equations: Extinction versus Acceleration. *SIAM Journal on Applied Mathematics*. Vol. 74, no. 6, p. 1919–1934 (cit. on pp. 58, 66, 94).

ALFARO, M; CARLES, R, 2014b. Explicit solutions for replicator-mutator equations: extinction versus acceleration. *SIAM Journal on Applied Mathematics*. Vol. 74, no. 6, p. 1919–1934 (cit. on p. 52).

ALFARO, M; CARLES, R, 2017. Replicator-mutator equations with quadratic fitness. *Proceedings of the American Mathematical Society*. Vol. 145, no. 12, p. 5315–5327 (cit. on pp. 52, 58, 66, 94, 97, 99, 141, 207).

ALFARO, M; VERUETE, M, 2018. Evolutionary branching via replicator-mutator equations. *Journal of Dynamics and Differential Equations*, p. 1–24 (cit. on pp. 52, 58, 66, 97, 99, 108, 141, 207).

ALFARO, Matthieu; COVILLE, Jerome; RAOUL, Gael, 2014. Bistable travelling waves for nonlocal reaction diffusion equations. *Disc Cont Dyn Systems A*. Vol. 34, p. 1775–1791 (cit. on p. 58).

ANCIAUX, Y; CHEVIN, L-M; RONCE, O; MARTIN, G, 2018. Evolutionary rescue over a fitness landscape. *Genetics*. Vol. 209, p. 265–279 (cit. on pp. 26, 89, 167, 181, 186–188, 238, 239, 335, 336).

ANCIAUX, Y; LAMBERT, A; RONCE, O; ROQUES, L, et al., 2019. Population persistence under high mutation rate: from evolutionary rescue to lethal mutagenesis. *Evolution*. Vol. 73, no. 8, p. 1517–1532 (cit. on p. 218).

ANDERS, H, 1998. *A History of Mathematical Statistics*. New York: John Wiley and Sons (cit. on p. 54).

ARONSON, D G; BESALA, P, 1967. Parabolic equations with unbounded coefficients. *Journal of Differential Equations*. Vol. 3, no. 1, p. 1–14 (cit. on pp. 58, 73, 117, 144).

ARONSON, D G; WEINBERGER, H F, 1978. Multidimensional nonlinear diffusion arising in population genetics. *Advances in Mathematics*. Vol. 30, no. 1, p. 33–76 (cit. on p. 284).

ARONSON, D; CRANDALL, M; PELETIER, L, 1982. Stabilization of solutions of a degenerate nonlinear diffusion problem. *Nonlinear Analysis*. Vol. 6, no. 10, p. 1001–1022 (cit. on p. 352).

AWERBUCH, T; KISZEWSKI, A; LEVINS, R, 2002. Surprise, nonlinearity and complex behaviour. *Health Impacts of Global Environmental Change: Concepts and Methods*, p. 96–102 (cit. on p. 31).

BAAKE, E; CASANOVA, A G; PROBST, S; WAKOLBINGER, A, 2019. Modelling and simulating Lenski's long-term evolution experiment. *Theoretical Population Biology* (cit. on p. 116).

BAAKE, E; WAKOLBINGER, A, 2015. Feller's Contributions to Mathematical Biology. *arXiv preprint arXiv:1501.05278* (cit. on pp. 43, 335).

BANSAYE, V; SIMATOS, F, 2015. On the scaling limits of Galton-Watson processes in varying environments. *Electronic Journal of Probability*. Vol. 20 (cit. on pp. 43, 336).

BARLES, G; MIRRAHIMI, S; PERTHAME, B, 2009. Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result. *Methods and Applications of Analysis*. Vol. 16, no. 3, p. 321–340 (cit. on pp. 99, 208).

BARTON, N H; ETHERIDGE, A M; VÉBER, A, 2017. The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*. Vol. 118, p. 50–73. ISSN 00405809. Disponible à l'adresse DOI: 10.1016/j.tpb.2017.06.001 (cit. on p. 285).

BARTON, N H; WHITLOCK, M C, 1997. The evolution of metapopulations. In: *The evolution of metapopulations*. *Metapopulation biology*. Elsevier, p. 183–210 (cit. on p. 237).

BARTON, N; ETHERIDGE, A, 2017. Establishment in a new habitat by polygenic adaptation. *Theoretical Population Biology*. Vol. 122, p. 110–127 (cit. on p. 163).

BEAUJEU-GARNIER, J, 1989. Malthus avait-il raison? *Espace Populations Sociétés*. Vol. 7, no. 3, p. 305–315 (cit. on p. 33).

BELL, G; GONZALEZ, A, 2009. Evolutionary rescue can prevent extinction following environmental change. *Ecology letters*. Vol. 12, no. 9, p. 942–948 (cit. on p. 26).

BELLOMO, N; DE ANGELIS, E; PREZIOSI, L, 2003. Multiscale modeling and mathematical problems related to tumor evolution and medical therapy. *Computational and Mathematical Methods in Medicine*. Vol. 5, no. 2, p. 111–136 (cit. on p. 49).

BENICHOU, O; CALVEZ, V; MEUNIER, N; VOITURIEZ, R, 2012. Front acceleration by dynamic selection in Fisherpopulation waves. *Physical review E*. Vol. 86, no. 4-1, p. 041908 (cit. on pp. 86, 274).

BERESTYCKI, H; HAMEL, F; NADIN, G, 2008. Asymptotic spreading in heterogeneous diffusive excitable media. *Journal of Functional Analysis*. Vol. 255, no. 9, p. 2146–2189 (cit. on p. 284).

BERESTYCKI, H; NIRENBERG, L; VARADHAN, S R S, 1994. The principal eigenvalue and maximum principle for second order elliptic operators in general domains. *Comm. Pure Appl. Math.* Vol. 47, p. 47–92 (cit. on p. 229).

BERESTYCKI, Henri; NADIN, Grégoire; PERTHAME, Benoit; RYZHIK, Lenya, 2009. The non-local Fisher-KPP equation: travelling waves and steady states. *Nonlinearity*. Vol. 22, no. 12, p. 2813–2844. http://stacks.iop.org/0951-7715/22/i=12/a=002?key=crossref.514ea7dbf016120e4b50c5ee77b5a5eb (cit. on p. 58).

BERESTYCKI, N; MOUHOT, C; RAOUL, G, 2015. Existence of self-accelerating fronts for a non-local reaction-diffsuion equations. *arXiv preprint arXiv:1512.00903* (cit. on pp. 86, 87, 276, 280, 284, 285, 304).

BERTHOULY-SALAZAR, C; RENSBURG, B J van; LE ROUX, Johannes J; VUUREN, B J van, et al., 2012. Spatial Sorting Drives Morphological Variation in the Invasive Bird, Acridotheris tristis. *PLOS ONE*. Vol. 7, no. 5, p. 1–9. Disponible à l'adresse DOI: 10.1371/journal.pone.0038145 (cit. on p. 284).

BESALA, P, 1979. Fundamental solution and Cauchy problem for a parabolic system with unbounded coefficients. *Journal of Differential Equations*. Vol. 33, no. 1, p. 26–38 (cit. on pp. 59, 220).

BIKTASHEV, V, 2014. A simple mathematical model of gradual Darwinian evolution: emergence of a Gaussian trait distribution in adaptation along a fitness gradient. *Journal of mathematical biology*. Vol. 68, no. 5, p. 1225–1248 (cit. on p. 51).

BIRZU, G; HALLATSCHEK, O; KOROLEV, K S, 2017. Neither pulled nor pushed: Genetic drift and front wandering uncover a new class of reaction-diffusion waves. *arXiv preprint arXiv:1709.01601* (cit. on p. 284).

BLACKBURN, T M; PYŠEK, P; BACHER, S; CARLTON, J T, et al., 2011. A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*. Vol. 26, no. 7, p. 333–339 (cit. on pp. 162, 165).

BLOWS, M W; HOFFMANN, A A, 2005. A reassessment of genetic limits to evolutionary change. *Ecology*. Vol. 86, no. 6, p. 1371–1384 (cit. on pp. 83, 237).

BORG, J; KLÆR, Lars Pødenphant; LECARPENTIER, C; GOLDRINGER, I, et al., 2018. Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research*. Vol. 221, p. 298–313 (cit. on p. 219).

BOUIN, E; BOURGERON, T; CALVEZ, V; CORRO, O, et al., 2018. Equilibria of quantitative genetics models beyond the Gaussian approximation I : maladaptation to a changing environment. *in preparation* (cit. on pp. 285, 296–298).

BOUIN, E; CALVEZ, V; MEUNIER, N; MIRRAHIMI, S, et al., 2012. Invasion fronts with variable motility: Phenotype selection, spatial sorting and wave acceleration. *Comptes Rendus Mathematique*. Vol. 350, no. 15–16, p. 761–766. ISSN 1631073X. Disponible à l'adresse DOI: 10.1016/j.crma.2012.09.010 (cit. on pp. 284, 285).

BOUIN, E; HENDERSON, C; RYZHIK, L, 2017. Super-linear spreading in local and non-local cane toads equations. *Journal de mathématiques Pures et Appliquées*. Vol. 108, no. 5, p. 724–750 (cit. on pp. 86, 87, 275–277, 280, 284, 285, 294, 304).

BRITANNICA, Encyclopaedia et al., 1993. *Encyclopædia britannica*. Chicago: University of Chicago (cit. on p. 33).

BROWN, J S; PAVLOVIC, N B, 1992. Evolution in heterogeneous environments: effects of migration on habitat specialization. *Evolutionary Ecology*. Vol. 6, no. 5, p. 360–382 (cit. on pp. 83, 237).

BULL, J J; SANJUAN, R; WILKE, C O, 2007. Theory of lethal mutagenesis for viruses. *Journal of Virology*. Vol. 81, no. 6, p. 2930–2939 (cit. on pp. 177, 186).

BULL, J J; WILKE, C O, 2008. Lethal mutagenesis of bacteria. *Genetics*. Vol. 180, no. 2, p. 1061–1070 (cit. on pp. 177, 186).

BULMER, M G, 1972. The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genetical Research*. Vol. 19, no. 01, p. 17. ISSN 0016-6723, 1469-5073. ISSN 0016-6723, 1469-5073. Disponible à l'adresse DOI: 10.1017/S0016672300014221 (cit. on p. 285).

BURGER, R, 1991. Moments, cumulants, and polygenic dynamics. *Journal of Mathematical Biology*. Vol. 30, no. 2, p. 199–213 (cit. on pp. 60, 169).

BÜRGER, R, 1986. On the maintenance of genetic variation: global analysis of Kimura's continuum-of-alleles model. *Journal of mathematical biology*. Vol. 24, no. 3, p. 341–351 (cit. on p. 344).

BURTON, O J; PHILLIPS, B L; TRAVIS, J M J, 2010. Trade-offs and the evolution of life-histories during range expansion. *Ecology Letters*. Vol. 13, no. 10, p. 1210–1220. Disponible à l'adresse DOI: 10.1111/j.1461-0248.2010.01505.x (cit. on p. 305).

CALVEZ, V; CREVAT, J; DEKENS, L; FABRÈGES, B, et al., 2019. Influence of the mode of reproduction on dispersal evolution during species invasion. *ESAIM: Proceedings and Surveys*. Accepted (cit. on pp. 87, 272, 285, 289, 294, 295).

CALVEZ, V; GARNIER, J; PATOUT, F, 2019. Asymptotic analysis of a quantitative genetics model with nonlinear integral operator. *Journal de l'Ecole polytechnique — Mathématiques*. Vol. 6, p. 537–579. ISSN 2270-518X. Disponible à l'adresse DOI: 10.5802/jep.100 (cit. on pp. 285, 295, 296, 298, 300).

CALVEZ, V; HENDERSON, C; MIRRAHIMI, S; TURANOVA, O, et al., 2018. *Non-local competition slows down front acceleration during dispersal evolution*. Disponible à l'adresse arXiv: 1810.07634 [math.AP] (cit. on pp. 86, 87, 275–277, 284, 285, 294, 304).

CANTRELL, R S; COSNER, C, 2003. *Spatial Ecology via Reaction-Diffusion Equations*. John Wiley & Sons Ltd, Chichester, UK (cit. on p. 213).

CAQUET, T; GASCUEL, C; TIXIER-BOICHARD, M, 2020. *Agroécologie: des recherches pour la transition des filières et des territoires*. Quae (cit. on p. 219).

CASANOVA, A G; KURT, N; WAKOLBINGER, A; YUAN, L, 2016. An individual-based model for the Lenski experiment, and the deceleration of the relative fitness. *Stochastic Processes and their Applications*. Vol. 126, no. 8, p. 2211–2252 (cit. on p. 116).

CHABROWSKI, J, 1970. Sur la construction de la solution fondamentale de l'équation parabolique aux coefficients non bornés. In: *Sur la construction de la solution fondamentale de l'équation parabolique aux coefficients non bornés*. *Colloquium Mathematicum*. Vol. 1, p. 141–148 (cit. on pp. 58, 73, 144).

CHAMPAGNAT, N; FERRIÈRE, R; MÉLÉARD, S, 2006. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical population biology*. Vol. 69, no. 3, p. 297–321 (cit. on p. 94).

COLAUTTI, R I; ALEXANDER, J M; DLUGOSCH, K M; KELLER, S R, et al., 2017. Invasions and extinctions through the looking glass of evolutionary ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*. Vol. 372, no. 1712, p. 20160031 (cit. on pp. 27, 75, 162).

CRÉCY-LAGARD, V A de; BELLALOU, J; MUTZEL, R; MARLIÈRE, P, 2001. Long term adaptation of a microbial population to a permanent metabolic constraint: overcoming thymineless death by experimental evolution of *Escherichia coli*. *BMC biotechnology*. Vol. 1, no. 1, p. 10 (cit. on p. 96).

CWYNAR, L C; MACDONALD, G M, 1987. Geographical variation of lodgepole pine in relation to population history. *The American Naturalist*. Vol. 129, no. 3, p. 463–469 (cit. on p. 284).

DAWKINS, R, 1995. *River out of Eden*. Londres, Weidenfeld and Nicolson (cit. on p. 54).

DAWSON, D A, 2017. Introductory lectures on stochastic population systems. *arXiv preprint arXiv:1705.03781* (cit. on pp. 43, 336).

DÉBARRE, F; RONCE, O; GANDON, S, 2013. Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology*. Vol. 26, no. 6, p. 1185–1202 (cit. on pp. 29, 76, 162, 163, 237–239).

DENNEHY, J J; FRIEDENBERG, N A; MCBRIDE, R C; HOLT, R D, et al., 2010. Experimental evidence that source genetic variation drives pathogen emergence. *Proceedings of the Royal Society B: Biological Sciences*. Vol. 277, no. 1697, p. 3113–3121 (cit. on pp. 24, 27, 185).

DEPEW, D J; WEBER, B H, 1996. Darwinism evolving: Systems dynamics and the genealogy of natural selection (cit. on pp. 21, 22).

DESAI, M M; FISHER, D S, 2011. The balance between mutators and nonmutators in asexual populations. *Genetics*. Vol. 188, no. 4, p. 997–1014 (cit. on pp. 72, 141).

DIEKMANN, O; JABIN, P-E; MISCHLER, S; PERTHAME, B, 2005. The dynamics of adaptation: an illuminating example and a Hamilton–Jacobi approach. *Theoretical population biology*. Vol. 67, no. 4, p. 257–271 (cit. on pp. 99, 208).

DRURY, K L S; DRAKE, J M; LODGE, D M; DWYER, G, 2007. Immigration events dispersed in space and time: Factors affecting invasion success. *Ecological Modelling*. Vol. 206, p. 63–78 (cit. on pp. 27, 76, 163).

ELENA, S F; LENSKI, R E, 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*. Vol. 4, no. 6, p. 457–469 (cit. on pp. 24, 26, 356).

ELENA, S; SANJUÁN, R, 2003. Climb Every Mountain? *Science*. Vol. 302, no. 5653, p. 2074–2075. ISSN 0036-8075. Disponible à l'adresse DOI: 10.1126/science.1093165 (cit. on p. 95).

EPSTEIN, C L; MAZZEO, R, 2010. Wright–Fisher diffusion in one dimension. *SIAM Journal on Mathematical Analysis*. Vol. 42, no. 2, p. 568–608 (cit. on p. 42).

EPSTEIN, C L; MAZZEO, R, 2013. *Degenerate diffusion operators arising in population biology*. Princeton University Press. No. 185 (cit. on pp. 337, 352).

EWENS, W J, 2012. *Mathematical Population Genetics 1: Theoretical Introduction*. Springer Science & Business Media (cit. on p. 332).

FANG, J; ZHAO, X, 2011. Monotone wavefronts of the nonlocal Fisher–KPP equation. *Nonlinearity*. Vol. 24, no. 11, p. 3043 (cit. on pp. 58, 274, 284).

FAVINI, A; MARINOSCHI, G, 2012. *Degenerate nonlinear diffusion equations*. Springer (cit. on pp. 337, 352).

FAYE, G; HOLZER, M, 2015. Modulated traveling fronts for a nonlocal Fisher-KPP equation: a dynamical systems approach. *J. Diff. Equations*. Vol. 258, no. 7, p. 2257–2289 (cit. on p. 58).

FELLER, W, 1951. Diffusion processes in genetics. In: *Diffusion processes in genetics. Proceedings of the second Berkeley symposium on mathematical statistics and probability* (cit. on pp. 42, 43, 333, 335).

FISHER, R, 1930. *The Genetical Theory of Natural Selection*. Oxford University Press (cit. on pp. 59, 60, 238, 275).

FISHER, R A, 1919. The correlation between relatives on the supposition of Mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*. Vol. 52, no. 2, p. 399–433 (cit. on p. 285).

FONG, Stephen S; PALSSON, Bernhard Ø, 2004. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nature genetics*. Vol. 36, no. 10, p. 1056 (cit. on p. 96).

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS, 2018. *The 10 Elements of Agroecology. Guiding the Transition to Sustainable Food and Agricultural Systems*. FAO (cit. on p. 219).

FRAÏSSE, C; WELCH, J J, 2019. The distribution of epistasis on simple fitness landscapes. *Biology letters*. Vol. 15, no. 4, p. 20180881 (cit. on p. 96).

FREEDMAN, H I; WALTMAN, P, 1977. Mathematical models of population interactions with dispersal. I: stability of two habitats with and without a predator. *SIAM Journal on Applied Mathematics*. Vol. 32, no. 3, p. 631–648 (cit. on pp. 83, 237).

FREIDLIN, Mark Iosifovich, 1985. *Functional Integration and Partial Differential Equations*. Princeton University Press. No. 109 (cit. on pp. 336, 338).

FRIEDMAN, A, 1964. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, NJ (cit. on p. 120).

FURRER, R D; PASINELLI, G, 2016. Empirical evidence for source–sink populations: a review on occurrence, assessments and implications. *Biological Reviews*. Vol. 91, no. 3, p. 782–795 (cit. on pp. 26, 162).

GANDON, S; MIRRAHIMI, S, 2017a. A Hamilton–Jacobi method to describe the evolutionary equilibria in heterogeneous environments and with non-vanishing effects of mutations. *Comptes Rendus Mathematique*. Vol. 355, no. 2, p. 155–160 (cit. on p. 29).

GANDON, S; MIRRAHIMI, S, 2017b. A Hamilton–Jacobi method to describe the evolutionary equilibria in heterogeneous environments and with non-vanishing effects of mutations. *Comptes Rendus Mathematique*. Vol. 355, no. 2, p. 155–160 (cit. on pp. 83, 99, 208, 237, 238).

GARNIER, J; ROQUES, L; HAMEL, F, 2012. Success rate of a biological invasion in terms of the spatial distribution of the founding population. *Bulletin of Mathematical Biology*. Vol. 74, p. 453–473 (cit. on pp. 76, 163).

GENIEYS, S; VOLPERT, V; AUGER, P, 2006a. Pattern and waves for a model in population dynamics with nonlocal consumption of resources. *Mathematical Modelling of Natural Phenomena*. Vol. 1, no. 1, p. 63–80 (cit. on pp. 274, 284).

GENIEYS, S; VOLPERT, V; AUGER, V, 2006b. Pattern and waves for a model in population dynamics with nonlocal consumption of resources. *Math. Model. Nat. Phenom.* Vol. 1, p. 63–80 (cit. on p. 58).

GERRISH, P J; SNIEGOWSKI, P D, 2012. Real time forecasting of near-future evolution. *Journal of the Royal Society Interface*. Vol. 9, no. 74, p. 2268–2278 (cit. on pp. 94, 169).

GERRISH, Philip J; COLATO, Alexandre; PERELSON, Alan S; SNIEGOWSKI, Paul D, 2007. Complete genetic linkage can subvert natural selection. *Proceedings of the National Academy of Sciences*. Vol. 104, no. 15, p. 6266–6271 (cit. on pp. 72, 141).

GIL, M-E; HAMEL, F; MARTIN, G; ROQUES, L, 2017. Mathematical properties of a class of integro-differential models from population genetics. *SIAM J Appl Math*. Vol. 77, no. 4, p. 1536–1561 (cit. on pp. 66, 72, 94, 141, 154, 343, 359, 369).

GIL, M-E; HAMEL, F; MARTIN, G; ROQUES, L, 2019. Dynamics of fitness distributions in the presence of a phenotypic optimum: an integro-differential approach. *Nonlinearity*, p. in press (cit. on pp. 72, 97, 100, 141, 343, 347).

GOMULKIEWICZ, R; HOLT, R D, 1995. When does evolution by natural selection prevent extinction? *Evolution*. Vol. 49, no. 1, p. 201–207 (cit. on pp. 34, 336).

GOMULKIEWICZ, R; HOLT, R D; BARFIELD, M, 1999. The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. *Theoretical Population Biology*. Vol. 55, no. 3, p. 283–296 (cit. on pp. 41, 75, 162, 164, 238).

GOMULKIEWICZ, R; HOLT, R D; BARFIELD, M; NUISMER, S L, 2010. Genetics, adaptation, and invasion in harsh environments. *Evolutionary Applications*. Vol. 3, no. 2, p. 97–108 (cit. on pp. 26, 76, 163, 238).

GOOD, B H; DESAI, M M, 2015. The impact of macroscopic epistasis on long-term evolutionary dynamics. *Genetics*. Vol. 85, p. 177–190 (cit. on pp. 96, 109, 114).

GOURLEY, S A, 2000. Travelling front solutions of a nonlocal Fisher equation. *Journal of mathematical biology*. Vol. 41, no. 3, p. 272–284 (cit. on pp. 58, 274, 284).

GUEZO, N; FIOGBE, E; TOBIAS, M, 2017. Evaluation of sodium chloride (NaCl) effects on water hyacinth Eichhornia crassipes development: Preliminary results. *EWASH & TI*. Vol. 1, no. 4, p. 34–40 (cit. on p. 24).

HADELER, K, 1981a. Diffusion in Fisher's population model. *The Rocky Mountain Journal of Mathematics*, p. 39–45 (cit. on p. 51).

HADELER, K P, 1981b. Stable polymorphisms in a selection model with mutation. *SIAM Journal on Applied Mathematics*. Vol. 41, no. 1, p. 1–7 (cit. on p. 51).

HAMEL, F; LAVIGNE, F; MARTIN, G; ROQUES, L, 2020. Dynamics of adaptation in an anisotropic phenotype-fitness landscape. *Nonlinear Analysis: Real World Applications*. Vol. 54, p. 103107 (cit. on pp. 19, 65, 72, 73, 78, 139, 141–145, 149, 153, 154, 156, 157, 207, 210, 218, 219, 235, 242, 244, 246, 254, 257–259).

HAMEL, F; RYZHIK, L, 2014a. On the nonlocal Fisher–KPP equation: steady states, spreading speed and global bounds. *Nonlinearity*. Vol. 27, no. 11, p. 2735 (cit. on p. 58).

HAMEL, F; RYZHIK, L, 2014b. On the nonlocal Fisher–KPP equation: steady states, spreading speed and global bounds. *Nonlinearity*. Vol. 27, no. 11, p. 2735 (cit. on pp. 274, 284).

HARMAND, N; GALLET, R; JABBOUR-ZAHAB, R; MARTIN, G, et al., 2017. Fisher's geometrical model and the mutational patterns of antibiotic resistance across dose gradients. *Evolution*. Vol. 71, no. 1, p. 23–37 (cit. on p. 187).

HARMAND, N; GALLET, R; MARTIN, G; LENORMAND, T, 2018. Evolution of bacteria specialization along an antibiotic dose gradient. *Evolution letters*. Vol. 2, no. 3, p. 221–232 (cit. on p. 24).

HASTINGS, A, 1983. Can spatial variation alone lead to selection for dispersal? *Theoretical Population Biology*. Vol. 24, no. 3, p. 244–251 (cit. on pp. 40, 83, 237).

HIETPAS, R T; BANK, C; JENSEN, J D; BOLON, D NA, 2013. Shifting fitness landscapes in response to altered environments. *Evolution*. Vol. 67, no. 12, p. 3512–3522 (cit. on pp. 167, 237, 239).

HISLOP, P D; SIGAL, I M, 2012. *Introduction to spectral theory: With applications to Schrödinger operators*. Springer Science & Business Media (cit. on p. 108).

HOLLING, C S, 1959. The components of predation as revealed by a study of small-mammal predation of the European pine sawfly. *The Canadian Entomologist*. Vol. 91, no. 5, p. 293–320 (cit. on p. 31).

HOLT, R D, 1985. Population dynamics in two-patch environments: some anomalous consequences of an optimal habitat distribution. *Theoretical population biology*. Vol. 28, no. 2, p. 181–208 (cit. on pp. 26, 40, 83, 237).

HOLT, R D, 1996a. Adaptive evolution in source-sink environments: direct and indirect effects of density-dependence on niche evolution. *Oikos*, p. 182–192 (cit. on p. 40).

HOLT, R D, 1996b. Demographic constraints in evolution: towards unifying the evolutionary theories of senescence and niche conservatism. *Evolutionary Ecology*. Vol. 10, no. 1, p. 1–11 (cit. on pp. 38, 242).

HOLT, R D, 2009. Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*. Vol. 106, no. Supplement 2, p. 19659–19665 (cit. on p. 163).

HOLT, R D; BARFIELD, M; GOMULKIEWICZ, R, 2004. Temporal variation can facilitate niche evolution in harsh sink environments. *The American Naturalist*. Vol. 164, no. 2, p. 187–200 (cit. on pp. 162–164).

HOLT, R D; BARFIELD, M; GOMULKIEWICZ, R, 2005. Theories of niche conservatism and evolution: could exotic species be potential tests. In: Sinauer Associates Sunderland, MA, p. 259–290 (cit. on p. 162).

HOLT, R D; GOMULKIEWICZ, R, 1997. How does immigration influence local adaptation? A reexamination of a familiar paradigm. *The American Naturalist*. Vol. 149, no. 3, p. 563–572 (cit. on pp. 83, 237).

HOLT, R D; GOMULKIEWICZ, R; BARFIELD, M, 2003. The phenomenology of niche evolution via quantitative traits in a “black-hole" sink. *Proceedings of the Royal Society of London B: Biological Sciences*. Vol. 270, no. 1511, p. 215–224 (cit. on pp. 162, 163).

HUFBAUER, R A; FACON, R; RAVIGNÉ, V; TURGEON, J, et al., 2012. Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*. Vol. 5, no. 1, p. 89–101 (cit. on p. 187).

IGLESIAS, S F; MIRRAHIMI, S, 2019. Selection and mutation in a shifting and fluctuating environment (cit. on p. 149).

JANSEN, M; COORS, A; STOKS, R; DE MEESTER, L, 2011. Evolutionary ecotoxicology of pesticide resistance: a case study in Daphnia. *Ecotoxicology*. Vol. 20, no. 3, p. 543–551 (cit. on pp. 27, 75, 162).

KAWECKI, T J, 1995. Demography of source—sink populations and the evolution of ecological niches. *Evolutionary Ecology*. Vol. 9, no. 1, p. 38–44 (cit. on pp. 38, 83, 237).

KAWECKI, T J, 2000. Adaptation to marginal habitats: contrasting influence of the dispersal rate on the fate of alleles with small and large effects. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. Vol. 267, no. 1450, p. 1315–1320 (cit. on pp. 83, 237).

KAWECKI, T J; HOLT, R D, 2002. Evolutionary consequences of asymmetric dispersal rates. *The American Naturalist*. Vol. 160, no. 3, p. 333–347 (cit. on pp. 30, 41, 83, 237).

KIMURA, M, 1954. Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. *Genetics*. Vol. 39, no. 3, p. 280 (cit. on pp. 42, 343).

KIMURA, M, 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences*. Vol. 54, no. 3, p. 731–736 (cit. on pp. 47, 55, 167, 240, 275).

KIRKPATRICK, M; BARTON, N H, 1997. Evolution of a species'range. *The American Naturalist*. Vol. 150, p. 1–23 (cit. on p. 162).

KOMAROVA, N, 2004. Replicator–mutator equation, universality property and population dynamics of learning. *Journal of Theoretical Biology*. Vol. 230, no. 2, p. 227–239 (cit. on pp. 51, 52).

KRYAZHIMSKIY, S; TKAČIK, G; PLOTKIN, J B, 2009. The dynamics of adaptation on correlated fitness landscapes. *Proceedings of the National Academy of Sciences*. Vol. 106, no. 44, p. 18638–18643 (cit. on pp. 89, 95).

LACROIX, R A; SANDBERG, T E; O'BRIEN, E J; UTRILLA, J, et al., 2015. Use of Adaptive Laboratory Evolution To Discover Key Mutations Enabling Rapid Growth of *Escherichia coli* K-12 MG1655 on Glucose Minimal Medium. *Applied and Environmental Microbiology*. Vol. 81, no. 1, p. 17–30. ISSN 0099-2240. Disponible à l'adresse DOI: 10.1128/AEM.02246-14 (cit. on p. 96).

LAMARCK, J-B de, 1873. *Philosophie zoologique*. F. Savy (cit. on p. 21).

LAMBERT, A, 2008. Population dynamics and random genealogies. *Stochastic Models*. Vol. 24, no. S1, p. 45–163 (cit. on pp. 43, 335).

LANDE, R, 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics*. Vol. 94, no. 1, p. 203–215 (cit. on pp. 55, 167, 240).

LANGE, K, 1978. Central limit theorems of pedigrees. *Journal of Mathematical Biology*. Vol. 6, no. 1, p. 59–66. ISSN 0303-6812, 1432-1416. ISSN 0303-6812, 1432-1416. Disponible à l'adresse DOI: 10.1007/BF02478517 (cit. on p. 285).

LAVIGNE, F; MARTIN, G; ANCIAUX, Y; PAPAIX, J, et al., 2020. When sinks become sources: adaptive colonization in asexuals. *Evolution*. Vol. 74, p. 29–42. Recommended by PCI (cit. on pp. 19, 76, 84, 88, 219, 235, 239, 244).

LAVIGNE, F; ROQUES, L, 2020. Extinction times of an inhomogenous Feller diffusion process: A PDE approach. *Expositiones Mathematicae* (cit. on pp. 19, 43).

LENORMAND, T, 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*. Vol. 17, no. 4, p. 183–189 (cit. on pp. 76, 83, 163, 164, 237, 356).

LENSKI, R E, 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME journal*. Vol. 11, no. 10, p. 2181 (cit. on pp. 24, 26, 70, 356).

LENSKI, R E; ROSE, M R; SIMPSON, S C; TADLER, S C, 1991. Experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *The American Naturalist*. Vol. 138, no. 6, p. 1315–1341 (cit. on pp. 24, 26, 33, 70, 100, 111, 142, 356).

LENSKI, R E; TRAVISANO, M, 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences*. Vol. 91, no. 15, p. 6808–6814 (cit. on pp. 70, 96, 100, 111, 142, 356).

LESLIE, P, 1945. On the use of matrices in certain population mathematics. *Biometrika*. Vol. 33, no. 3, p. 183–212 (cit. on p. 38).

LEVIN, S A, 1974. Dispersion and population interactions. *The American Naturalist*. Vol. 108, no. 960, p. 207–228 (cit. on pp. 83, 237).

LEVINS, R, 1966. The strategy of model building in population biology. *American scientist*. Vol. 54, no. 4, p. 421–431 (cit. on p. 31).

LOCKWOOD, J L; CASSEY, P; BLACKBURN, T, 2005. The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*. Vol. 20, no. 5, p. 223–228 (cit. on p. 186).

LOREAU, M; DAUFRESNE Tand Gonzalez, A; GRAVEL, D; GUICHARD, F, et al., 2013. Unifying sources and sinks in ecology and Earth sciences. *Biological Reviews*. Vol. 88, no. 2, p. 365–379 (cit. on pp. 25, 26, 162).

LORZ, A; MIRRAHIMI, S; PERTHAME, B, 2011. Dirac mass dynamics in multidimensional nonlocal parabolic equations. *Communications in Partial Differential Equations*. Vol. 36, no. 6, p. 1071–1098 (cit. on pp. 99, 208).

LUNARDI, A, 1998. Schauder theorems for linear elliptic and parabolic problems with unbounded coefficients in $\mathbb{R}^n$. *Studia Mathematica*. Vol. 128, p. 171–198 (cit. on pp. 58, 117).

MACLEAN, R C; HALL, A R; PERRON, G G; BUCKLING, A, 2010. The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nature Reviews Genetics*. Vol. 11, no. 6, p. 405 (cit. on p. 163).

MARTIN, G, 2014. Fisher's geometrical model emerges as a property of complex integrated phenotypic networks. *Genetics*. Vol. 197, no. 1, p. 237–255 (cit. on pp. 95, 97, 167, 241, 346).

MARTIN, G; ELENA, S F; LENORMAND, T, 2007. Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nature Genetics*. Vol. 39, no. 4, p. 555 (cit. on pp. 95, 163).

MARTIN, G; LENORMAND, T, 2006a. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution*. Vol. 60, no. 5, p. 893–907 (cit. on p. 99).

MARTIN, G; LENORMAND, T, 2006b. The fitness effect of mutations across environments: a survey in light of fitness landscape models. *Evolution*. Vol. 60, no. 12, p. 2413–2427 (cit. on pp. 96, 238).

MARTIN, G; LENORMAND, T, 2015. The fitness effect of mutations across environments: Fisher's geometrical model with multiple optima. *Evolution*. Vol. 69, no. 6, p. 1433–1447 (cit. on pp. 56, 95, 188, 207, 239).

MARTIN, G; ROQUES, L, 2016. The Non-stationary Dynamics of Fitness Distributions: Asexual Model with Epistasis and Standing Variation. *Genetics*. Vol. 204, no. 4, p. 1541–1558 (cit. on pp. 57, 60, 84, 89, 97, 98, 106–109, 142, 154, 164, 168–171, 183, 188, 190, 218, 235, 244, 270, 347, 353–355).

MAUPERTUIS, P de, 1745. *Vénus physique* (cit. on p. 21).

MCCANDLISH, D M; STOLTZFUS, A, 2014. Modeling evolution using the probability of fixation: history and implications. *The Quarterly Review of Biology*. Vol. 89, no. 3, p. 225–252 (cit. on p. 89).

MESZÉNA, G; CZIBULA, I; GERITZ, S, 1997. Adaptive dynamics in a 2-patch environment: a toy model for allopatric and parapatric speciation. *Journal of Biological Systems*. Vol. 5, no. 02, p. 265–284 (cit. on pp. 83, 237).

MESZÉNA, G; GYLLENBERG, M; JACOBS, F J; METZ, Johan A J, 2005. Link between population dynamics and dynamics of Darwinian evolution. *Physical review letters*. Vol. 95, no. 7, p. 078105 (cit. on p. 33).

MIRRAHIMI, S, 2012. Adaptation and migration of a population between patches. *arXiv preprint arXiv:1204.0801* (cit. on p. 238).

MIRRAHIMI, S; GANDON, S, 2019. Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration. *Genetics* (cit. on p. 29).

MIRRAHIMI, S; GANDON, S, 2020. Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration. *Genetics*. Vol. 214, no. 2, p. 479–491 (cit. on pp. 80, 83, 207, 219, 237, 238).

MIRRAHIMI, S; RAOUL, G, 2013. Dynamics of sexual populations structured by a space variable and a phenotypical trait. *Theoretical population biology*. Vol. 84, p. 87–103 (cit. on pp. 284, 285).

MORSE, S S, 2001. Factors in the emergence of infectious diseases. In: *Factors in the emergence of infectious diseases. Plagues and politics*. Springer, p. 8–26 (cit. on p. 186).

NOVELLA, I S; DUARTE, E A; ELENA, S F; MOYA, A, et al., 1995. Exponential increases of RNA virus fitness during large population transmissions. *Proceedings of the National Academy of Sciences*. Vol. 92, no. 13, p. 5841–5844 (cit. on p. 96).

NOWAK, M; KOMAROVA, N; NIYOGI, P, 2002. Computational and evolutionary aspects of language. *Nature*. Vol. 417, no. 6889, p. 611 (cit. on p. 52).

ORR, H A, 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution*. Vol. 52, no. 4, p. 935–949 (cit. on p. 207).

ORR, H A; UNCKLESS, R L, 2014. The population genetics of evolutionary rescue. *PLoS Genetics*. Vol. 10, no. 8, p. e1004551 (cit. on p. 186).

PAGE, K; NOWAK, M, 2002. Unifying evolutionary dynamics. *Journal of theoretical biology*. Vol. 219, no. 1, p. 93–98 (cit. on p. 51).

PATTISON, S; MITCHELL, C; LADE, S; LEONG, T, et al., 2017. Early relapses after adjuvant chemotherapy suggests primary chemoresistance in diffuse gastric cancer. *PloS one*. Vol. 12, no. 9 (cit. on p. 24).

PEISCHL, S; DUPANLOUP, I; KIRKPATRICK, M; EXCOFFIER, L, 2013. On the accumulation of deleterious mutations during range expansions. *Molecular Ecology*. Vol. 22, no. 24, p. 5972–5982. ISSN 09621083. Disponible à l'adresse DOI: 10.1111/mec.12524 (cit. on p. 305).

PEISCHL, S; GILBERT, K J, 2020. Evolution of dispersal can rescue populations from expansion load. *The American Naturalist*. Vol. 195, no. 2, p. 000–000 (cit. on p. 305).

PERFEITO, L; SOUSA, A; BATAILLON, T; GORDO, I, 2014. Rates of fitness decline and rebound suggest pervasive epistasis. *Evolution*. Vol. 68, no. 1, p. 150–162 (cit. on p. 96).

PERRON, G G; GONZALEZ, A; BUCKLING, A, 2007. Source-sink dynamics shape the evolution of antibiotic resistance and its pleiotropic fitness cost. *Proceedings of the Royal Society of London B: Biological Sciences*. Vol. 274, no. 1623, p. 2351–2356 (cit. on p. 185).

PERTHAME, B, 2009. Quelques équations de transport apparaissant en biologie. *Boletín SEMA*. No. 28 (cit. on p. 49).

PERTHAME, B; BARLES, G, 2008. Dirac concentrations in Lotka-Volterra parabolic PDEs. *Indiana University Mathematics Journal*, p. 3275–3301 (cit. on pp. 99, 208).

PHILLIPS, B; BROWN, G; WEBB, J; SHINE, R, 2006. Invasion and the evolution of speed in toads (cit. on p. 284).

PROTTER, M H; WEINBERGER, H F, 1967. *Maximum Principles in Differential Equations*. Prentice-Hall, Englewood Cliffs, NJ (cit. on pp. 58, 117, 118).

PROVINE, W B, 2001. *The origins of theoretical population genetics: With a new afterword*. University of Chicago Press (cit. on p. 23).

PULLIAM, H R, 1988. Sources, sinks, and population regulation. *The American Naturalist*. Vol. 132, no. 5, p. 652–661 (cit. on pp. 26, 83, 162, 237).

RAOUL, G, 2017. Macroscopic limit from a structured population model to the Kirkpatrick-Barton model. *arXiv:1706.04094 [math]*. http://arxiv.org/abs/1706.04094. arXiv: 1706.04094 (cit. on p. 285).

REX CONSORTIUM, 2013. Heterogeneity of selection and the evolution of resistance. *Trends in Ecology & Evolution*. Vol. 28, no. 2, p. 110–118 (cit. on p. 187).

RONCE, O; KIRKPATRICK, M, 2001. When sources become sinks: migrational meltdown in heterogeneous habitats. *Evolution*. Vol. 55, no. 8, p. 1520–1531 (cit. on pp. 29, 237, 238).

ROQUES, L; PATOUT, F; BONNEFON, O; MARTIN, G, 2020. Adaptation in general temporally changing environments. *arXiv preprint arXiv:2002.09542* (cit. on pp. 72, 139, 143, 149–152, 154, 207, 235).

ROQUES, Lionel, 2013. *Modèles de réaction-diffusion pour l'écologie spatiale*. Editions Quae (cit. on p. 137).

ROSENZWEIG, F; SHERLOCK, G, 2014. Experimental evolution: prospects and challenges. *Genomics*. Vol. 104, no. 6 0 0, p. v (cit. on p. 94).

SCHOUSTRA, S; HWANG, S; KRUG, J; VISSER, J A GM de, 2016. Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. *Proceedings of the Royal Society B: Biological Sciences*. Vol. 283, no. 1837, p. 20161376 (cit. on p. 95).

SEGEL, L A; LEVIN, S A, 1976. Application of nonlinear stability theory to the study of the effects of diffusion on predator-prey interactions. In: *Application of nonlinear stability theory to the study of the effects of diffusion on predator-prey interactions*. *AIP Conference Proceedings*. Vol. 27, p. 123–152. No. 1 (cit. on p. 30).

SHI, J; ABID, A D; KENNEDY, I M; HRISTOVA, K R, et al., 2011. To duckweeds (Landoltia punctata), nanoparticulate copper oxide is more inhibitory than the soluble copper in the bulk solution. *Environmental Pollution*. Vol. 159, no. 5, p. 1277–1282 (cit. on p. 24).

SHINE, R; BROWN, G P; PHILLIPS, B L, 2011. An evolutionary process that assembles phenotypes through space rather than through time. *Proceedings of the National Academy of Sciences*. Vol. 108, no. 14, p. 5708–5711 (cit. on p. 284).

SHMIDA, A; ELLNER, S, 1984. Coexistence of plant species with similar niches. *Vegetatio*. Vol. 58, no. 1, p. 29–55 (cit. on p. 26).

SIRAKOV, B, 2009. Some estimates and maximum principles for weakly coupled systems of elliptic PDE. *Nonlinear Analysis: Theory, Methods and Applications*. Vol. 70, no. 8, p. 3039–3046 (cit. on pp. 59, 228).

SNIEGOWSKI, P D; GERRISH, P J, 2010. Beneficial mutations and the dynamics of adaptation in asexual populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*. Vol. 365, no. 1544, p. 1255–1263 (cit. on pp. 72, 94, 141).

SOKURENKO, E V; GOMULKIEWICZ, R; DYKHUIZEN, D E, 2006. Source-sink dynamics of virulence evolution. *Nature Reviews Microbiology*. Vol. 4, no. 7, p. 548 (cit. on pp. 26–28, 75, 162).

STADLER, P F; SCHUSTER, P, 1992. Mutation in autocatalytic reaction networks. *Journal of mathematical biology*. Vol. 30, no. 6, p. 597–632 (cit. on p. 52).

SWEERS, G, 1992. Strong positivity in C () for elliptic systems. *Math. z*. Vol. 209, no. 2, p. 251–271 (cit. on pp. 59, 82, 213).

TAYLOR, P; JONKER, L, 1978. Evolutionary stable strategies and game dynamics. *Mathematical biosciences*. Vol. 40, no. 1-2, p. 145–156 (cit. on p. 51).

TENAILLON, O, 2014. The utility of Fisher's geometric model in evolutionary genetics. *Annual Review of Ecology, Evolution, and Systematics*. Vol. 45, p. 179–201 (cit. on pp. 55, 95, 163, 167, 168, 207, 239, 241).

THOMAS, C D; BODSWORTH, E J; WILSON, R J; SIMMONS, A D, et al., 2001. Ecological and evolutionary processes at expanding range margins. *Nature*. Vol. 411, no. 6837, p. 577–581 (cit. on p. 284).

TRAN, T; HOFRICHTER, J; JOST, J, 2013. An introduction to the mathematical structure of the Wright–Fisher model of population genetics. *Theory in Biosciences*. Vol. 132, no. 2, p. 73–82 (cit. on p. 42).

TRAN, T; HOFRICHTER, J; JOST, J, 2014. The evolution of moment generating functions for the Wright-Fisher model of population genetics. *Mathematical biosciences*. Vol. 256, p. 10–17 (cit. on pp. 60, 332).

TRAVIS, J MJ; DYTHAM, C, 2002. Dispersal evolution during invasions. *Evolutionary Ecology Research*. Vol. 4, no. 8, p. 1119–1129 (cit. on p. 284).

TRAVIS, J MJ; MUSTIN, K; BENTON, T G; DYTHAM, C, 2009. Accelerating invasion rates result from the evolution of density-dependent dispersal. *Journal of theoretical biology*. Vol. 259, no. 1, p. 151–158 (cit. on p. 284).

TRINDADE, S; PERFEITO, L; GORDO, I, 2010. Rate and effects of spontaneous mutations that affect fitness in mutator Escherichia coli. *Philosophical Transactions of the Royal Society B: Biological Sciences*. Vol. 365, no. 1544, p. 1177–1186 (cit. on p. 183).

TRINDADE, S; SOUSA, A; GORDO, I, 2012. Antibiotic resistance and stress in the light of Fisher's model. *Evolution: International Journal of Organic Evolution*. Vol. 66, no. 12, p. 3815–3824 (cit. on pp. 163, 167, 239).

TSIMRING, L S; LEVINE, H; KESSLER, D A, 1996. RNA virus evolution via a fitness-space model. *Physical review letters*. Vol. 76, no. 23, p. 4440–4443 (cit. on pp. 26, 66, 72, 94, 98, 141, 207).

TUFTO, J, 2000. Quantitative genetic models for the balance between migration and stabilizing selection. *Genetics Research*. Vol. 76, no. 3, p. 285–293 (cit. on p. 285).

TURCHIN, P, 1998. *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sinauer, Sunderland, MA (cit. on p. 137).

TURELLI, M, 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical population biology*. Vol. 25, no. 2, p. 138–193 (cit. on p. 89).

TURELLI, M, 2017. Commentary: Fisher's infinitesimal model: A story for the ages. *Theoretical Population Biology*. Vol. 118, p. 46–49. ISSN 00405809. Disponible à l'adresse DOI: 10.1016/j.tpb.2017.09.003 (cit. on p. 285).

TURELLI, M; BARTON, N H, 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? *Genetics*. Vol. 138, no. 3, p. 913–941 (cit. on p. 285).

VON HOLLE, B; SIMBERLOFF, D, 2005. Ecological resistance to biological invasion overwhelmed by propagule pressure. *Ecology*. Vol. 86, no. 12, p. 3212–3218 (cit. on p. 186).

WATT, K E F, 1956. The choice and solution of mathematical models for predicting and maximizing the yield of a fishery. *Journal of the Fisheries Board of Canada*. Vol. 13, no. 5, p. 613–645 (cit. on p. 31).

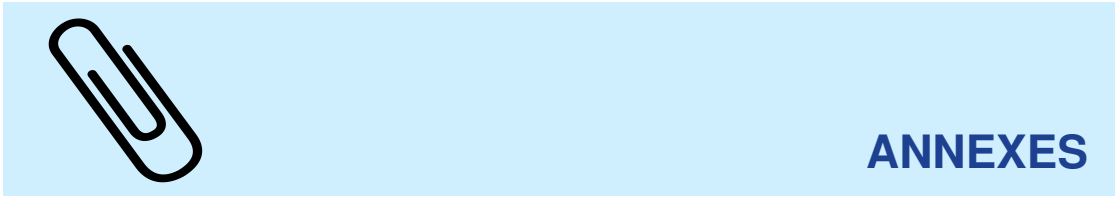
WAXMAN, D; PECK, J R, 1998. Pleiotropy and the preservation of perfection. *Science*. Vol. 279, no. 5354, p. 1210–1213 (cit. on p. 97).

WEINBERGER, H F, 1975. Invariant sets for weakly coupled parabolic and elliptic systems. *Rend. Mat*. Vol. 8, no. 6, p. 295–310 (cit. on pp. 59, 220, 221).

WILLIAMS, J L; HUFBAUER, R A; MILLER, T E X, 2019. How Evolution Modifies the Variability of Range Expansion. *Trends in Ecology & Evolution*. Vol. 34, no. 10, p. 903–913. ISSN 0169-5347. Disponible à l'adresse DOI: https://doi.org/10.1016/j.tree.2019.05.012 (cit. on pp. 285, 306).

WISER, M J; RIBECK, N; LENSKI, R E, 2013. Long-term dynamics of adaptation in asexual populations. *Science*, p. 1243357 (cit. on pp. 24, 26, 96, 100, 111, 112, 142, 356).

YAKYMIV, A L, 2011. A generalization of the Curtiss theorem for moment generating functions. *Math. Notes*. Vol. 90, p. 920–924 (cit. on p. 132).


ANNEXES

# A. Stochastic differential equations satisfied by the phenotypic distribution under the Fisher model

Evolutionary biology tries to understand the demography of populations. Several differential and stochastic models take into account some parameters as the number $n$ of phenotypic traits. The case $n = 1$ was introduced by Fisher, while the general case has just been a direct generalization of the previous case, or studied by a stochastic point of view. In this annex, we use differential results to find the associated infinitesimal generator, after developping the SDE satisfied by the phenotype proportions.

Evolutionary theories are based on different stochastic processes, as the Galton-Watson model. Ewens 2012 and Tran; Hofrichter; Jost 2014 have computed the infinitesimal generator of this process, using stochastic tools (Chapman-Kolmogorov equation, master equation, *etc.*):

$$A\ f(p) = \frac{p(1-p)}{2N}\ \partial^2_{pp} f(p).$$

These techniques have also let to generalize this formula for $(n+1)$-alleles:

$$A\ f(\mathbf{p}) = \sum_{i,j=1}^{n} \frac{(\delta_{i,j} - p_j)\ p_i}{2N}\ \partial_{p_i p_j} f(\mathbf{p}).$$

The generic stochastic methods do not give a stochastic differential equation (SDE) satisfied by the proportions, and often we do not need it: the process is in fact totally characterized by its generator. However, for a better understanding of the evolution of the process, it could be helpful to study the proportions of individuals by a SDE.

## A.1. One equation satisfied by the fitness frequencies

We take a population, in which each individual has a given phenotype $1 \leq i \leq n$, and we assume that the population size is not constant. While the stochasticity is neglected, the number $N_i(t)$ of individuals with phenotype $i$ at time $t$ satisfies an ordinary differential equation (see Section 1.2.1 for more details):

$$\forall t \geq 0,\ N_i'(t) = r_i\ N_i(t),$$

where $r_i$ is the absolute Malthusian fitness of phenotype $i$. However, populations suffer random events, named *genetic drift*. The last equation is transformed into

the following stochastic differential equation (Feller 1951):

$$\mathrm{d}N_i(t) = r_i\ N_i(t)\ \mathrm{d}t + \sqrt{\sigma_i\ N_i(t)}\ \mathrm{d}W_i(t),$$

where $\sigma_i$ is the demographic variance and $(\mathrm{d}W_i(t))_{1\leq i\leq n, t\geq 0}$ is a family of independent Wiener processes. Assuming a demographic isotropy (*i.e.*, $\sigma_i = \sigma$ for all $i$), the independence of the Wiener processes yields a SDE satisfied by the total population size $N(t)$:

$$\mathrm{d}N(t) = \overline{r}(t)\ N(t)\ \mathrm{d}t + \sum_{i=1}^{n} \sqrt{\sigma_i\ N_i(t)}\ \mathrm{d}W_i(t) = \overline{r}(t)\ N(t)\ \mathrm{d}t + \sqrt{\sigma\ N(t)}\ \mathrm{d}W(t),$$

where $r(t) = \sum_{i=1}^{n} r_i N_i(t)/N(t)$ is the mean absolute Malthusian fitness of the population and $(\mathrm{d}W(t))_{t\geq 0}$ a Wiener process. However, it is more convenient to study the frequencies of phenotype $p_i(t) = N_i(t)/N(t)$. We can check that:

$$\mathrm{d}p_i(t) = (r_i - \overline{r}(t))\ p_i(t)\ \mathrm{d}t + \sqrt{\frac{\sigma\ p_i(t)}{N(t)}}\ \mathrm{d}W_i(t) - \sum_{j=1}^{n} \sqrt{\frac{\sigma p_i^2(t) p_j(t)}{N(t)}}\ \mathrm{d}W_j(t).$$

## A.2. Associated infinitesimal generator

To understand an Itô diffusion process, we have to compute its infinitesimal generator. For this, we recall a famous result:

**Theorem 45.**
*Let a stochastic process $X_t(\omega) : \mathbb{R}^+ \times \Omega \to \mathbb{R}^n$ satisfying the stochastic diffusion equation $\mathrm{d}X_t = b(X_t)\ \mathrm{d}t + \Sigma(X_t)\ \mathrm{d}\mathbf{W}(t)$, with $(\mathrm{d}\mathbf{W}(t))_{t\geq 0}$ is a $m$-dimensional Wiener process.*
*Then the infinitesimal generator $A$ of $X_t$ is defined on $D_A \supseteq C_0^2(\mathbb{R}^n)$, and:*

$$\forall f \in C_0^2(\mathbb{R}^n),\ \forall x \in \mathbb{R}^n,\ A\ f(x) = \sum_{i=1}^{n} b_i(x)\ \partial_{x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^{n} (\Sigma\Sigma^T)_{i,j}(x)\ \partial^2_{x_i x_j} f(x).$$

This result yields the following infinitesimal generator for $\mathbf{p}(t) = (p_i(t))_{1\leq i\leq n}$.

**Proposition 46.** *The generator associated to the proportions vector of phenotype satisfies:*

$$\forall f \in C_0^2(\mathbb{R}^n),\ \forall \mathbf{p} \in \mathbb{S}^{n-1},\ A\ f(\mathbf{p}) = \sum_{i=1}^{n}(r_i - \overline{r}(\mathbf{p}))\ p_i\ \partial_{p_i} f(\mathbf{p}) + \frac{1}{2N(t)} \sum_{i,j=1}^{n} (\delta_{i,j} - p_j)\ p_i\ \partial_{p_i p_j} f(\mathbf{p}).$$

*where* $\overline{r}(\mathbf{p}) = \sum_{i=1}^{n} r_i p_i$.

*Proof.* Thanks to Theorem 45, we have just to compute $(\Sigma\Sigma^T)_{i,j}$, with:

$$\Sigma_{i,j}(\mathbf{p}) = \delta_{i=j}\sqrt{\frac{\sigma p_i}{N(t)}} - \sqrt{\frac{\sigma p_i^2 p_j}{N(t)}}.$$

Let two indices $i \neq j$. Then we have for all $\mathbf{p} \in \mathbb{S}^{n-1}$:

$$\begin{aligned}
(\Sigma\Sigma^T)_{i,j}(\mathbf{p}) &= \sum_{k=1}^{n} \left[\delta_{i=k}\sqrt{\frac{\sigma p_i}{N(t)}} - \sqrt{\frac{\sigma p_i^2 p_k}{N(t)}}\right] \cdot \left[\delta_{j=k}\sqrt{\frac{\sigma p_j}{N(t)}} - \sqrt{\frac{\sigma p_j^2 p_k}{N(t)}}\right], \\
&= -\sqrt{\frac{\sigma p_i}{N(t)}} \cdot \sqrt{\frac{\sigma p_j^2 p_i}{N(t)}} - \sqrt{\frac{\sigma p_i^2 p_j}{N(t)}} \cdot \sqrt{\frac{\sigma p_j}{N(t)}} + \sum_{k=1}^{n} \sqrt{\frac{\sigma p_i^2 p_k}{N(t)}} \cdot \sqrt{\frac{\sigma p_j^2 p_k}{N(t)}}, \\
&= -2\ \frac{\sigma p_j p_i}{N(t)} + \sum_{k=1}^{n} p_k \cdot \frac{\sigma p_j p_i}{N(t)} = -\ \frac{\sigma p_j p_i}{N(t)},
\end{aligned}$$

as $\mathbf{p} \in \mathbb{S}^{n-1}$. By the same kind of arguments, we can check that:

$$(\Sigma\Sigma^T)_{ii}(\mathbf{p}) = \frac{\sigma p_i(1 - p_i)}{N(t)},$$

which ends the proof. □

# B. Extinction times of an inhomogenous Feller diffusion process: a PDE approach

**Florian LAVIGNE[a,b,c] and Lionel ROQUES[a]**

[a] BioSP, INRA, 84914, Avignon, France
[b] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
[c] ISEM (UMR 5554), CNRS, 34095, Montpellier, France

**Abstract**

We focus on the distribution of the extinction times of a population whose size $N(t)$ follows a Feller diffusion process with inhomogeneous growth term $r(t)$. Obtaining a precise description of the extinction times and of their dependence with respect to $r(t)$ has important applications in adaptive biology, for understanding "evolutionary rescue" phenomena. A formula for the distribution of the extinction times has been recently obtained, through probabilistic arguments. The aim of this note is to propose a new derivation of this formula, based on the analysis of a degenerate parabolic partial differential equation.

Keywords: *Partial differential equations; Birth-death process; Extinction times; Population dynamics; Evolutionary dynamics*

## B.1. Introduction

The Itô stochastic differential equation (SDE),

$$\mathrm{d}N(t) = r\,N(t)\mathrm{d}t + \sqrt{\sigma^2 N(t)}\mathrm{d}W(t),\ N(0) = N_0 \geq 0, \tag{.30}$$

emerges as the diffusion limit of a birth-death process (Baake; Wakolbinger 2015; Feller 1951; Lambert 2008). This equation, also known as Feller's diffusion, describes the evolution of the size $N(t)$ of a population with birth rate $b = k + \frac{b_1}{N}$ and death rate $d = k + \frac{d_1}{N}$, $b_1, d_1 \geq 0$ and $k > 0$. The term $r = b_1 - d_1$ is the growth rate, $\sigma^2 = 2\,k$ is the stochastic reproductive variance and $W$ a standard Brownian motion.

When $N(t)$ reaches $0$, we keep it equal to $0$ (the process so defined still satisfies (.30)), and the population becomes extinct. Here, we are interested in the distribution of the extinction times of a population, whose size satisfies the time-inhomogeneous SDE:

$$\mathrm{d}N(t) = r(t)\,N(t)\mathrm{d}t + \sqrt{\sigma^2 N(t)}\mathrm{d}W(t),\ N(0) = N_0 \geq 0, \tag{.31}$$

with $r(t)$ a continuous function. This equation naturally arises (Anciaux; Chevin; Ronce; Martin 2018) in the context of genetic adaptation. Due to adaptation, the

growth term $r(t)$ changes over time. Thus, a population with a negative initial growth term $r(0) < 0$ may ultimately survive if $r(t)$ becomes positive at larger times. This phenomenon is known as “evolutionary rescue" (Anciaux; Chevin; Ronce; Martin 2018; Gomulkiewicz; Holt 1995). To characterize the extinction times, we define the first exit time from $(0, +\infty)$:

$$\tau = \inf\{t \geq 0,\ \exists s \leq t, N(s) = 0\}.$$

By definition of $\tau$ and of the stochastic process $N(t)$, we have $N(t) = 0$ for all $t \geq \tau$.

When $r$ is constant, as in (.30), the distribution of the extinction times $\tau$ is well-known (*e.g.*, proposition 4.4 in section 4.3 of Dawson 2017). The cumulative distribution function of $\tau$ is:

$$\forall\, t > 0,\ \forall\, N_0 \geq 0,\ \mathbb{P}(\tau \geq t|N_0) = 1 - \exp\left[-\frac{2rN_0\sigma^{-2}}{1 - e^{-rt}}\right].$$

In the time-inhomogeneous case of (.31), the computation of the distribution of $\tau$ is more involved. Recent works (Bansaye; Simatos 2015) based on general probabilistic results on the convergence of Galton-Watson branching processes in varying environments show that:

$$\forall\, t > 0,\ \forall\, N_0 \geq 0,\ \mathbb{P}(\tau \geq t|N_0) = 1 - \exp\left[-\frac{2\, N_0\sigma^{-2}}{\int_0^t \exp\left(-\int_0^\xi r(\gamma)d\gamma\right) d\xi}\right]. \quad (.32)$$

The objective of this note is to propose a new method, based on partial differential equation techniques, to compute the cumulative distribution function of $\tau$, $\mathbb{P}(\tau \geq t|N_0)$.

## B.2. Main results

Standard theory shows that the function $u(s, x; t) = \mathbb{P}(\tau \geq t|N(t - s) = x)$ ($0 \leq s \leq t$, $x \geq 0$) emerges as a representation of the solution of a degenerate parabolic partial differential equation (see theorem 2.3., page 133 in Freidlin 1985):

$$\begin{cases} \partial_s u(s, x; t) &= \frac{\sigma^2 x}{2}\, \partial^2_{xx} u(s, x; t) + r(t - s)x\, \partial_x u(s, x; t),\ 0 < s < t,\ x > 0, \\ u(0, x; t) &= 1,\ x > 0. \end{cases} \quad (.33)$$

More precisely, $t > 0$ being fixed, if the problem (.33) admits a solution $u(s, x; t) \in C^{1,2}((0, t) \times (0, +\infty))$ ($C^1$ in time and $C^2$ in space), such that:

$$\forall\, h \in (0, t),\ \exists\, \mu > 0 \text{ such that } \|u\|_{C^{1,2}((h,t)\times(0,+\infty))} < \mu, \quad (.34)$$

with $\|\cdot\|_{C^{1,2}((h,t)\times(0,+\infty))}$ the usual norm on $C^{1,2}((h,t)\times(0,+\infty))$, then $u(s,x;t) = \mathbb{P}(\tau \geq t | N(t-s) = x)$ for all $0 \leq s \leq t$ and $x > 0$. Thus, to obtain an explicit expression for $\mathbb{P}(\tau \geq t | N(t-s) = x)$, we need to find an explicit solution of (.33) with the above regularity and boundedness conditions.

Due to the degenerate diffusion term $\frac{\sigma^2 x}{2}\,\partial^2_{xx} u(s,x)$, standard parabolic theory does not apply. To the best of our knowledge, there is no standard result that shows the existence of a classical solution of (.33) (but see Epstein; Mazzeo 2013 and Favini; Marinoschi 2012 for some results on other degenerate parabolic equations with constant advection term). In this note, we derive some necessary conditions on $u$ that lead to an explicit expression, and we show that this expression satisfies (.33) and (.34).

We begin with the derivation of an explicit expression for the Laplace transform of $u(s,\cdot;t)$. For a bounded measurable function $g$, we recall that the Laplace transform (according to $x$) is given by:

$$\mathcal{L}g(z) = \int_0^{+\infty} e^{-zx} g(x) dx \quad \text{for all } z > 0.$$

We denote by $M(s,z;t) = \mathcal{L}u(s,z;t)$ the Laplace transform of $u(s,\cdot;t)$.

**Proposition 47.** *If $u$ satisfies (.33)-(.34), then its Laplace transform $M(s,z;t)$ satisfies for $0 < s < t$ and $z > 0$:*

$$\partial_s M(s,z;t) = -\left[r(t-s)z + \frac{\sigma^2}{2}z^2\right]\partial_z M(s,z;t) - [\sigma^2 z + r(t-s)]M(s,z;t). \tag{.35}$$

*Moreover, for all times $0 \leq s \leq t$ and $z \geq 0$, we have:*

$$M(s,z;t) = z^{-1} \Big/ \left[1 + \frac{\sigma^2 z}{2}\int_{t-s}^{t} \exp\left(-\int_{t-s}^{\xi} r(\gamma)d\gamma\right) d\xi\right].$$

In order to compute the inverse Laplace transform of $M(s,\cdot;t)$, we first note that $M(s,z;t)$ satisfies, for all $0 \leq s \leq t$ and $z \geq 0$:

$$M(s,z;t) = \frac{1}{z} - \frac{1}{\frac{1}{\alpha_t(s)} + z},$$

with $\alpha_t(s) = \frac{\sigma^2}{2}\int_{t-s}^{t} \exp\left(-\int_{t-s}^{\xi} r(\gamma)d\gamma\right) d\xi$. Moreover, we know that the Laplace transform of an exponential function $h : x \mapsto \exp(-\beta x)$, for $\beta \geq 0$, is:

$$\mathcal{L}h(z) = (\beta + z)^{-1}.$$

We deduce that $M(s,z;t)$ is the Laplace transform of the function:

$$u(s,x;t) = 1 - \exp\left[ - \frac{2x\sigma^{-2}}{\int_{t-s}^{t} \exp\left( - \int_{t-s}^{\xi} r(\gamma) d\gamma \right) d\xi} \right].$$

We readily observe that this function is a classical solution of (.33)-(.34). Thanks to theorem 2.3. page 133 in Freidlin 1985, we finally get that $\mathbb{P}(\tau \geq t | N(t-s) = x) = u(s,x;t)$. Consequently,

**Corollary 10.**
*For all times $t > 0$, $0 \leq s \leq t$ and for all initial population size $N_0 \geq 0$, we have:*

$$\mathbb{P}(\tau \geq t | N(0) = N_0) = u(t, N_0; t) = 1 - \exp\left[ - \frac{2\, N_0\, \sigma^{-2}}{\int_{0}^{t} \exp\left( - \int_{0}^{\xi} r(\gamma) d\gamma \right) d\xi} \right].$$

This last result is consistent with the formula (.32).

## B.3. Proof of Proposition 47

From the assumption (.34) on $u$, we know that, for each $s > 0$ and $t > 0$, there exists $\mu > 0$ such that $|e^{-zx} u(s,x;t)| \leq \mu\ e^{-zx}$ for all $z,\ x > 0$. Thus $M(s,z;t) = \int_0^\infty e^{-zx} u(s,x;t)\, dx$ is well-defined. By the same kind of arguments, we can check that $M$ is $C^1$ with respect to $t$ and $C^2$ with respect to $z$. Thus, multiplying (.33) by $e^{-z\,x}$, and integrating between $0$ and $+\infty$, we have for all $t > 0$ and $z > 0$:

$$\begin{aligned} \partial_s M(s,z;t) &= \int_0^{+\infty} e^{-zx} \partial_t u(s,x;t)\, dx, \\ &= \int_0^{+\infty} \tfrac{\sigma^2 x}{2} e^{-zx} \partial_{xx} u(s,x;t) dx + r(t-s) \int_0^{+\infty} x e^{-zx} \partial_x u(s,x;t) dx. \end{aligned}$$

Integrating by parts, we get:

$$\begin{aligned} \partial_s M(s,z;t) = \frac{\sigma^2}{2} &\left[ \left[ x e^{-zx} \partial_x u(s,x) \right]_0^{+\infty} - \int_0^{+\infty} (1 - zx) e^{-zx} \partial_x u(s,x;t) dx \right] \\ &+ r(t-s) \left[ \left[ x e^{-zx} u(s,x;t) \right]_0^{+\infty} - \int_0^{+\infty} (1 - zx) e^{-zx} u(s,x;t) dx \right]. \end{aligned}$$

Thanks to (.34), this expression reduces to:

$$\partial_s M(s,z;t) = -\frac{\sigma^2}{2} \int_0^{+\infty} (1 - zx) e^{-zx} \partial_x u(s,x;t) dx - r(t-s)[M(s,z;t) + z \partial_z M(s,z;t)].$$

Integrating again by parts, we get:

$$\partial_s M(s,z;t) = -\frac{\sigma^2}{2}\left[[(1-zx)e^{-zx}u(s,x;t)]_0^{+\infty} - \int_0^{+\infty}(z^2x-2z)e^{-zx}u(s,x;t)dx\right] \\ - r(t-s)[M(s,z;t)+z\partial_z M(s,z;t)],$$

and so the hypothesis (.34) yields that:

$$\partial_s M(s,z;t) = -\left[\frac{\sigma^2}{2}z^2 + r(t-s)z\right]\partial_z M(s,z;t) - [\sigma^2 z + r(t-s)]\, M(s,z;t).$$

We have shown that $M(s,z)$ satisfies (.35).

Now to solve this equation, we use the method of characteristics. Fix a time $t>0$ and a parameter $z_0>0$. Let $\chi(s)$ and $v(s)$ be defined as the solution of:

$$\begin{cases} \chi'(s) &= r(t-s)\chi(s) + \frac{\sigma^2}{2}\chi^2(s),\ 0<s<t, \\ \chi(0) &= z_0, \end{cases} \tag{.36}$$

and:

$$v(s) = M(s,\chi(s);t).$$

First, the function $w(s) = \exp\left(-\int_0^s r(t-\xi)d\xi\right)\chi(s)$ satisfies:

$$w'(s) = \frac{\sigma^2}{2}\chi^2(s)\exp\left(-\int_0^s r(t-\xi)d\xi\right) = \frac{\sigma^2}{2}\exp\left(\int_0^s r(t-\xi)d\xi\right)w^2(s).$$

This implies that $w(s) = 1/\left[c - \frac{\sigma^2}{2}\int_0^s \exp\left(\int_0^\xi r(t-\gamma)d\gamma\right)d\xi\right]$ with $c = w(0)^{-1} = \chi(0)^{-1} = z_0^{-1}$. Thus we get:

$$\chi(s) = \exp\left(\int_0^s r(t-\xi)d\xi\right) / \left[c - \frac{\sigma^2}{2}\int_0^s \exp\left(\int_0^\xi r(t-\gamma)d\gamma\right)d\xi\right]. \tag{.37}$$

Second, as $v'(s) = -[\sigma^2\chi(s) + r(t-s)]v(s)$, we have:

$$\begin{aligned} v(s) &= v(0)\exp\left(-\int_0^s r(t-\xi)d\xi - \sigma^2\int_0^s \chi(\xi)d\xi\right), \\ &= v(0)\exp\left[-\int_0^s r(t-\xi)d\xi + 2\ln\left(1 - \frac{\sigma^2}{2c}\int_0^s \exp\left(\int_0^\xi r(t-\gamma)d\gamma\right)d\xi\right)\right]. \end{aligned}$$

Thus, we have:

$$v(s) = v(0)\left[1 - \frac{\sigma^2 z_0}{2}\int_0^s \exp\left(\int_0^\xi r(t-\gamma)d\gamma\right)d\xi\right]^2 \exp\left(-\int_0^s r(t-\xi)d\xi\right). \tag{.38}$$

Let us choose $z>0$ and $s>0$. By formula (.37), there exists $z_0>0$ such that

$\chi(s)$ exists and $\chi(z) = z$ (so that $v(s) = M(s, \chi(s)) = M(s, z)$):

$$z_0 = z / \left[ \exp\left( \int_0^s r(t-\xi) d\xi \right) + \frac{\sigma^2 z}{2} \int_0^s \exp\left( \int_0^\xi r(t-\gamma) d\gamma \right) d\xi \right]. \tag{.39}$$

As $v(0) = M(0, z_0; t)$ and $u(0, \cdot; t) = 1$, we have $v(0) = 1/z_0$. Finally, using $M(s, z; t) = M(s, \chi(s); t) = v(s)$ and (.38)-(.39), we obtain:

$$\begin{aligned} M(s,z;t) &= \exp\left( \int_0^s r(t-\xi) d\xi \right) z^{-1} / \left[ \exp\left( \int_0^s r(t-\xi) d\xi \right) + \frac{\sigma^2 z}{2} \int_0^s \exp\left( \int_0^\xi r(t-\gamma) d\gamma \right) d\xi \right], \\ &= z^{-1} / \left[ 1 + \frac{\sigma^2 z}{2} \int_0^s \exp\left( - \int_{t-s}^{t-\xi} r(\gamma) d\gamma \right) d\xi \right], \\ &= z^{-1} / \left[ 1 + \frac{\sigma^2 z}{2} \int_{t-s}^t \exp\left( - \int_{t-s}^{\xi} r(\gamma) d\gamma \right) d\xi \right], \end{aligned}$$

which ends the proof of Proposition 47.

# C. Cas de noyau mutationnel isotrope

Nous nous intéressons ici à l'évolution d'une population asexuée, qui se développe dans un unique environnement, à partir d'un moment $t = 0$, soit parce qu'elle y était déjà présente, soit parce qu'elle y a été introduite. Chaque individu souffre de la sélection, des mutations et de la dérive génétique, comme cela fut présenté dans le Chapitre 1. Cette population aura une taille constante au cours du temps : l'effet de la démographie sera ainsi négligé. De plus, cet effectif sera supposé suffisament grand, pour pouvoir négliger les effets de la dérive génétique.

Nous ferons une hypothèse supplémentaire sur les mutations. Le noyau $K$ sera supposé être *isotrope* dans le sens où,

**(H-iso)** La fonction $K$ est radiale :
$\forall\, x \in \mathbb{R}^n,\ K(\mathbf{x}) = K(\|\mathbf{x}\|)$.

Ainsi, dans la Section C.1, nous étudierons le modèle EDP suivant :

$$\partial_t q(t,\mathbf{x}) = [m(\mathbf{x}) - \overline{m}(t)]q(t,\mathbf{x}) + U\left[\int_{\mathbb{R}^n} K(\|\mathbf{x}-\mathbf{y}\|)\, q(t,\mathbf{y})\,\mathrm{d}\mathbf{y} - q(t,\mathbf{x})\right], \quad (.40)$$

pour $\overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})\, q(t,\mathbf{x})\,\mathrm{d}\mathbf{x}$. Nous dériverons de cette EDP, d'autres modèles basés sur la distribution $p(t,m)$ des fitness (sous-entendu Malthusienne relative), ce qui nous permettra d'obtenir des résultats analytiques sur $\overline{m}(t)$.

La Section C.2 sera dédiée à l'étude de l'approximation diffusive présentée dans la Section 1.3.2, en appliquant les mêmes méthodes que dans la Section C.1, pour obtenir des résultats analytiques, que nous comparerons avec des données expérimentales dans la Section C.3.

# Sommaire

## C.1. Modèle intégral

Nous faisons l'hypothèse que l'évolution de la population asexuée considérée ne subit pas la démographie (son effectif $N$ ne dépend pas du temps $t \geq 0$) et que le nombre d'individus est suffisament important pour négliger la dérive génétique. Dans ce cas, nous avons déjà vu que l'évolution de la distribution des phénotypes $\mathbf{x} \in \mathbb{R}^n$ pouvait être modélisée par l'EDP (.40). Cependant, il semble impossible d'avoir directement la trajectoire évolutive de la population : la fonction $\overline{m}(t)$ ne peut être explicitée en tout temps. Il faudra utiliser l'appoximation diffusive (décrite en Section 1.3.2) et faire intervenir une nouvelle fonction : la distribution de fitness $p(t, m)$ (voir Section C.1.2).

### C.1.1. Propriétés sur la distribution des phénotypes

Penchons-nous sur le problème de Cauchy suivant :

$$\begin{cases} \partial_t q(t,\mathbf{x}) = U\left[\displaystyle\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\, q(t,\mathbf{y})\,\mathrm{d}\mathbf{y} - q(t,\mathbf{x})\right] + [m(\mathbf{x}) - \overline{m}(t)]q(t,\mathbf{x}), & \forall t>0,\ \forall \mathbf{x}\in\mathbb{R}^n, \\ q(0,\mathbf{x}) = q_0(\mathbf{x}), & \forall \mathbf{x}\in\mathbb{R}^n. \end{cases} \tag{.41}$$

Nous admettrons que, pour des distributions de probabilité initiales $q_0$ adéquates, il existe une unique solution $q \in C^1(\mathbb{R}_+, L^\infty(\mathbb{R}^n))$ positive avec $\overline{m}(t) \in C(\mathbb{R}_+)$ et :

$$\forall t > 0,\ \int_{\mathbb{R}^n} q(t,\mathbf{x})\mathrm{d}\mathbf{x} = 1.$$

Ce résultat semble assez naturel, vue le théorème 2.2 de l'article de Gil; Hamel; Martin; Roques 2017, qui traite de ce problème dans le cas $n = 1$ et $m$ linéaire. Le même type de preuve semble fonctionner. Cependant, il manque à cette extension la donnée de sur-et-sous-solutions explicites.

Bien que ce résultat ne semble pas simple à démontrer, nous pourrions prouver grâce à la distribution de fitness (cf. Section C.1.2) et à Gil; Hamel; Martin; Roques 2019, que si la condition initiale est radiale, alors le problème de Cauchy admet une unique solution radiale positive. En effet, si $q(t,\cdot)$ est radiale en tout temps $t \geq 0$, il existe alors une bijection entre les distributions $q$ et $p$. Or la distribution $p$ vérifie une équation du type de celle étudiée dans Gil; Hamel; Martin; Roques 2019, ce qui donnerait l'existence (et l'unicité) d'une telle solution $p$ et donc $q$.

Revenons sur les résultats connus sur cette équation différentielle. Bien qu'elle ne fut que partiellement étudiée (Kimura 1954, étant le premier à poser ce modèle, décide de l'approximer par un terme diffusif, comme cela est décrit en Section 1.3.2), il faut attendre 1986, pour obtenir les premiers travaux sur cette équation, portant sur les états stationnaires, dans le cas $n = 1$ :

**Théorème 10. (BÜRGER 1986)**
*Il existe une unique distribution stationnaire $\widetilde{q}$ à l'équation (.41). Elle est strictement positive, symmétrique, et satisfait :*

$$\int_{\mathbb{R}} x^2 \widetilde{q}(x)\,\mathrm{d}x < +\infty.$$

*De plus, quelque soit la condition initiale positive $q_0 \in L^2(\mathbb{R})$ telle que $\sqrt{1+x^2}q_0 \in L^2(\mathbb{R})$, la solution $q(t,\cdot)$ (qui existe) converge vers $\widetilde{q}$, quand $t \to +\infty$.*

Remarquons qu'estimer $\int_{\mathbb{R}} x^2\widetilde{q}(x)\,\mathrm{d}x$ revient à étudier la première valeur propre d'un opérateur différentiel. En effet, nous avons que pour tout $x \in \mathbb{R}$ :

$$\frac{x^2}{2}\widetilde{q}(x) - U\int_{\mathbb{R}} K(x-y)\,\widetilde{q}(y)\,\mathrm{d}y = [\lambda - U]\widetilde{q}(x), \quad \text{avec} \quad \lambda = \int_{\mathbb{R}} \frac{x^2}{2}\widetilde{q}(x)\,\mathrm{d}x.$$

Ainsi la preuve proposée par Bürger 1986 implique que :

$$0 < \int_{\mathbb{R}} \frac{x^2}{2}\widetilde{q}(x)\,\mathrm{d}x < U.$$

Aucune formule exacte ne semble être simple à trouver, comme le fait remarquer l'auteur. Cependant, il est possible de l'approximer par un algorithme numérique (voir Bürger 1986 pour plus de détails).

Finalement, le problème (.41) ne permet nullement de tirer d'informations directes sur l'évolution et l'adaptation de la population.

### C.1.2. Distribution de fitness

Bien que nous ayons des estimations sur la valeur finale de $\overline{m}$ (dans le cas $n = 1$), nous n'arrivons pas à avoir le comportement de la fitness moyenne $\overline{m}(t)$ en tout temps $t \geq 0$, ni sa limite exacte. L'astuce que nous utiliserons ici va être de définir la distribution de fitness $p(t,m)$, qui est la mesure image de $q$ par la fonction $\mathbf{x} \mapsto -\|\mathbf{x}\|^2/2$. La représentation en strates de $q$ apporte la proposition suivante :

**Proposition 48.** *En tout temps $t \geq 0$, et pour toute fitness $m < 0$, la distribution de fitness $p$ est donnée par :*

$$p(t,m) = (2|m|)^{n/2-1} \int_{\mathbb{S}^{n-1}} q\Big(t, \sqrt{2|m|}\,\sigma\Big)\,\mathrm{d}\sigma = (2|m|)^{n/2-1}\, Q\Big(t, \sqrt{2|m|}\Big),$$

*avec $\mathbb{S}^{n-1}$ la sphère unité euclidienne de $\mathbb{R}^n$ et la fonction $Q$ définie par :*

$$\forall\, t \geq 0,\ \forall\, r \geq 0,\ Q(t,r) = \int_{\mathbb{S}^{n-1}} q(t, r\,\sigma)\,\mathrm{d}\sigma.$$

*De plus, nous avons :*

$$\forall\, t \geq 0,\ \int_{-\infty}^{0} p(t,m)\,\mathrm{d}m = 1,\ \textit{et}\ \overline{m}(t) = \int_{-\infty}^{0} m\,p(t,m)\,\mathrm{d}m,$$

*où les intégrales ci-dessus convergent.*

Comme les fonctions $K$ et $q$ sont positives, le théorème de Fubini implique que pour tout $t \geq 0$ et $r \geq 0$ :

$$\begin{aligned}
\int_{\mathbb{S}^{n-1}} \int_{\mathbb{R}^n} K(\|r\sigma - \mathbf{y}\|)\, q(t,\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\sigma &= \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \int_0^{+\infty} K(\|r\sigma - r'\sigma'\|)\, q(t,r'\sigma')\,\mathrm{d}r'\,\mathrm{d}\sigma'\,\mathrm{d}\sigma, \\
&= \int_{\mathbb{S}^{n-1}} \int_0^{+\infty} \int_{\mathbb{S}^{n-1}} K(\|r\sigma - r'\sigma'\|)\,\mathrm{d}\sigma\, q(t,r'\sigma')\,\mathrm{d}r'\,\mathrm{d}\sigma', \\
&= \int_{\mathbb{S}^{n-1}} \int_0^{+\infty} \int_{\mathbb{S}^{n-1}} K(\|r\sigma - r'\|)\,\mathrm{d}\sigma\, q(t,r'\sigma')\,\mathrm{d}r'\,\mathrm{d}\sigma', \\
&= \int_0^{+\infty} \int_{\mathbb{S}^{n-1}} K(\|r\sigma - r'\|)\,\mathrm{d}\sigma \int_{\mathbb{S}^{n-1}} q(t,r'\sigma')\,\mathrm{d}\sigma'\,\mathrm{d}r', \\
&= \int_0^{+\infty} \int_{\mathbb{S}^{n-1}} K(\|r\sigma - r'\|)\,\mathrm{d}\sigma\, Q(t,r')\,\mathrm{d}r'.
\end{aligned}$$

Ainsi la fonction $Q$ est solution de :

$$\forall\, t > 0,\ \forall r \geq 0,\ \partial_t Q(t,r) = U\left[\int_0^{\infty} K'(r,r')\,Q(t,r')\,\mathrm{d}r' - Q(t,r)\right] + \left[-\frac{r^2}{2} - \overline{m}(t)\right] Q(t,r),$$

avec $K'(r,r') = \int_{\mathbb{S}^{n-1}} K(\|r\sigma - r'\|)\,\mathrm{d}\sigma$. Finalement, ces arguments mis bout-à-bout apportent une EDP dont $p$ est solution :

**Théorème 11.**
*La distribution de fitness $p(t,m)$ est une solution classique de classe $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}_-^*)$ de l'EDP :*

$$\forall\, t>0,\, \forall m<0,\, \partial_t p(t,m) = U\left[\int_{-\infty}^{0} K'\left(\sqrt{2|m|},\sqrt{2|m'|}\right)\, p(t,m')\,\mathrm{d}m' - p(t,m)\right] + (m-\overline{m}(t))\,p(t,m), \quad (.42)$$

*avec le noyau mutationnel étant défini par :*

$$\forall\, r\geq 0,\, \forall\, r'\geq 0,\, K'(r,r') = r^{n-2}\int_{\mathbb{S}^{n-1}} K(\|r\sigma - r'\|)\,\mathrm{d}\sigma, \quad (.43)$$

*et la condition initiale :*

$$p_0(m) = (2|m|)^{n/2-1}\int_{\mathbb{S}^{n-1}} q_0\left(\sqrt{2|m|}\,\sigma\right)\mathrm{d}\sigma. \quad (.44)$$

Dans le cadre du FGM – $K$ une gaussienne centrée isotrope de variance $\lambda>0$ – nous pouvons expliciter le noyau $K'$ :

$$\forall\, m\leq 0,\, \forall\, m'\leq 0,\, K'\left(\sqrt{2|m|},\sqrt{2|m'|}\right) = \frac{|m|^{(n-2)/2}}{\lambda^{n/2}\Gamma(n/2)}\exp\left(\frac{m+m'}{\lambda}\right)\ {}_0F_1\left(;\frac{n}{2};\frac{mm'}{\lambda^2}\right).$$

Les calculs sont détaillés dans la Section C.1.4. On retrouve les résultats présents dans l'article de Martin 2014.

Bien que le terme quadratique $m(\mathbf{x})$ (voir (.41)) ait été remplacé ici par un terme linéaire $m$, cette équation n'admet pas d'étude simple de la solution. Cependant nous allons pouvoir l'utiliser pour étudier une transformée de $p$ : ses fonctions génératrices $M_p(t,z)$ et $C_p(t,z)$.

### C.1.3. Etudes des fonctions génératrices

Nous étendons les définitions de $M_p$ et $C_p$ données dans la Section 1.5.2 dans le cas d'un ensemble continu de phénotypes :

$$\forall\, z\geq 0,\, M_p(t,z) = \int_{-\infty}^{0} e^{zm}p(t,m)\,\mathrm{d}m, \quad \text{et} \quad C_p(t,z) = \log M_p(t,z).$$

La fonction $M_p$ vérifie en règle générale une équation différentielle ouverte :

$$\partial_t M_p(t,z) = U\left[\int_{-\infty}^{0} p(t,m')\int_{-\infty}^{0} e^{zm}K'\left(\sqrt{2|m|},\sqrt{2|m'|}\right)\mathrm{d}m\,\mathrm{d}m' - M_p(t,z)\right] + \partial_z M_p(t,z) - \overline{m}(t)M_p(t,z),$$

avec $\overline{m}(t) = \partial_z M_p(t,0)$. Le choix de $K$ (et donc de $K'$) devient alors crucial pour la suite des calculs. C'est pourquoi nous ferons maintenant le choix de considérer le modèle FGM, pour lequel l'équation devient fermée du même type que celui donné dans Gil; Hamel; Martin; Roques 2019 (voir Section C.1.4 pour une dérivation de l'EDP ou Martin; Roques 2016 qui propose un raisonnement direct sur la variation de la MGF).

**Proposition 49.** *La fonction $C_p \in C^1(\mathbb{R}_+ \times \mathbb{R}_+)$ est une solution classique du système :*

$$\begin{cases} \partial_t C_p(t,z) = \partial_z C_p(t,z) - \overline{m}(t) + U\left(e^{C_p(t,z+\omega(z))-C_p(t,z)} M_*(z) - 1\right), & \forall\, t>0,\, \forall\, z \geq 0, \\ C_p(0,z) = C_{p_0}(z), & \forall\, z \geq 0, \\ C_p(t,0) = 0, & \forall\, t \geq 0, \end{cases} \tag{.45}$$

*avec la fitness moyenne $\overline{m}(t) = \partial_z C_p(t,0)$, la condition initiale :*

$$C_{p_0}(z) = \log\left[\int_{-\infty}^{0} e^{zm} p_0(m)\,\mathrm{d}m\right],$$

*et les fonctions $M_*$ et $\omega$ définies par :*

$$\forall\, z>0,\ M_*(z) = (1+\lambda z)^{-n/2}, \qquad \text{et} \qquad \omega(z) = -\frac{\lambda z^2}{1+\lambda z}.$$

Là-encore, cette équation ne nous permet pas d'avoir des résultats sur l'évolution de la fitness moyenne $\overline{m}$. Cependant, en faisant l'hypothèse que la famille $(p(t,\cdot))_{t\geq 0}$ converge faiblement vers une mesure de Radon $p_\infty$, il est possible d'avoir des résultats sur l'état d'équilibre $p_\infty$, en faisant peu d'hypothèse sur la dimension (contrairement au Theorème 10).

**Théorème 12. (Gil; Hamel; Martin; Roques 2019)**
*Supposons que pour tout $\alpha > 0$, $\lim_{m\to-\infty} e^{\alpha|m|} p_0(m) = 0$. Alors la fitness moyenne converge vers $\overline{m}_\infty := \int_{-\infty}^0 m\, p_\infty(m)\,\mathrm{d}m \geq -U$ et $p_\infty$ peut s'écrire sous la forme :*

$$p_\infty = \varrho\delta_0 + (1-\varrho)p^*,$$

*pour $\varrho \in [0,1]$ et $p^*$ une distribution de probabilité supportée par $\mathbb{R}_-$, et n'ayant pas de masse en 0. De plus :*

- $\varrho = 0$ *et* $\overline{m}_\infty > -U$, *si* $n \leq 2$ ;
- $\varrho > 0$ *et* $\overline{m}_\infty = -U$, *si* $n > 2$ *et* $U \leq \lambda(n-2)/2$ ;
- $\varrho = 0$ *et* $\overline{m}_\infty > -U$, *si* $n > 2$ *et* $U > \lambda(n-2)/2$.

Nous avons donc pu obtenir quelques informations supplémentaires sur le comportement asymptotique de la distribution des fitness $p$ et la fitness moyenne $\overline{m}$. Cependant, il est possible d'avoir des résultats analytiques sur l'évolution de ces quantités, en utilisant une approximation – l'approximation diffusive – ce que nous allons voir dans la section suivante.

### C.1.4. Distributions des effets mutationnels sur la fitness

Dans cette partie, nous détaillons les calculs précédents dans le cadre du FGM, c'est-à-dire quand le noyau $K$ est de la forme :

$$K(\mathbf{x}) = (2\pi\lambda)^{-n/2} \exp\left(-\frac{\|\mathbf{x}\|^2}{2\lambda}\right),$$

pour $\lambda^2 > 0$ la variance des effets mutationnels sur les phénotypes. La première chose que nous devons calculer est le noyau $K'$ défini par (.43) :

$$\forall\, r \geq 0,\, \forall\, r' \geq 0,\; K'(r,r') = r^{n-2} \int_{\mathbb{S}^{n-1}} K(\|r\sigma - r'\|)\, \mathrm{d}\sigma.$$

Nous devons donc paramétrer la sphère $\mathbb{S}^{n-1}$, pour pouvoir calculer cette intégrale. Nous allons, pour ce faire, employer les coordonnées hypersphériques en dimension $n$ :

$$K'(r,r') = (2\pi\lambda)^{-n/2}\, r^{n-2} \exp\left(-\frac{r^2 + r'^2}{2\lambda}\right) \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} \exp\left(\frac{rr'}{\lambda}\cos(\Phi_1)\right) \prod_{j=1}^{n-2} \sin^{n-1-j}(\Phi_j)\, \mathrm{d}\Phi_{n-1} \cdots \mathrm{d}\Phi_1.$$

Notons d'une part que pour tout indice $j \geq 1$ :

$$\int_0^\pi \sin^j(\Phi)\, \mathrm{d}\Phi = \sqrt{\pi}\, \frac{\Gamma[(j+1)/2]}{\Gamma[(j+2)/2]},$$

avec $\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} \mathrm{d}t$ la fonction Gamma, ce qui permet de simplier :

$$\int_0^\pi \cdots \int_0^\pi \prod_{j=2}^{n-2} \sin^{n-1-j}(\Phi_j)\, \mathrm{d}\Phi_{n-2} \cdots \mathrm{d}\Phi_2 = \prod_{j=1}^{n-3} \left[\int_0^\pi \sin^j(\Phi)\, \mathrm{d}\Phi\right] = \frac{\pi^{(n-3)/2}}{\Gamma[(n-1)/2]}.$$

Enfin il nous reste à exprimer l'intégrale $I = \int_0^\pi \exp\left(\frac{rr'}{\lambda}\cos(\Phi)\right) \sin^{n-2}(\Phi)\, \mathrm{d}\Phi$. Le changement de variable $x = \cos(\Phi)$ implique que :

$$I = \int_{-1}^1 \exp\left(\frac{rr'}{\lambda}\, x\right) (1-x^2)^{(n-3)/2}\, \mathrm{d}x.$$

En utilisant le développement en série de la fonction $\exp$ et le théorème de Fubini, nous avons :

$$\begin{aligned}
I &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{rr'}{\lambda}\right)^k \int_{-1}^{1} x^k \, (1-x^2)^{(n-3)/2} \, \mathrm{d}x, \\
&= 2 \sum_{k=0}^{\infty} \frac{1}{(2k)!} \left(\frac{rr'}{\lambda}\right)^{2k} \int_{0}^{1} x^{2k} \, (1-x^2)^{(n-3)/2} \, \mathrm{d}x, \\
&= \sum_{k=0}^{\infty} \frac{1}{(2k)!} \left(\frac{rr'}{\lambda}\right)^{2k} \frac{\Gamma(k+1/2)\Gamma[(n-1)/2]}{\Gamma(k+n/2)}, \\
&= \sum_{k=0}^{\infty} \frac{\Gamma[(n-1)/2]}{\Gamma(k+n/2)} \frac{\sqrt{\pi}}{k!} \left(\frac{rr'}{2\lambda}\right)^{2k}, \\
&= \sqrt{\pi} \frac{\Gamma[(n-1)/2]}{\Gamma(n/2)} \sum_{k=0}^{\infty} \frac{1}{(n/2)_k} \frac{1}{k!} \left(\frac{rr'}{2\lambda}\right)^{2k}, \\
&= \sqrt{\pi} \frac{\Gamma[(n-1)/2]}{\Gamma(n/2)} \; {}_0F_1\left(;\frac{n}{2};\left(\frac{rr'}{2\lambda}\right)^2\right).
\end{aligned}$$

Ainsi les mutations agissent sur les fitness d'après le noyau :

$$K'(r,r') = \frac{2(2\lambda)^{-n/2}}{\Gamma(n/2)} \, r^{n-2} \, \exp\left(-\frac{r^2+r'^2}{2\lambda}\right) \; {}_0F_1\left(;\frac{n}{2};\left(\frac{rr'}{2\lambda}\right)^2\right).$$

Cette formule explicite du noyau permet d'obtenir une EDP fermée vérifiée par la MGF $M_p$ (ce qui implique une autre EDP satisfaite par $C_p$). D'après la remarque faite en début de Section C.1.3, nous devons tout d'abord exprimer l'intégrale suivante pour tout $m' < 0$ et $z \geq 0$ :

$$\begin{aligned}
\int_{-\infty}^{0} e^{zm} K'(\sqrt{2|m|}, \sqrt{2|m'|}) \, \mathrm{d}m &= \int_0^{+\infty} \frac{e^{-zx} x^{(n-2)/2}}{\Gamma(n/2)\lambda^{n/2}} e^{(m'-x)/\lambda} \sum_{k=0}^{\infty} \frac{1}{(n/2)_k} \frac{1}{k!} \left(\frac{x|m'|}{\lambda^2}\right)^k \mathrm{d}x, \\
&= \sum_{k=0}^{\infty} \frac{e^{m'/\lambda}}{\Gamma(k+n/2)} \frac{1}{k!} \frac{|m'|^k}{\lambda^{2k+n/2}} \int_0^{+\infty} e^{-(z+1/\lambda)x} x^{k+(n-2)/2} \mathrm{d}x,
\end{aligned}$$

vue que $(n/2)_k = \Gamma(k+n/2)/\Gamma(n/2)$. La définition de la fonction $\Gamma$ implique que :

$$\begin{aligned}
\int_{-\infty}^{0} e^{zm} K'(\sqrt{2|m|}, \sqrt{2|m'|}) \, \mathrm{d}m &= \sum_{k=0}^{\infty} \frac{e^{m'/\lambda}}{\Gamma(k+n/2)} \frac{1}{k!} \frac{|m'|^k}{\lambda^{2k+n/2}} \frac{\Gamma(k+n/2)}{(z+1/\lambda)^{k+n/2}}, \\
&= \exp\left[\frac{1}{\lambda}\left(m' - \frac{m'}{1+\lambda z}\right)\right] (1+z\lambda)^{-n/2}, \\
&= \exp\left[\frac{m'z}{1+\lambda z}\right] (1+z\lambda)^{-n/2} = M_*(z) \exp[m'(z+\omega(z))].
\end{aligned}$$

Ainsi la fonction génératrice des moments $M_p$ vérifie pour tout $t > 0$ et $z > 0$ :

$$\partial_t M_p(t,z) = \partial_z M_p(t,z) - \overline{m}(t) M_p(t,z) + U \left[ \int_{-\infty}^{0} e^{(z+\omega(z))m} p(t,m)\,\mathrm{d}m - M_p(t,z) \right],$$

puis l'équation non locale :

$$\partial_t M_p(t,z) = \partial_z M_p(t,z) - \overline{m}(t) M_p(t,z) + U \left[ M_p(z+\omega(z)) - M_p(t,z) \right].$$

## C.2. Approximation diffusive

Dans cette section, nous étudions le système différentiel, provenant de l'approximation diffusive (voir Section 1.3.2) :

$$\partial_t q(t,\mathbf{x}) \approx \sum_{i=1}^{n} \frac{\mu_i^2}{2} \partial_{x_i x_i} q(t,\mathbf{x}) + [m(\mathbf{x}) - \overline{m}(t)] q(t,\mathbf{x}).$$

avec $\mu_i^2 = U\,\lambda_i$.

### C.2.1. Propriétés sur la distribution des phénotypes

Comme $K$ est radiale (par hypothèse), la distribution $K$ vérifie :

$$\forall\, 1 \leq i \leq n,\ \mu_i = \int_{\mathbb{R}^n} x_i^2\, K(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} x_1^2\, K(\mathbf{x})\,\mathrm{d}\mathbf{x} = \mu_1 =: \mu.$$

Ainsi l'approximation diffusive revient au problème de Cauchy :

$$\begin{cases} \partial_t q(t,\mathbf{x}) = \dfrac{\mu^2}{2} \Delta q(t,\mathbf{x}) + [m(\mathbf{x}) - \overline{m}(t)] q(t,\mathbf{x}), & \forall t > 0,\ \forall \mathbf{x} \in \mathbb{R}^n, \\ q(0,\mathbf{x}) = q_0(\mathbf{x}), & \forall \mathbf{x} \in \mathbb{R}^n. \end{cases} \tag{.46}$$

Cette équation est un cas particulier du cas anisotrope (détaillé dans le Chapitre 2).

**Théorème 13.**
*Il existe une unique solution $q \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$ positive du problème de Cauchy :*

$$\left\{ \begin{array}{rcl} \partial_t q(t,\mathbf{x}) & = & \frac{\mu^2}{2}\Delta q(t,\mathbf{x}) + [m(\mathbf{x}) - \overline{m}(t)]\, q(t,\mathbf{x}), \quad \forall\, t \geq 0,\ \forall\, \mathbf{x} \in \mathbb{R}^n, \\ q(0,\mathbf{x}) & = & q_0(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n, \end{array} \right. \tag{.47}$$

*telle que $q \in L^\infty((0,T) \times \mathbb{R}^n)$ pour tout $T > 0$, et la fonction :*

$$t \mapsto \overline{m}(t) = \int_{\mathbb{R}^n} m(\mathbf{x})\, q(t,\mathbf{x})\, d\mathbf{x},$$

*soit bien définie et continue sur $\mathbb{R}_+$. De plus, la masse est conservée :*

$$\forall\, t \geq 0,\ \int_{\mathbb{R}^n} q(t,\mathbf{x})\, d\mathbf{x} = 1.$$

Quand initialement les phénotypes sont distribués selon une gaussienne, une solution explicite à (.46) peut être trouvée :

**Corollaire 2.**
*Supposons que la distribution phénotypique est initialement gaussienne, id est :*

$$\forall\, \mathbf{x} \in \mathbb{R}^n,\ q_0(\mathbf{x}) = (2\pi)^{-n/2}\Big(\prod_{i=1}^n (s_i^0)^{-1/2}\Big)\, \exp\left(-\sum_{i=1}^n \frac{(x_i - \overline{q}_i^0)^2}{2s_i^0}\right), \tag{.48}$$

*avec $\overline{q}_i^0 \in \mathbb{R}$ et $s_i^0 > 0$. Alors en tout temps $t \geq 0$, la distribution phénotypique $q(t,\mathbf{x})$ du problème de Cauchy (.47) est gaussienne :*

$$\forall\, t \geq 0,\ \forall\, \mathbf{x} \in \mathbb{R}^n,\ q(t,\mathbf{x}) = (2\pi)^{-n/2}\Big(\prod_{i=1}^n (s_i(t))^{-1/2}\Big)\, \exp\left(-\sum_{i=1}^n \frac{(x_i - \overline{q}_i(t))^2}{2s_i(t)}\right), \tag{.49}$$

*pour :*

$$\overline{q}_i(t) = \frac{\mu \overline{q}_i^0}{\mu\cosh(\mu t) + s_i^0 \sinh(\mu t)}, \quad \textit{et} \quad s_i(t) = \mu\, \frac{\mu\sinh(\mu t) + s_i^0\cosh(\mu t)}{\mu\cosh(\mu t) + s_i^0\sinh(\mu t)}. \tag{.50}$$

*De plus, la fitness malthusienne relative moyenne est donnée par :*

$$\overline{m}(t) = -\sum_{i=1}^n \frac{\overline{q}_i^2(t) + s_i(t)}{2}.$$

Ce corollaire implique la convergeance de la distribution $q(t,\mathbf{x})$ vers une distribution gaussienne de moyenne $\overline{q}_\infty = 0$ et de variances $s_{\infty,i} = \mu$, quand $t \to +\infty$. Cependant, un tel résultat ne semble pas possible, pour une autre condition

initiale.

### C.2.2. Distribution de fitness

Nous allons donc nous intéresser à la distribution de fitness $p(t,m)$ définie par la Proposition 48. Elle vérifie une EDP parabolique, dite dégénérée, dans le sens où le coefficient de diffusion peut s'annuler :

**Théorème 14.**
*La distribution de fitness $p(t,m)$ est une solution classique de classe $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}_-^*)$ de l'EDP :*

$$\forall\, t>0,\, \forall\, m<0,\, \partial_t p(t,m) = -\mu^2\, m\, \partial_{mm} p(t,m) + \mu^2 \left(\frac{n}{2} - 2\right) \partial_m p(t,m) + (m - \overline{m}(t))\, p(t,m), \quad (.51)$$

*avec la condition initiale définie par (.44).*

Nous connaissons peu de chose sur de telles équations. Il existe quelques résultats d'existence et d'unicité de solutions (Aronson; Crandall; Peletier 1982; Epstein; Mazzeo 2013; Favini; Marinoschi 2012) et même des schémas numériques pour approximer cette solution. Cependant, aucune formule analytique ne semble avoir été trouvée directement.

### C.2.3. Etudes des fonctions génératrices

Nous avons déjà vu que la distribution $p$ ne pouvait pas nous donner d'informations analytiques dans la Section C.1.2. Cependant, les fonctions génératrices, définies en Section C.1.3, nous avaient déjà permis d'obtenir des résultats sur le comportement de la fitness moyenne $\overline{m}$ en temps grand. Ainsi nous allons développer un raisonnement similaire, à partir de l'équation dégénérée (14), sur la fonction génératrice des cumulants $C_p$.

**Corollaire 3.**
*La fonction génératrice des cumulants $C_p$ de la distribution des fitness $p$ est une solution de classe $C^{1,1}(\mathbb{R}_+ \times \mathbb{R}_+)$ de :*

$$\begin{cases} \partial_t C_p(t,z) = (1-\mu^2 z^2)\partial_z C_p(t,z) - \dfrac{n}{2}\mu^2 z - \overline{m}(t), & t \geq 0,\ z \in \mathbb{R}_+, \\ C_p(0,z) = C_{p_0}(z), & z \in \mathbb{R}_+, \end{cases} \quad (.52)$$

*où la condition initiale $p_0$ est donnée par (.44) et $\overline{m}(t) = \partial_z C_p(t,0)$.*

Cette EDP a déjà été trouvée par Martin; Roques 2016, en modélisant directement la variation de la MGF $M_p$, sous l'effet de la sélection et de la mutation.

**Remarque 5.** *En dérivant l'EDP vérifiée par $C_p$ par rapport à $z$, nous obtenons :*

$$\partial_t \partial_z C_p(t,z) = -2\mu^2 z \partial_z C_p(t,z) + (1-\mu^2 z^2)\partial_{zz} C_p(t,z) - \frac{n}{2}\mu^2,$$

*ce qui implique, en évaluant ensuite en $z=0$, le théorème fondamental de Fisher (voir Théorème 1).*

La méthode présentée dans la Section 1.5.2.1 nous permet alors d'avoir une solution explicite au problème de Cauchy (.52) :

**Corollaire 4.**
*La fonction génératrice des cumulants $C_p$ de la distribution des fitness $p$ est donnée par :*

$$C_p(t,z) = \frac{n}{2}\log\left(\frac{\cosh(\mu\, t)\cosh(\operatorname{atanh}(\mu\, z))}{\cosh(\mu\, t + \operatorname{atanh}(\mu\, z))}\right) + C_{p_0}(\varphi(t,z)) - C_{p_0}(\varphi(t,0)),$$

*pour $t \geq 0$ et $0 \leq z < 1/\mu$, avec $\varphi(t,z) = (1/\mu)\tanh\Big(\mu\, t + \operatorname{atanh}(\mu\, z)\Big)$. De plus, nous avons :*

$$\overline{m}(t) = (1-\tanh^2(\mu t))\,\partial_z C_{p_0}\left(\frac{\tanh(\mu t)}{\mu}\right) - \frac{n\mu}{2}\tanh(\mu t). \tag{.53}$$

**Remarque 6.** *Nous venons de voir que $\overline{m}(t) \to -\frac{n\mu}{2} =: \overline{m}_\infty$ quand $t \to +\infty$. De plus, comme $C_p(t,z)$ tend vers $\frac{n}{2}\log[\cosh(\operatorname{atanh}(\mu z))]$, quand $t$ tend vers $+\infty$, nous pouvons montrer que la suite $(p(t,\cdot))_{t\geq 0}$ converge faiblement vers $p_\infty$, avec :*

$$\forall\, m < 0,\ p_\infty(m) = \frac{|m|^{\frac{n}{2}-1}}{\Gamma(n/2)\,\mu^{n/2}}\,\exp\left(\frac{m}{\mu}\right).$$

Ainsi par le biais d'une approximation du terme intégrale $\int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y})\,q(t,\mathbf{y})\,\mathrm{d}\mathbf{y}$, nous avons pu obtenir une approximation de la fitness Malthusienne relative moyenne $\overline{m}$ en tout temps. Comparons numériquement cette formule avec une approximation numérique de la solution de (.47).

### C.2.4. Validation numérique de l'appoximation diffusive

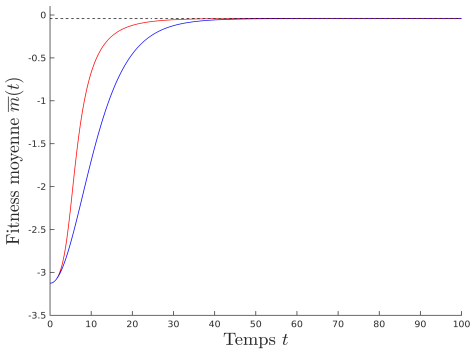
Cette section est dédiée à la comparaison entre les deux modèles précédents, décrits par les équations (.47) et (.52), dans le cas $n = 1$. Nous supposons que la distribution phénotypique initiale est une mesure de Dirac en $\mathbf{x}_0$. Dans ce cas, la distribution de fitness initiale est $p_0(m) = \delta_{m=-x_0^2/2}$ ce qui implique que la fonction génératrice des cumulants initiale est donnée par :

$$\forall\, z > 0,\, C_{p_0}(z) = -\frac{\|\mathbf{x}_0\|^2}{2}\, z.$$

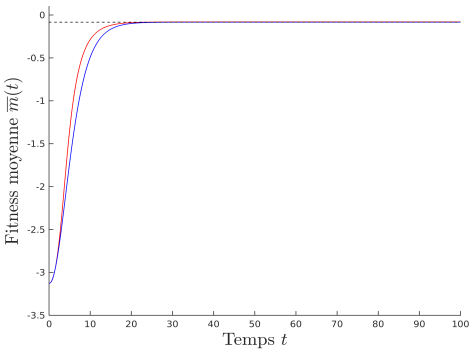
On suppose que la formule (.53) est encore vérifiée dans ce cadre, c'est-à-dire :

$$\overline{m}(t) = -\frac{\|\mathbf{x}_0\|^2}{2}\left(1 - \tanh^2(\mu t)\right) - \frac{n\mu}{2}\tanh(\mu t). \tag{.54}$$

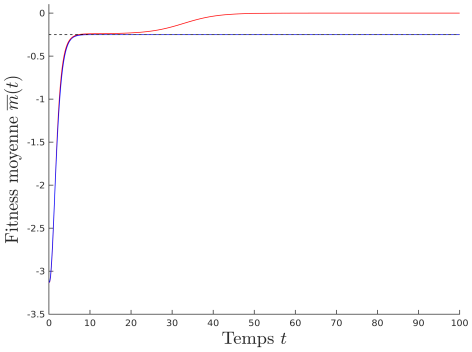
Grâce aux Fig. .12 (a) et (b), nous remarquons que l'approximation diffusive est d'autant meilleure que le taux de mutation $U$ est important : plus précisément, l'approximation diffusive donne une évolution de la fitness moyenne cohérente avec celle donnée par l'IDE (.41), pour des valeurs de $U$ très grandes devant un seuil critique $U_c = \lambda n^2/4$ (*cf.* Martin; Roques 2016). Pour de faibles variances $\lambda$ (ici $\lambda = 1/30$), la formule analytique de la limite $\overline{m}_\infty$ semble décrire correctement celle de la fitness moyenne du problème (.41). Par contre, si la variance $\lambda$ est trop importante (*e.g.* $\lambda = 0.1$), dans ce cas, l'approximation de la limite $\overline{m}_\infty = \mu n/2$ n'est plus valide (voir Fig. .12 (c)).




(a) $U = 25\,U_c$ et $\lambda = 1/30$




(b) $U = 50\,U_c$ et $\lambda = 1/30$




(c) $U = 50\,U_c$ et $\lambda = 0.1$

Figure .12. – **Evolution de la fitness moyenne** $\overline{m}(t)$. On suppose que la population initiale est distribuée selon une masse de Dirac $\delta_{x_0}$ pour $x_0 = 2.5$. On suppose que les mutations apparaissent avec un taux $U$ et sont distribuées de façon gaussienne avec une variance $\lambda$, avec (a) $U = 25\,U_c$ et $\lambda = 1/300$, (b) $U = 50\,U_c$ et $\lambda = 1/30$, (c) $U = 50\,U_c$ et $\lambda = 0.1$, pour $U_c = \lambda n^2/4$ valeur critique de validité donnée par Martin; Roques 2016. Les lignes rouges représentent la fitness moyenne $\overline{m}(t)$ donnée par une approximation numérique de la solution de (.47), les bleues correspondent à la formule analytique (.52), tandis que les lignes en pointillé est l'asymptote analytique $\overline{m}_\infty = \mu n/2$, pour $\mu = \sqrt{U\lambda}$.

## C.3. Expérience de Lenski

Cette dernière section est dédiée à la plus grande expérience évolutive menée le siècle dernier, et toujours en cours d'étude, nommée *Long-Term Evolutionary Experiment* ou LTEE.

### C.3.1. Contexte expérimental

En 1988, l'équipe de R. Lenski a lancé l'évolution de 12 populations d'*Escherichia Coli* (E. Coli) provenant de la même souche. Chacune évolue dans un environnement contenant une quantité limitée de glucose. Chaque jour $1\%$ de chaque population est prélevé pour la remettre dans un nouvel environnement identique à celui du début de l'expérience, et ainsi étudier l'évolution de la population. Le reste de chaque échantillon est congelé pour garder une trace, à chaque génération, de l'évolution de cette longue expérience.

Le premier but de cette expérience était de savoir si les nouvelles générations de bactéries resteraient identiques à leurs ancêtres, ou bien si elles évolueraient. De plus, si l'évolution avait lieu, chacune des 12 populations initiales évolueraient-elles toutes de la même manière ? Pour répondre à cela, la fitness est mesurée en comparant les individus de chaque prélèvement, avec les échantillons fossiles, grâce à un marqueur génétique neutre (qui n'a pas d'effet sur l'évolution des bactéries). Plus précisément, la fitness relative est définie par R. Lenski comme étant le quotient entre le taux de croissance de la population ayant évolué, et la précédente, qui sont en compétition l'un face à l'autre. En fait, cette mesure correspond plus précisément à la fitness Darwinienne $w(\text{gen})$, faite à la génération gen. En considérant qu'un cycle expérimental (une journée) – soit approximativement $6.64$ générations – correspond à une unité temporel $t$, nous avons la relation suivante entre la fitness Malthusienne relative $\overline{m}(t)$ et la fitness Darwinienne relative :

$$\overline{m}(t+2) - \overline{m}(2) = \ln\left[w(6.64t)\right].$$

Cette formule peut se retrouver à partir de la formule (1a) de Lenski; Rose; Simpson; Tadler 1991. Plus de détails biologiques sur la LTEE peuvent se trouver dans les différents articles, co-écrits par R. Lenski (*e.g.* Elena; Lenski 2003; Lenormand 2002; Lenski 2017; Lenski; Rose; Simpson; Tadler 1991; Lenski; Travisano 1994).

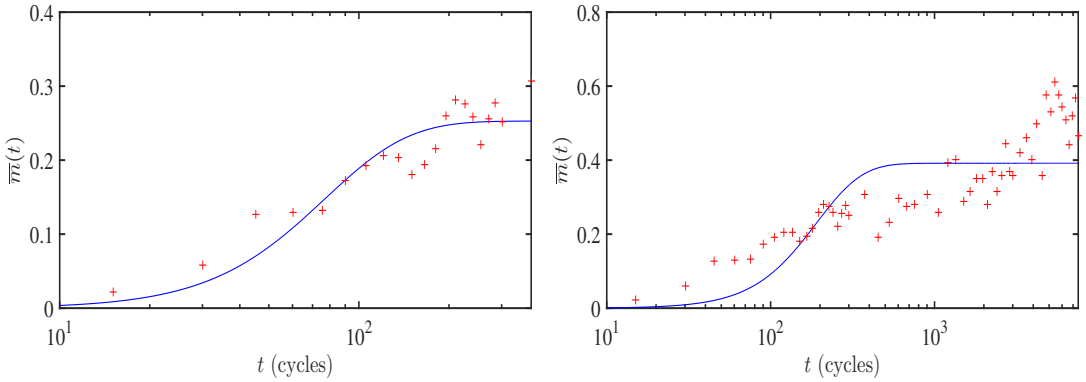
### C.3.2. Analytique vs. données

Comme mentionné dans la section précédente, les données récoltées recouvrent 12 populations différentes. Ici, nous n'utiliserons que les données d'une seule population – la population Ara-1 (Wiser; Ribeck; Lenski 2013). Les mesures ont été obenues à intervalle de 100 générations pour les 2 000 premières générations, puis à intervalle de 500.

En utilisant une méthode des moindres carrés non linéaire de Matlab® Curve Fitting Toolbox®, on affiche la meilleure approximation donnée par (.53), de la fonction $f(t) = \overline{m}(t+2) - \overline{m}(2)$. Nous supposerons que l'espace phénotypique est de dimension $n = 1$ et que la distribution initiale est une masse de Dirac en un phénotype $x_0$. Seulement deux paramètres doivent être estimés à savoir la

fitness $m_0$ du phénotype initial $x_0$ et le coefficient mutationnel $\mu$, ce qui donne les Figures .13.

Nous pouvons remarquer que le modèle isotrope décrit parfaitement bien l'évolution de la fitness, sur les premières générations ($\approx$ 400 cycles). Cela est confirmé par la valeur du $R^2$ qui vaut 0.88. Cependant, depuis peu de temps, la fitness moyenne connaît une nouvelle croissance, après avoir atteint un premier palier. Cela implique que la formule analytique donnée par (.54) ne traduit plus la réalité : le $R^2$ devient égal à 0.57. Ces dernières données remettent en question le modèle géométrique de Fisher et donc la présence d'un optimum phénotypique. Une autre possibilité, développée dans le prochain chapitre, est la présence d'anisotropie dans le noyau mutationnel.




(a) Premières générations (b) Toutes les générations

Figure .13. – **Comparaison de trajectoire adaptative de $\overline{m}$ dans le cas isotrope, avec les données de la LTEE.** La fonction $f(t) = \overline{m}(t + 2) - \overline{m}(2)$, avec $\overline{m}$ donnée par la formule (.54), est estimée avec les données de la population Ara-1, $\ln[w(6.64t)]$, de l'expérience LTEE (représentées par les croix rouges sur les graphiques). Elles sont données par la méthode des moindres carrées, et correspondent aux courbes bleues. En ne considérant que les premières générations, les paramètres données par la méthode des moindres carrés valent $m_0 = -0.2594$ et $\mu = 1.317 \cdot 10^{-2}$, tandis que si nous considérons toutes les générations, les valeurs sont $m_0 = -0.3968$ et $\mu = 5.33 \cdot 10^{-3}$.

# D. Blak-hole sink without epistasy

In this annex, the mutation effect on the fitness are not assumed epistatic. This hypothesis implies that the model (.42) is transformed into the differential equation:

$$\partial_t p(t,m) = U\left[\int_{-\infty}^{0} J(m-y)\,p(t,y)\,\mathrm{d}y - p(t,m)\right] + [m-\overline{m}(t)]\,p(t,m), \qquad (.55)$$

for all $t>0$ and $m<0$, where $J\in L^1(\mathbb{R})$ is a given mutational kernel. This kernel is assumed to satisfy:

$$\begin{cases} \int_{\mathbb{R}} J(m)dm = 1, \\ J(m)\geq 0, \text{ for a.e. } m\in\mathbb{R}, \\ \int_{\mathbb{R}} J(m)e^{x|m|}dm<\infty,\ \forall x>0. \end{cases} \qquad (.56)$$

The last hypothesis is equivalent to say that $J$ decays faster than any exponential function as $|m|\to+\infty$. By the same arguments as in Chapter C and 4, we can check that the CGF $C_p(t,z)$ is a solution of the problem:

$$\begin{cases} \partial_t C_p(t,z) = \partial_z C_p(t,z) - \overline{m}(t) + \beta(z) + \frac{d}{N(t)}\left(\mathbb{E}\left[e^{m_{migr}\ z}\right]e^{-C_t(z)}-1\right), & \forall\, t,\ z>0, \\ N'(t) = N(t)\,\overline{m}(t) + d, & \forall t>0, \\ C_p(t,0)=0, & \forall t\geq 0, \\ C_p(0,z)=C_{p_0}(z), & \forall z\geq 0, \\ N(0)=N_0, \end{cases} \qquad (.57)$$

where $d\geq 0, N_0\geq 0$, $\overline{m}(t)=\partial_z C_p(t,0)$ and the function $\beta$ is given by:

$$\beta(z) := U\left(\int_{\mathbb{R}} J(m)\,e^{m\ z}\mathrm{d}m - 1\right). \qquad (.58)$$

**Proposition 50.** *The problem (.57) admits a unique solution $C_p(t,z)$ defined for all $t \geq 0$ and $z \geq 0$. Futhermore, for every $t \geq 0$, the expected mean fitness is equal to:*

$$\overline{m}(t) = \frac{N_0 \left(\partial_z C_{p_0}(t) + \beta(t)\right) e^{C_{p_0}(t)} + d\, \mathbb{E}[e^{m_{migr}\ t}] - d\, e^{-\int_0^t \beta(s)\mathrm{d}s}}{N_0\, e^{C_{p_0}(t)} + d \int_0^t \mathbb{E}[e^{\tau m_{migr}}] e^{-\int_\tau^t \beta(s)\mathrm{d}s}\mathrm{d}\tau}, \tag{.59}$$

*and the expected population size:*

$$N(t) = \left[N_0 + d\int_0^t e^{-\int_0^\tau \overline{m}(s)\mathrm{d}s}\mathrm{d}\tau\right] e^{\int_0^t \overline{m}(s)\mathrm{d}s}. \tag{.60}$$

*The solution $C_p(t,z)$ of (.57) is given by the expression:*

$$\begin{aligned} C_p(t,z) = &\int_0^t \beta(z+\tau)\mathrm{d}\tau - \log\left(N_0 + d\int_0^t e^{-\int_0^\tau \overline{m}(s)\mathrm{d}s}\mathrm{d}\tau\right) - \int_0^t \overline{m}(\tau)\mathrm{d}\tau \\ &+ \log\left[N_0 e^{C_{p_0}(z+t)} + d\int_0^t \mathbb{E}[e^{(z+\tau)\ m_{migr}}] \exp\left(-\int_0^\tau \beta(z-s)\mathrm{d}s\right)\mathrm{d}\tau\right]. \end{aligned} \tag{.61}$$

The proof of this proposition is similar as in Chapter 4.

**Remark 8.** *Taking $d = 0$, we find the formula given in Gil; Hamel; Martin; Roques 2017:*

$$\overline{m}(t) = \partial_z C_{p_0}(t) + \beta(t).$$

We have already seen that the variation in mean fitness $\overline{m}'(t)$ and the variance in fitness $V(t)$ are connected by the Fisher's fundamental theorem (see Theorem 1). In the presence of non-epistatic mutations, formula of Theorem (1) extends to:

$$\overline{m}'(t) = V(t) + U\,\mu_J, \text{ with } \mu_J = \int_{\mathbb{R}} m\, J(m)\,\mathrm{d}m.$$

There we extend Fisher's fundamental theorem to take into account non-epistatic mutations and immigration events.

**Proposition 51.** *Let $\mu_J = \int_{\mathbb{R}} m\, J(m)\,\mathrm{d}m$. Then we have for all time $t$:*

$$\overline{m}'(t) = V(t) + U\mu_J + \frac{d}{N(t)}\left(\mathbb{E}[m_{migr}] - \overline{m}(t)\right). \tag{.62}$$

## D.1. Finitness of the characteristic time

We first give a precise meaning to the notion of "characteristic time":

$$t_0 = \sup\{t \geq 0, \overline{m}(t) \leq 0 \text{ in } [0,t]\}.$$

In the sequel, we assume that all of the individuals initially present in the sink population are maladapted, *id est*:

$$m_0 = \sup(\text{supp}(p_0)) < 0.$$

This implies that $\overline{m}(0) < 0$ and so $t_0 > 0$. We recall that $\mathbb{E}[m_{migr}] < 0$. We study here the finitness of the characteristic time $t_0$ depending on the type of mutation kernel $J$ (purely deleterious or including beneficial mutations) and of the distribution $p_{migr}$ of fitness $m_{migr}$ of the migrants.

**Theorem 52. (FINITNESS OF THE CHARACTERISTIC TIME)**

- ◇ *If the mutation kernel $J$ includes some beneficial mutations, then the characteristic time $t_0$ is finite;*
- ◇ *If the mutation kernel $J$ is purely deleterious:*

$$\textit{supp } J \cap \mathbb{R}^+ = \emptyset,$$

*and if $\mathbb{E}[e^{m_{migr}t}]e^{-Ut}$ tends to $+\infty$ as $t$ tends to $+\infty$, then the charateristic time $t_0$ is finite.*

- ◇ *If $\text{supp}(p_{migr}) \cap \mathbb{R}_+ = \emptyset$ and if $J$ is purely deleterious, then the characteristic time is infinite.*

*Proof.* See Section D.4. □

In the last case, the population can never adapt in the new environment. All of the immigrants are indeed maladapted and the mutation cannot improve the offspring fitness. However, the population size stabilizes: for $t$ large enough, $N'(t)$ is approximately equal to $\overline{m}(+\infty)N(t) + d$, and so:

$$\lim_{t \to +\infty} N(t) = d\,\mathbb{E}[(U - m_{migr})^{-1}].$$

**Remarque 7.** *Assume that $m_{migr}$ is Gaussian distributed, which implies that:*

$$\mathbb{E}[e^{m_{migr}t}]e^{-Ut} = \exp\left((\mu - U)t + \frac{\sigma^2 t^2}{2}\right),$$

*which diverges to $+\infty$ as $t \to +\infty$. So, by the second case of Theorem 52, we know that the characteristic time is finite, while $J$ is purely deleterious: although all of the mutations are deleterious, selection is strong enough to drive $\overline{m}(t)$ towards positive values.*

## D.2. Explicit formulae for the characteristic time

We will assume that the hypotheses of Theorem 52 are satisfied so that the characteristic time is finite. For the sake of simplicity, we assume that the sink population is initially empty before the first immigration event, which occurs at $t = 0$.

Thanks to equation (.59), we have:

$$\mathbb{E}[e^{m_{migr}\ t_0}] = e^{-\int_0^{t_0}\beta(s)\mathrm{d}s}, \tag{.63}$$

with:

$$\beta(z) = U\left(\int_{\mathbb{R}} J(m)\ e^{zm}\mathrm{d}m - 1\right).$$

To analyse the dependence of $t_0$ with respect to the model parameters $d$, $U$, $p_{migr}$ and $J$, we will see it as a function of four variables:

$$t_0 = t_0(d, U, p_{migr}, J).$$

For the sake of clarity, we will not write the parameters of the function $t_0$.

**Remarque 8.** *The function $\beta$ is strictly convex.*

For kernels $J$ including beneficial mutations, and with positive mean effect on fitness, we have the upper bound for $t_0$:

**Proposition 53.** *If $\mu_J = \int_{\mathbb{R}} m\, J(m)\, \mathrm{d}m > 0$, then we have:*

$$t_0 \leq -2\,\frac{\mathbb{E}[m_{migr}]}{U\mu_J}.$$

*Proof.* See Section D.4. □

For kernels $J$ with negative mean, we obtain the following bounds:

**Proposition 54.** *Assume that $\mu_J = \int_{\mathbb{R}} m\, J(m)\, \mathrm{d}m < 0$. Then there exists a unique $\tau > 0$ such that $\beta(\tau) = 0$. Moreover, we have $\beta'(\tau) > 0$ and:*

$$\tau < t_0 < \tau - \frac{\mathbb{E}[m_{migr}]}{\beta'(\tau)} + \sqrt{\left(\frac{\mathbb{E}[m_{migr}]}{\beta'(\tau)} - \tau\right)^2 - \tau^2\, \frac{\beta'(0)}{\beta'(\tau)}}.$$

*Proof.* See Section D.4. □

### Dirac distribution of $m_{migr}$

Now, we will assume that $m_{migr}$ has a Dirac distribution $\delta_{m_{migr}}$. Using the equation (.63) we immediately get:

**Proposition 55. (CLONAL MIGRANTS, GENERAL MUTATION KERNEL)**

$$m_{migr}\, t_0 = -\int_0^{t_0} \beta(s)\, \mathrm{d}s, \tag{.64}$$

*where $\beta(z) = U\,(\int_{\mathbb{R}} J(m)\, e^{zm} \mathrm{d}m - 1)$.*

So we have:

$$\int_0^{t_0} \beta(s)\, \mathrm{d}s \geq 0, \quad \beta(t_0) > 0 \quad \text{and} \quad \beta'(t_0) > 0.$$

Using expression (.64), we easily describe the dependence of the characteric time with respect to the model parameters.

**Proposition 56.** *The function $t_0(d, U, m_{migr}, J)$ does not depend to $d$. Moreover, it satisfies:*

$$\frac{\partial t_0}{\partial m_{migr}} = -\frac{t_0}{\beta(t_0) + m_{migr}} < 0, \tag{.65}$$

*and:*

$$\frac{\partial t_0}{\partial U} = \frac{1}{U}\frac{m_{migr}\, t_0}{\beta(t_0) + m_{migr}} < 0. \tag{.66}$$

*Proof.* See Section D.4. □

These results show that the highter the mutation rate, the more there are mutations, and so the less the population needs time to adapt. Also if $m_{migr}$ is near zero, then the population will need less time to adapt, than if $|m_{migr}|$ is high.

We now derive exact and approximate formulae for three different types of mutation kernels.

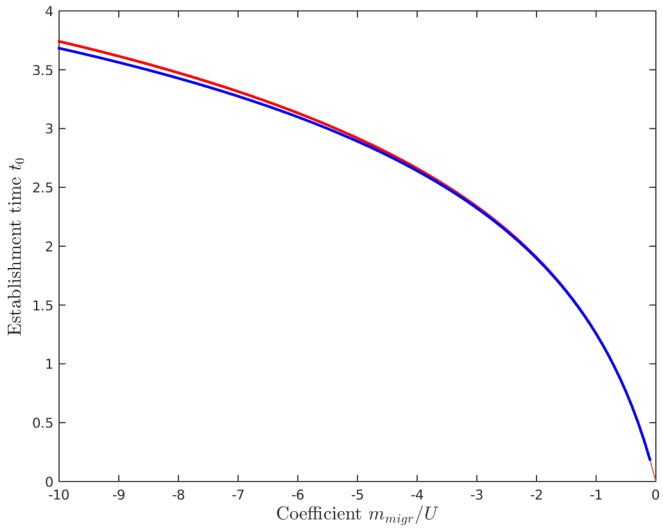



Figure .14. – **Dynamics of $t_0$ with respect to $\frac{m_{migr}}{U}$.** We asume here a Dirac distribution of migrant fitnesses ($p_{migr} = \delta_{m_{migr}}$) and a Dirac mutation kernel $J = \delta_{\mu_J}$ ($\mu_J = 1$). The dark curve represents the exact solution (.67) and the dashed red curve is the approximation (.68).

Here, we assume $J = \delta_{\mu_J}$, with $\mu_J > 0$ (otherwise, $t_0 = +\infty$, by Theorem 52). Thus, we have $\beta(t) = U\left(e^{\mu_J t} - 1\right)$ and so the characteristic time satisfies (thanks to (.64)):

$$e^{\mu_J t_0} + \frac{\mu_J}{U}\left(m_{migr} - U\right) t_0 - 1 = 0.$$

The solutions of this equation are given by:

$$t_0 = \frac{U}{\mu_J\left(m_{migr} - U\right)} - \frac{1}{\mu_J} W\left[\frac{U}{m_{migr} - U} \exp\left(\frac{U}{m_{migr} - U}\right)\right],$$

where $W$ is the Lambert-$W$ function (see Appendix 1.5.4).

However, if the branch is $W_0$ then $t_0 = 0$, which is impossible. Hence, we get:

$$t_0 = \frac{U}{\mu_J\left(m_{migr} - U\right)} - \frac{1}{\mu_J} W_{-1}\left[\frac{U}{m_{migr} - U} \exp\left(\frac{U}{m_{migr} - U}\right)\right]. \tag{.67}$$

As $\beta(t_0) + 4\beta(t_0/2) \approx -6m_{migr}$, we can approximate the establishment time $t_0$ by:

$$t_0 \approx \frac{2}{\mu_J} \log\left(-2 + \sqrt{9 - \frac{6\ m_{migr}}{U}}\right). \tag{.68}$$

- **Gaussian mutation kernel**

Here, we assume that the distribution of mutation effects on fitness, $J(m)$, is a Gaussian function mean value $\mu_J$ and variance $\sigma_J^2$:

$$J(m) = \frac{1}{\sqrt{2\pi\sigma_J^2}} \exp\left(-\frac{(m-\mu_J)^2}{2\ \sigma_J^2}\right).$$

Then, we have:

$$\beta(t) = U\left[\exp\left(\mu_J\ t + \frac{\sigma_J^2 t^2}{2}\right) - 1\right].$$

The equation (.64) yields:

$$-m_{migr}\ t_0 = \int_0^{t_0} \beta(s)\mathrm{d}s = U\sqrt{\frac{\pi}{2\sigma_J^2}}\ e^{-\frac{\mu_J^2}{2\sigma^2}}\left[\text{Erfi}\left(\frac{\mu_J + \sigma_J^2 t_0}{\sqrt{2\sigma_J^2}}\right) - \text{Erfi}\left(\frac{\mu_J}{\sqrt{2\sigma_J^2}}\right)\right] - U\ t_0,$$

where Erfi is the imaginary error function:

$$\text{Erfi}(x) = \frac{\text{Erf}(ix)}{i} = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}\mathrm{d}t.$$

We can get approximate $t_0$ by:

$$t_0 \approx -\frac{\mu_J}{\sigma_J^2} + \sqrt{\frac{\mu_J^2}{\sigma_J^4} + \frac{2}{\sigma_J^2}\log\left(1 - 2\frac{m_{migr}}{U}\right)}.$$

- **Symetrized gamma mutation kernel**

Let $k \in \mathbb{N} \setminus \{1\}$, and $\theta$ and $s$ be two positive real numbers. Now we assume that for all $m \in \mathbb{R}$:

$$J(m) = \begin{cases} \dfrac{(s-m)^{k-1}}{\Gamma(k)\ \theta^k} \exp\left(-\dfrac{s-m}{\theta}\right), & \text{if } m < s, \\ 0, & \text{if not.} \end{cases}$$

This means that the kernel $J$ contains beneficial mutations, with a maximum

effect on fitness $s$ . Thus, we get:

$$\beta(t) = U\left(\frac{e^{st}}{(1+t\theta)^k} - 1\right).$$

For example, for $k = 2$, we have:

$$\begin{aligned}\int_0^{t_0} \frac{e^{sx}}{(1+x\theta)^2}\mathrm{d}x &= \left[-\frac{e^{sx}}{\theta(1+\theta x)}\right]_0^{t_0} + \frac{1}{\theta}\int_0^{t_0}\frac{e^{sx}}{1+x\theta}\mathrm{d}x,\\ &= \frac{1}{\theta}\left[1 - \frac{e^{st_0}}{1+\theta t_0}\right] + \frac{s}{\theta}e^{-s/\theta}\int_{a/b}^{a(bt_0+1)/b}\frac{e^x}{x}\mathrm{d}x,\\ &= \frac{1}{\theta}\left[1 - \frac{e^{st_0}}{1+\theta t_0}\right] + \frac{s}{\theta}e^{-s/\theta}\left[\mathrm{Ei}\left(a\,\frac{bt_0+1}{b}\right) - \mathrm{Ei}\left(\frac{a}{b}\right)\right],\end{aligned}$$

where Ei is the exponential integral:

$$\mathrm{Ei}(x) = \int_{-\infty}^{x}\frac{e^t}{t}\mathrm{d}t.$$

Again, we can compute the approximation of $t_0$:

$$t_0 \approx -\frac{1}{\theta} - \frac{2}{s}W_{-1}\left[-\frac{s}{2\theta}\sqrt{\frac{U}{U-m_{migr}}}\exp\left(-\frac{s}{2\theta}\right)\right].$$

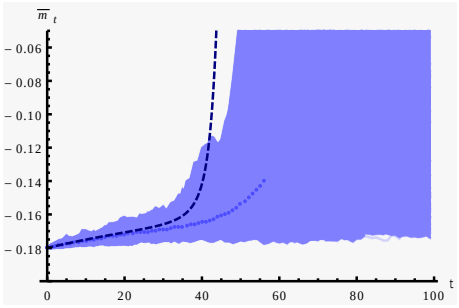
## D.3. Comparison with the IBM

We assume that the ditribution of the migrants $p_{migr}$ is a Dirac at $m_{migr} = -0.18$. We take the migration rate $d$ equal to 90 (such that $d = -500m_{migr}$), and the mean of the mutation kernel $\mu_J$ to 0.1. We focus on two cases: Dirac mutation kernel and Gaussian mutation kernel, with variance $\sigma_J^2 = 0.01$. To study weak and strong mutation events, we take the mutation rate equals respectively to either $0.1 \times \mu_J$ $(= 0.01)$ or $\mu_J$.

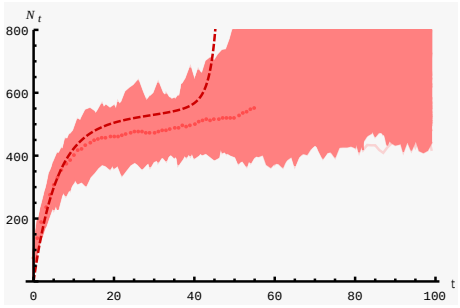
Our PDE framework (.57) gives accurate results for both the prediction of $\overline{m}(t)$ and $N(t)$ (respectively (.59) and (.60)) at small times (Figures .15 and .16). Then, the accuracy of the predictions of our PDE framework at layer times depends on the value of the mutation rate. For low mutation rates (upper panels in Figures .15 and .16), the mean fitness $\overline{m}(t)$ tends to be overestimated by our theory: this is probably a consequence of the "large population size assumption". As mutations are rare, they often do not even occur in the Individual-Based Model framework, whereas the PDE framework assumes that they always occur, but with a low rate. For higher mutation rates (lower panels in Figures .15 and .16) the PDE theory gives satisfactory results, even at large times.

Futhermore, the dynamics in the sink has three phases, with an inversion of the convexity of $N(t)$: for a short time, the quantity is rapidly increasing,

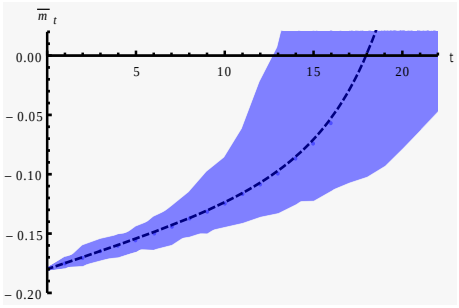
thanks to the migration events, then becomes stable (approximately equal to $-d/m_{migr}$), because of the selection, and finally is rapidly increasing again, with an exponential growth, due to mutations (right panels in Figures .15 and .16).



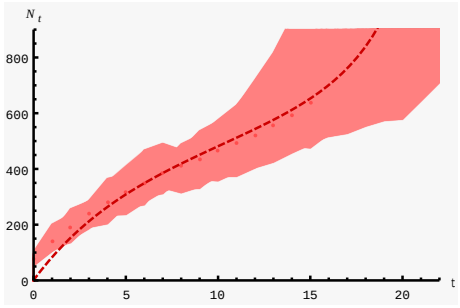

(a) $U = 0.01$




(b) $U = 0.01$

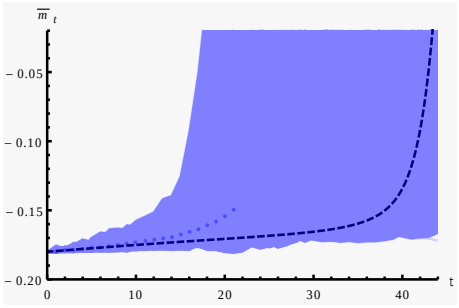



(c) $U = 0.1$

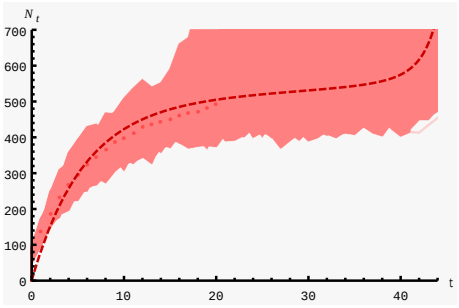



(d) $U = 0.1$

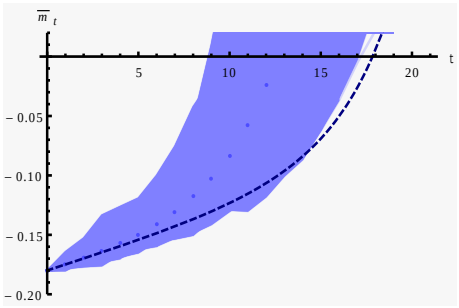
Figure .15. – **Individual-Based simulations vs PDE theory**. We assume here a Dirac distribution of migrant fitnesses ($p_{migr} = \delta_{m_{migr}}$) and a Dirac mutation kernel $J = \delta_{\mu_J}$. (a,c): dynamics of the expected mean fitness $\overline{m}(t)$ in the sink. (b,d): dynamics of the expected population size $N(t)$ in the sink. The dark dashed lines represent the solution given by our PDE theory (respectively (.59) and (.60)). The light dotted lines represent the mean values obtained by simulations. The shaded regions correspond to higher and lower values obtained by Individual-Based simulations. The parameters have been chosen as $\mu_J = 0.1$, $m_{migr} = -0.18$ and $d = 90$.



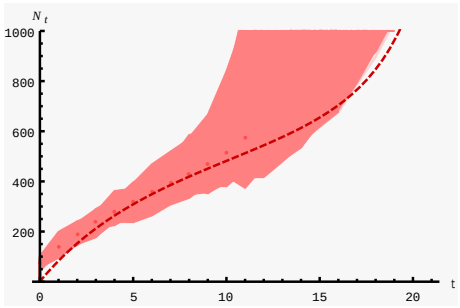

(a) $U = 0.01$




(b) $U = 0.01$

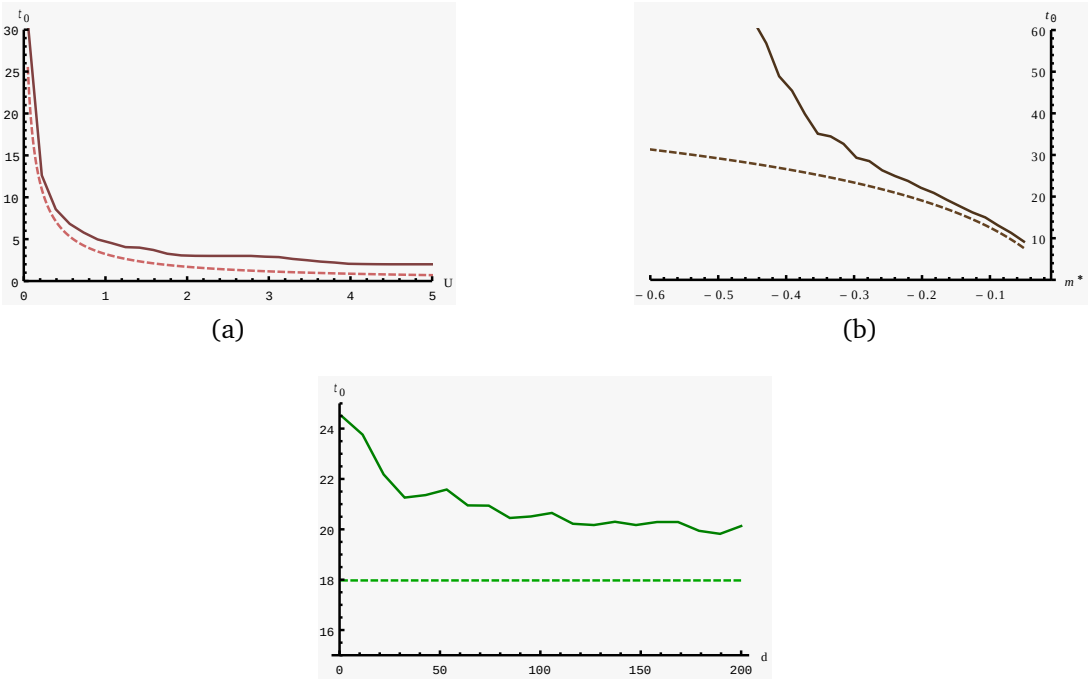



(c) $U = 0.1$




(d) $U = 0.1$

Figure .16. – **Individual-Based simulations vs PDE theory**. We assume here a Dirac distribution of migrant fitnesses ($p_{migr} = \delta_{m_{migr}}$) and a Gaussian mutation kernel $J(m) = \frac{1}{\sqrt{2\pi\sigma_J^2}} \exp\left(-\frac{(m-\mu_J)^2}{2\sigma_J^2}\right)$. (a,c): dynamics of the expected mean fitness $\overline{m}(t)$ in the sink. (b,d): dynamics of the expected population size $N(t)$ in the sink. The dark dashed lines represent the solution given by our PDE theory (respectively (.59) and (.60)). The light dotted lines represent the mean values obtained by Individual-Based simulations. The shaded regions correspond to higher and lower values obtained by Individual-Based simulations. The parameters have been chosen as $\mu_J = 0.1$, $\sigma_J^2 = 0.01$, $m_{migr} = -0.18$ and $d = 90$.




(a) (b)

(c)

Figure .17. – **Dynamics of the characteristic time.** We assume here a Dirac distribution of migrant fitnesses ($p_{migr} = \delta_{m_{migr}}$) and a Dirac mutation kernel $J = \delta_{\mu_J}$ with $\mu_J = 0.1$. (a) with respect to $d$ ($U = \mu_J$, $m_{migr} = -0.18$). (b) with respect to $U$ ($d = 90$, $m_{migr} = -0.18$). (c) with respect to $m_{migr}$ ($d = 90$, $U = \mu_J$). The plain lines represent the mean value obtained with Individual-Based simulations, and the dashed lines are the analytic value of $t_0$, given by (.67).

Now we analyse the dependence of $t_0$ with respect to the model parameters, namely the migration rate ($d$), migrant fitness ($m_{migr}$) and the mutation rate ($U$) (Figure .17). The PDE theory and the Individual-Based simulations give consistent results regarding the dependence of $t_0$ with respect to $U$ (Figure .17 (b)) and $m_{migr}$ (Figure .17 (c)). However, our PDE theory predicts (when $N_0 = 0$) that $t_0$ is independent of $d$, which is apparently not the case in our simulations (Figure .17 (a)). The function $t_0(d)$ obtained with the simulations seems to converge towards the value predicted by the PDE theory for large $d$, as in the epistatic case (see Chapter 4).

## D.4. Proofs of the results

In this section, we develop the different proofs of the results given in Sections D.1 and D.2.

First, let us remark that, from the properties (.56) of $J$, we get for all $z \geq 0$:

$$\beta(z) = U\left(\int_{\mathbb{R}} J(m)\, e^{z\ m}\mathrm{d}m - 1\right) \text{ and } \beta'(z) = U\int_{R} m\ J(m)\ e^{z\ m}\mathrm{d}m.$$

In particular, we have $\beta'(0) = U\int_{\mathbb{R}} m\ J(m)\mathrm{d}m = U\mu_J$, with $\mu_J$ the mean of the mutation rate.

**Proposition 57.** *Assume that $J$ includes some beneficial mutations:*

$$\mathit{supp}(J) \cap (0, \infty) \neq \emptyset.$$

*Then* $\lim_{z\to\infty} \beta(z) = \lim_{z\to\infty} \beta'(z) = +\infty$.

*Proof.* See Gil; Hamel; Martin; Roques 2017. □

**Proposition 58.** *Let $m_0 \in (-\infty, +\infty]$ be defined by:*

$$m_0 = \sup(\mathit{supp}(p_0)).$$

*The function $C_{p_0}$ is convex and $\partial_z C_{p_0}(z) \to m_0$ as $z \to +\infty$. Furthermore, if $m_0 < +\infty$, then:*

$$\lim_{z\to\infty} \partial_{zz} C_{p_0}(z) = 0.$$

*Proof.* See Gil; Hamel; Martin; Roques 2017. □

*Proof.* of Theorem 52.

- ◇ Assume that $J$ includes some beneficial mutations. Thanks to Propositions 57 and 58, we know that for $t$ large enough:

$$N_0(\partial_z C_{p_0}(t) + \beta(t))e^{C_{p_0}(t)} > 0.$$

Thus, using Proposition .59, we get:

$$\begin{aligned}\overline{m}(t) &= \frac{N_0\ (\partial_z C_{p_0}(t) + \beta(t))\ e^{C_{p_0}(t)} + d\ \mathbb{E}[e^{m_{migr}\ t}] - d\ e^{-\int_0^t \beta(s)\mathrm{d}s}}{N_0\ e^{C_{p_0}(t)} + d\int_0^t \mathbb{E}[e^{\tau m_{migr}}]e^{-\int_\tau^t \beta(s)\mathrm{d}s}\mathrm{d}\tau},\\ &\geq \frac{d\ \mathbb{E}[e^{m_{migr}\ t}]e^{\int_0^t \beta(s)\mathrm{d}s} - d}{N_0\ e^{\int_0^t (\partial_z C_{p_0}(s)+\beta(s))\mathrm{d}s} + d\int_0^t \mathbb{E}[e^{m_{migr}\tau}]e^{\int_0^\tau \beta(s)ds}\mathrm{d}\tau}.\end{aligned}$$

Thanks to Proposition 57, we know that $\beta$ diverges to $+\infty$ and so:

$$\mathbb{E}[e^{m_{migr}t}]e^{\int_0^t \beta(s)\mathrm{d}s} \geq \int_0^{+\infty} e^{ty} p_{migr}(y)\mathrm{d}y \; e^{\int_0^t \beta(s)\mathrm{d}s}.$$

diverges too. Thus for $t$ large enough $\overline{m}(t)$ is positive, which gives the results.

- ◇ We prove the expected results, by the same arguments and using the fact that $\beta(t) \to -U$ as $t$ tends to $+\infty$.
- ◇ Let us prove that the numerator (denoted here by $h(t)$) in the formula (.59) is negative for all $t \geq 0$.

  First, we know that $\beta$ is convex, with $\beta(0) = 0$. Additionally, we can prove that $\beta(t)$ tends to $-U$ as $t$ tends to $+\infty$: so $\beta(t)$ is nonpositive for all $t \geq 0$. Second, Proposition 58 yields that for all $t \geq 0$, $\partial_z C_{p_0}(t) < 0$, thanks to the hypothesis $m_0 < 0$. So we have:

$$f(t) < d\mathbb{E}[e^{m_{migr}}] - de^{-\int_0^t \beta(s)\mathrm{d}s}.$$

  This upper bound is a decreasing function, and at $t = 0$ is equal to 0. Therefore, we get $f(t) < 0$ which gives $\overline{m}(t) < 0$ and so the characteristic time is infinite.

□

*Proof.* of Proposition 53.

By convexity of $\beta$, we have for all $t > 0, \beta(t) \geq \beta'(0)t$. Thus we get, thanks to $\beta'(0) = U\mu_J$:

$$\int_0^{t_0} \beta(s)\mathrm{d}s \geq U\mu_J \, \frac{t_0^2}{2}.$$

Jensen's inequality (thanks to the convexity of the exponential function) yields so that:

$$e^{t_0 \, \mathbb{E}[m_{migr}]} \leq \mathbb{E}\left[e^{t_0 m_{migr}}\right] \leq e^{-U\mu_J \, \frac{t_0^2}{2}},$$

which gives the result (because $t_0 > 0$). □

*Proof.* of Proposition 54.

Thanks to the strict convexity of $\beta$, the existence and uniqueness of $\tau > 0$ such that $\beta(\tau) = 0$ follows from the properties:

$$\beta(0) = 0, \quad \beta'(0) = U\mu_J < 0, \quad \text{and} \quad \beta(+\infty) = +\infty.$$

Always by convexity, it is clear that $\beta'(\tau) > 0$ and $t_0 > 0$.

Furthermore, the convexity of $\beta$ yields that:

$$\begin{aligned}\int_0^{t_0} \beta(s)\mathrm{d}s &\geq \int_0^{\tau} \beta(s)\mathrm{d}s + \int_{\tau}^{t_0} \beta(s)\mathrm{d}s,\\ &\geq \int_0^{\tau} \beta'(0)\ s\mathrm{d}s + \int_{\tau}^{t_0} \beta'(\tau)\ (s-\tau)\mathrm{d}s,\\ &\geq \beta'(0)\ \frac{\tau^2}{2} + \beta'(\tau)\ \frac{(t_0-\tau)^2}{2}.\end{aligned}$$

Thanks to (.63), the Jensen inequality yields that:

$$e^{t_0\ \mathbb{E}[m_{migr}]} \leq \mathbb{E}\left[e^{t_0 m_{migr}}\right] \leq e^{-\beta'(0)\ \frac{\tau^2}{2} - \beta'(\tau)\ \frac{(t_0-\tau)^2}{2}}.$$

Hence we have:

$$P(t_0) = t_0^2 + 2\left(\frac{\mathbb{E}[m_{migr}]}{\beta'(\tau)} - \tau\right) t_0 + \tau^2\ \frac{\beta'(0)}{\beta'(\tau)} \leq -\tau^2 < 0.$$

The dicriminant of $P$ is:

$$\Delta = 4\left[\left(\frac{\mathbb{E}[m_{migr}]}{\beta'(\tau)} - \tau\right)^2 - \tau^2\ \frac{\beta'(0)}{\beta'(\tau)}\right],$$

which is negative (because $\beta'(0) = U\mu_J < 0$ and $\beta'(\tau) > 0$). Therefore $P$ has two roots $x_- < 0$ and $x_+ > 0$ defined by:

$$x_\pm = \tau - \frac{\mathbb{E}[m_{migr}]}{\beta'(\tau)} \pm \sqrt{\left(\frac{\mathbb{E}[m_{migr}]}{\beta'(\tau)} - \tau\right)^2 - \tau^2\ \frac{\beta'(0)}{\beta'(\tau)}}.$$

The sign of $P(t_0)$ gives the result. □

Now, let us turn to the case, when the migrant fitness distribution is a Dirac distribution at $m_{migr}$.

**Proposition 59.** *We have:*

$$\frac{\beta(t_0)\ t_0}{2} > \int_0^{t_0} \beta(s)\mathrm{d}s = -m_{migr}\ t_0.$$

*Proof.* This is given by strict convexity of $\beta$. □

*Proof.* of Proposition 56.

The relation (.64) yields clearly the independence of $t_0$ to $d$. The equality (.64) yields:

$$t_0 + m_{migr}\,\frac{\partial t_0}{\partial m_{migr}} = -\beta(t_0)\,\frac{\partial t_0}{\partial m_{migr}},$$

which gives the formula (.65). The sign is a consequence of it, using Proposition 59.

Thanks to the definition of $\beta$, we get, for all $U > 0$:

$$m_{migr}\,\frac{\partial t_0}{\partial U} = -\frac{1}{U}\int_0^{t_0} \beta(s)\mathrm{d}s - \beta(t_0)\frac{\partial t_0}{\partial U}.$$

Adding the relation (.64), we get:

$$\frac{\partial t_0}{\partial U} = \frac{1}{U}\frac{m_{migr}\;t_0}{\beta(t_0) + m_{migr}}.$$

By Proposition 59, we get $\frac{\partial t_0}{\partial U} < 0$. □