Spatial statistics and stochastic partial differential equations: A mechanistic viewpoint

Lionel Roques, Denis Allard, Samuel Soubeyrand

Spatial Statistics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Spatial statistics and stochastic partial differential equations: A mechanistic viewpoint

Lionel Roques, Denis Allard *, Samuel Soubeyrand

Biostatistics and Spatial Processes (BioSP), INRAE, 84914 Avignon, France

ARTICLE INFO

Article history:
Received 1 November 2021
Received in revised form 4 January 2022
Accepted 5 January 2022
Available online xxxx

Keywords:
Fractional Laplacian
Matérn covariance
PDEs
Gaussian random fields
Dispersal kernels
Spatio-temporal point processes

ABSTRACT

The Stochastic Partial Differential Equation (SPDE) approach, now commonly used in spatial statistics to construct Gaussian random fields, is revisited from a mechanistic perspective based on the movement of microscopic particles, thereby relating pseudo-differential operators to dispersal kernels. We first establish a connection between Lévy flights and PDEs involving the Fractional Laplacian (FL) operator. The corresponding Fokker–Planck PDEs will serve as a basis to propose new generalisations by considering a general form of SPDE with terms accounting for dispersal, drift and reaction. We detail the difference between the FL operator (with or without linear reaction term) associated with a fat-tailed dispersal kernel and therefore describing long-distance dependencies, and the damped FL operator associated with a thin-tailed kernel, thus corresponding to short-distance dependencies. Then, SPDE-based random fields with non-stationary external spatially and temporally varying force are illustrated and nonlinear bistable reaction terms are introduced. The physical meaning of the latter and possible applications are discussed. Returning to the microscopic interpretation of the above-mentioned equations, we describe in a relatively simple case their links with point processes. We unravel the nature of the point processes they generate and show how such mechanistic models, associated to a probabilistic observation model, can be used in a hierarchical setting to estimate the parameters of the particle dynamics.

* Corresponding author.
E-mail address: denis.allard@inrae.fr (D. Allard).

https://doi.org/10.1016/j.spasta.2022.100591
2211-6753/

1. Introduction

Diffusion processes are commonly used for modelling natural space–time processes in climate sciences, ecology, epidemiology and population dynamics, to name a few. For example, the diffusion equation, $\partial_t u(x,t) = \Delta u(x,t)$ with $x \in \mathbb{R}^2$, $t > 0$, one of the most elementary partial differential equations (PDEs) known as the heat equation, describes the macroscopic behaviour of many micro-particles resulting from their random movements. How PDEs such as the heat equation, but also more complex ones, interact with spatial statistics is the focus of this contribution.

PDEs and statistics have long been connected through the usual question of parameter estimation in mechanistic models (see e.g., Wikle, 2003; Soubeyrand and Roques, 2014, and the references therein). About ten years ago, Lindgren et al. (2011) revisited the explicit link between stochastic partial differential equations (SPDEs) and spatial statistics already established in Whittle (1954, 1963). They introduced a new paradigm for the statistical analysis of spatial data that consists in modelling a spatial variable as the solution of a certain class of SPDE for which the representation on a finite grid presents interesting Markov properties, making it computable for very large grids and capable of handling very large datasets. This approach was based on the observation made in Whittle (1954, 1963) that, on $\mathbb{R}^d$, the solution of the SPDE

$$(\kappa^2 - \Delta)^{\alpha/2} U = W_S, \tag{1}$$

is a Gaussian random field whose covariance function is proven to be a Matérn covariance. In (1) $\alpha > d/2$, Δ is the usual Laplacian, W_S is a Gaussian (spatial) White Noise process and U is the unknown stationary random field. For $\alpha \neq 2$, the pseudo-differential operator in (1) was constructed on an ad hoc basis to get a Matérn covariance. As Whittle (1954) himself indicated "it is difficult to visualise a physical mechanism which would lead to (1)". However, in contrast to the previous statistically oriented constructions, the SPDE approach should also lead to a physically grounded construction, for which the parameters would carry traditional physical interpretation such as diffusivity, reaction and transport. The SPDE approach also allows the construction of models with interesting properties, such as non-symmetry and space–time non-separability (Brown et al., 2000). Bolin and Kirchner (2020) proposed a rational SPDE approach for Gaussian random fields defined by (1) with general smoothness parameter $\alpha > d/4$ and Bolin and Wallin (2020) proposed multivariate non-Gaussian extensions based on systems of SPDEs with additive type G noise whose marginal covariance functions are of Matérn type.

Non-stationarity can also easily be accounted for by letting the parameters in (1) be space-dependent, see e.g. Lindgren et al. (2011) and Fuglstad et al. (2015). Space–time generalisation is not straightforward. In particular, the space–time covariance is not always accessible (but the spectral density is) and the definition of the space–time white noise requires some care. Recent advances have been made in Bakka et al. (2020) and Carrizo Vergara et al. (2021), where space–time models admitting Matérn covariance functions in space have been proposed.

Carrizo Vergara et al. (2021) proposed a rigorous mathematical framework, based on the theory of Generalised Random Fields (Itô, 1954; Gelfand and Shilov, 1964), to construct spatio-temporal random fields arising from a very broad class of linear SPDEs. They considered the general class of linear SPDEs of the form $\mathcal{L}_g U = X$ defined through operators of the form

$$\mathcal{L}_g(\cdot) := \mathcal{F}^{-1}(g\mathcal{F}(\cdot)), \tag{2}$$

where, in the real case, $g : \mathbb{R}^d \to \mathbb{R}$ is a continuous, symmetric function bounded by a polynomial, $\mathcal{F}$ the Fourier transform on $\mathbb{R}^d$ and $\mathcal{F}^{-1}$ the inverse Fourier transform. The function g in (2) is referred to as the *symbol function* of the operator. If, in addition, $|g|$ is bounded from below by the inverse of a strictly positive polynomial, there exists a unique Generalised Random Field solution to the equation $\mathcal{L}_g U \overset{\text{2nd.o.}}{=} X$, where '2nd. o.' means that the solution has to be understood in the second order sense (see Carrizo Vergara et al., 2021, for more details). In this case, the spectral measures of X and U are connected by:

$$d\mu_X = |g|^2 d\mu_U^X,$$

which leads to the covariance $\rho_U = \mathcal{F}(|g|^{-2}\, d\mu_X)$. Thus, the relationship between the covariances of X and U is

$$\rho_X = \mathcal{L}_{|g|^2}\rho_U.$$

When X reduces to a white noise W, its spectral measure is $d\mu_X(\xi) = (2\pi)^{-d/2}\, d\xi$, and the covariance ρ_X reduces to the Dirac δ. Hence, the covariance ρ_U can be seen as a Green's Function of the operator $\mathcal{L}_{|g|^2}$. This construction offers the possibility to build and characterise models far beyond the Matérn family which is currently the covariance model considered within most SPDE implementations.

In this contribution, our aim is to further explore the intimate interplay between partial differential equations, their stochastic extensions, namely the SPDEs, and spatial statistics, by adopting a mechanistic perspective based on the movement of microscopic particles. Most of the literature about PDEs focused on elliptic and parabolic operators involving a Laplacian or its anisotropic extension. Although the fractional Laplacian has recently generated much attention in the PDE community, its microscopic interpretation is much less classical than that of the standard Laplacian. As a preliminary to this study, we propose in Section 2 an introduction to the connection between Lévy flights and PDEs involving the fractional Laplacian operator. The corresponding Fokker–Planck PDEs will serve as a basis to revisit in Section 3 the original SPDE framework associated with Eq. (1), but with a mechanistic perspective. We propose new generalisations by considering a general form of SPDE with terms accounting for dispersal, drift and reaction. Regarding the dispersal term, we first show that there is an equivalent integral representation offering a mechanistic interpretation of the operator $-L^{\alpha/2} := -(\kappa^2 - \Delta)^{\alpha/2}$ in the left-hand-side of Eq. (1). As a side result, we show the profound difference between the fractional Laplacian $-(-\Delta)^{\alpha/2}$, which is associated with a fat-tailed dispersal kernel (for $\alpha \in (0, 2)$) and therefore describes long-distance dependencies, and the damped fractional Laplacian $-L^{\alpha/2}$, which is associated with a thin-tailed kernel, corresponding to short-distance dependencies. We then introduce the fractional Laplacian operator with a linear reaction term $-(-\Delta)^{\alpha/2} - \kappa^2$ and we also consider kernel-based dispersal operators. In each case, we derive the associated covariance function and we compare its properties with those of the Matérn family. We then illustrate SPDEs with non-stationary drift terms representing external spatially and temporally varying forces. Finally, we introduce nonlinear bistable reaction terms; we discuss their physical meaning in a deterministic framework and the possible applications of the corresponding SPDE, the stochastic Allen–Cahn equation. Lastly, in Section 4, we revisit the classical use of statistical models for parameter estimation in PDE models, with which we started this introduction. In this original approach, we combine the Fokker–Planck PDEs of Section 2 with inhomogeneous point processes. Section 5 concludes with some elements of discussion, with a particular focus on the form of the diffusion operator.

In the following, $U(x, t)$ will represent a random field defined over space and time, with $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $t \in \mathbb{R}^+$. Partial derivatives along time will be denoted ∂_t and $\Delta = \sum_{i=1}^{d} \partial^2/\partial x_i^2$ is the Laplacian. The Fourier transform is defined, as follows: $\mathcal{F}(f)(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-i\xi\cdot x} f(x)dx$. As already mentioned in the first paragraph, we will often adopt a microscopic point of view, in which solutions of PDEs or SPDEs describe the density of particles subject to diffusion equations.

2. Lévy flights and PDEs with fractional Laplacian

Since Einstein's theory of Brownian motion (Einstein, 1905), the parabolic PDE $\partial_t u(x, t) = D\Delta u(x, t)$ has a well-established physical interpretation. The solution $u(x, t)$ can indeed be seen as the density of particles following independent Brownian motions. More generally, Fokker–Planck equations act as a natural bridge between parabolic PDEs and Itô diffusion stochastic differential equations (SDEs) (e.g., Gardiner, 2009). These SDEs extend Brownian motion by introducing a drift (or similarly a transport) coefficient and spatio-temporal heterogeneities in the movement of particles. The solution of parabolic-like equations but with a fractional Laplacian $-(-\Delta)^{\alpha/2}$ (with $\alpha \in (0, 2)$) instead of a standard Laplacian ($\alpha = 2$) can also be interpreted as a density of particles. But in this case the particles follow independent Lévy flights. As this physical interpretation is

less standard, we propose below a short introduction to the connection between Lévy flights and Fokker–Planck PDEs involving a fractional Laplacian.

We consider a particle whose position $X_t \in \mathbb{R}^d$ follows a d-dimensional α-stable Lévy process, for some $\alpha \in (0, 2]$ (see e.g., Cont and Tankov, 2003; Applebaum, 2009, for precise definitions). This can be described by a Langevin-like equation (e.g. Schertzer et al., 2001; Applebaum, 2009):

$$dX_t = B(t, X_t)\, dt + \Sigma(t, X_t) dL_t, \tag{3}$$

where L is an α-stable rotation-invariant d-dimensional Lévy process with scale parameter γ and dL_t is a forward increment in time of this process. As mentioned above, when $\alpha = 2$, L is a standard Brownian motion and Eq. (3) is a standard Itô diffusion SDE (for a discussion of the differences between Itô and Stratonovitch SDEs, the reader can refer to Gardiner, 2009; Horsthemke and Lefever, 2006). The vector field $B(t, X_t) = (B_1(t, X_t), \ldots, B_d(t, X_t)) \in \mathbb{R}^d$ is the drift coefficient. It is related to the existence of some external force acting on the particles, sometimes also called *bias*, e.g. due to wind or attraction/repulsion. For simplicity, we assume an isotropic and time-independent random driving force (but see Remark 1), which means that the term Σ (the diffusion matrix when $\alpha = 2$) satisfies $\Sigma(t, X_t) = \sigma(X_t) I_d$ where σ is a scalar function and I_d is the identity matrix of size d. Note that, in general, the coefficients B and Σ may depend on the current position X_t. For example, the particle motion could be slower in some regions of space, which would be modelled with a smaller value of σ in these regions. Note that the existence and uniqueness of the solution of (3) are guaranteed under some technical growth conditions on the coefficients (see e.g., Arnold, 1974; Schertzer et al., 2001; Øksendal, 2003).

The Fokker–Planck equation describes the dynamics of the transition probability density $P(X(t) = x|X(0) = x_0)$. We denote by $p(x, t)$ this transition probability. In the standard diffusion case $\alpha = 2$, the dynamics of $p(x, t)$ are described by the parabolic PDE (see e.g. Gardiner, 2009):

$$\partial_t p(x, t) = \frac{1}{2} \Delta(\sigma^2(x)\, p(x, t)) - \operatorname{div}\left(B(x, t)\, p(x, t)\right), \quad t > 0, \ x \in \mathbb{R}^d, \tag{4}$$

where the divergence operator 'div' acts on the vector fields $V \in C^1(\mathbb{R}^d, \mathbb{R}^d)$, $\operatorname{div}(V) := \sum_{i=1}^d \partial V_i / \partial x_i$. When the movement of the particles is not driven by a Gaussian noise (i.e. when $\alpha < 2$), the Fokker–Planck equation involves a fractional Laplacian (Schertzer et al., 2001):

$$\partial_t p(x, t) = -\gamma\, (-\Delta)^{\alpha/2}(\sigma^\alpha(x)\, p(x, t)) - \operatorname{div}\left(B(x, t)\, p(x, t)\right), \quad t > 0, \ x \in \mathbb{R}^d, \tag{5}$$

see Section 3 for more details about this operator; see also Kwaśnick (2017) for alternative definitions of this operator. Notice that, with $\alpha = 2$ and $\gamma = 1/2$ (Brownian diffusion), Eq. (5) is equivalent to (4).

Remark 1. In the general anisotropic case, when $\alpha = 2$ and $\Sigma(t, X_t)$ is not necessarily of the form $\sigma(X_t) I_d$, we obtain an equation similar to (4), but the term $\frac{1}{2}\Delta(\sigma^2(x)\, p(x, t))$ must be replaced by $\sum_{i,j=1}^d \partial^2 / \partial_{x_i} \partial_{x_j} (D_{ij}(x, t) p(x, t))$ with D_{ij} the entries of the matrix $\mathbf{D}(x, t) := \Sigma\, \Sigma^{\mathrm{T}}/2$. When $\mathbf{D}$ does not depend on x, we note that this operator reduces to $\operatorname{div}(\mathbf{D}(t) \nabla p(x, t))$. Otherwise, it can be written as a sum of a 'Fickian' diffusion term (see also the discussion in Section 5) and a drift term: $\sum_{i,j=1}^N \partial^2 / \partial_{x_i} \partial_{x_j} (D_{ij}(x, t) p(x, t)) = \operatorname{div}(\mathbf{D}(x, t) \nabla p(x, t)) + \operatorname{div}[\operatorname{div}(\mathbf{D}(x, t)) p(x, t)]$, with $\operatorname{div}(\mathbf{D}(x, t))$ the vector with coordinates $\sum_{i=1}^N \partial / \partial_{x_i} D_{ij}(x, t)$.

The transition probability density $P(X(t) = x|X(0) = x_0)$ involves an initial condition $\delta(x - x_0)$, where δ is a Dirac measure at 0. Assuming N_0 independent particles, with common initial distribution p_0, it is then straightforward to check that the expected density of particles $u(x, t)$ satisfies the PDE (5), with the initial condition $u_0(x) := u(x, 0) = N_0\, p(x, 0)$.

Assume further that the particles have a life expectancy $1/\kappa^2$ for some $\kappa > 0$ and disappear at an exponential rate. Consider any given particle. We denote by τ its "death" time (or removal time: we do not only focus on living organisms). With a slight abuse of language, we define $h(x, t)$ as the probability that the particle is at the position x and still dispersing at time t, that is, for any measurable set $\Omega \subset \mathbb{R}^d$,

$$\int_\Omega h(x, t) dx := P(\{X_t \in \Omega\} \cap \{t < \tau\}).$$

Since dispersal and death are defined as two independent processes, $h(x,t) = p(x,t)P(t < \tau) = p(x,t)\,e^{-\kappa^2 t}$. Therefore, $h(x,t)$ and consequently the expected particle density $u(x,t)$ satisfy the following equation:

$$\partial_t u(x,t) = -\gamma\,(-\Delta)^{\alpha/2}(\sigma^\alpha(x)\,u(x,t)) - \operatorname{div}(B(x,t)\,u(x,t)) - \kappa^2\,u(x,t),\ \ t>0,\ \ x\in\mathbb{R}^d. \tag{6}$$

This construction highlights two interesting connections between PDEs and stochastic processes. On the one hand, it illustrates that some PDEs with fractional Laplacian arise quite naturally as the Fokker–Planck equation of SDEs with Lévy flights. On the other hand, the PDE in (6) suggests interesting generalisations of the SPDE in (1) which will be the subject of the next section.

3. SPDEs and random fields: a mechanistic point of view

Inspired by (6), we consider a quite general class of time-dependent SPDE models that rely on mechanistic assumptions concerning the underlying variable. These models take the following general form:

$$\partial_t U(x,t) = \underbrace{\mathcal{D}(x,t,[U])}_{\text{dispersal}} + \underbrace{\mathcal{B}(x,t,[U])}_{\text{drift}} + \underbrace{f(x,t,U)}_{\text{reaction}} + \underbrace{\sigma\,W_t(x)}_{\text{noise}},\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+. \tag{7}$$

The random field $U(x,t)$ is thus governed by a dispersal term $\mathcal{D}$, a drift term $\mathcal{B}$, a reaction term f describing the local feedback of the random field U on itself and a spatio-temporal white noise $\sigma\,W_t$ defined as the derivative of a Wiener process, and thus verifying $E[W_t(x)W_{t'}(x')] = \delta(t-t')\delta(x-x')$. The square bracket symbol $[U]$ is used to underline that the operator may depend on the whole function $U(\cdot,t)$, and not only on its value at x. The well-posedness of (7) depends on its precise form and on the space dimension d. When the reaction term is linear and with coefficients that are constant in (x,t), the existence and uniqueness of a stationary distribution-valued solution generally follow from the results in Carrizo Vergara et al. (2021), at least in the examples that we treat below. With nonlinear reaction terms $f(x,t,U)$ (nonlinear with respect to U), the theory is also well-established in the case $d=1$ (e.g. Da Prato and Zabczyk, 2014). The picture is less clear when $f(x,t,U)$ is nonlinear and $d\geq 2$: although such equations are regularly used in applied sciences, mathematical results tend to show that additive white noise leads to ill-posed equations (Hairer et al., 2012; Ryser et al., 2012).

In the particular case $\mathcal{D}(x,t,[U]) = D\,\Delta U(x,t)$, $\mathcal{B}(x,t,[U]) = 0$ (no drift) and $f(x,t,u) = -\kappa^2\,U(x,t)$ for some $D,\kappa>0$ the solution $U(x,t)$ of (7) satisfies the "standard" evolving Matérn equation, already discussed in Lindgren et al. (2011) and Carrizo Vergara et al. (2021):

$$\partial_t U(x,t) = \underbrace{D\Delta U(x,t)}_{\text{dispersal}} \underbrace{-\kappa^2\,U(x,t)}_{\text{reaction}} + \underbrace{\sigma\,W_t(x)}_{\text{noise}},\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+, \tag{$\mathcal{S}$}$$

In this equation, there is a diffusive dispersal term and an absorption term (the reaction) which forces the solution to remain close to 0, see Section 2 for a microscopic interpretation. When the noise does not depend on time (spatial white noise W_S), the SPDE (1) in Lindgren et al. (2011) is a time-independent solution of ($\mathcal{S}$) for $\alpha = 2$. In the general case $\alpha\neq 2$ the SPDE (1) involves the "damped" fractional operator $-L^{\alpha/2} := -(\kappa^2-\Delta)^{\alpha/2}$, which does not exactly fit in (7).

We will discuss below three extensions of the "standard" SPDE-GMRF that go beyond the framework ($\mathcal{S}$) set in Lindgren et al. (2011) and their possible applications in environmental and ecological sciences. These extensions either deal with the dispersal term (Section 3.1), the drift term (Section 3.2) or the reaction term (Section 3.3).

3.1. Dispersal operators

In equations of the form (7), nonlocal dispersal operators

$$\mathcal{D}(x,t,[U]) = D\,(\mathcal{J}\star U - U)(x,t) := D\int_{\mathbb{R}^d}\mathcal{J}(x-y)(U(y,t)-U(x,t))\,dy, \tag{8}$$

Table 1
Summary of the main models surveyed in Section 3.1. The (generalised) covariance functions correspond to the time-independent case ($\partial_t U = 0$, spatial white noise W_S).

Operator	Dispersal kernel $\mathcal{J}(\|x\|)$	Absorption term	Spatial symbol $g(\xi)$	Covariance function
$-L^{\alpha/2}U$	$\frac{K_{(\alpha+d)/2}(\kappa\,\|x\|)}{\|x\|^{(\alpha+d)/2}}$ $\underset{+\infty}{\sim} c\,\frac{e^{-\kappa\,\|x\|}}{\|x\|^{(1+\alpha+d)/2}}$ $\alpha \in (0,2)$	$-h_\kappa\, U$ (cf (15))	$(\|\xi\|^2+\kappa^2)^{\alpha/2}$	Matérn (12) ($\alpha \geq d/2$)
$-M^{\alpha/2}U$	$c_{d,\alpha}/\|x\|^{d+\alpha}$	$-\kappa^2\, U$	$\|\xi\|^\alpha+\kappa^2$	$d=1, \alpha=1$: (17) $\underset{+\infty}{\sim} \frac{2}{\kappa^6\,\pi\,\|h\|^2}$ $d=2, \alpha=1$: (19) $\underset{+\infty}{\sim} \frac{1}{\kappa^6\,\pi\,\|h\|^3}$ $\kappa=0$, $\alpha<d/2$: $\propto \frac{1}{\|\|h\|\|^{d-2\alpha}}$
$D(\mathcal{J}\star U-U)-\kappa^2 U$	$\mathcal{J}(\|x\|)$	$-\kappa^2\, U$	$D(1-(2\pi)^{d/2}\mathcal{F}(\mathcal{J})(\xi))+\kappa^2$	$\mathcal{J}(x)=e^{-\|x\|/\beta}$, $d=1$: "Nugget"+3/2-Matérn

can describe short or long-distance dispersal, depending on the precise shape of the kernel $\mathcal{J}$ (Klein et al., 2006; Garnier, 2011). When imposing that the integral of $\mathcal{J}$ is equal to 1, the quantity $\mathcal{J}(x-y)$ can be interpreted as the probability density that a dispersing particle initially located at y moves to the position x (see e.g., Roques, 2013), and the coefficient $D > 0$ is the rate of dispersal. The frequency of long-distance dispersal events is governed by the tail of the kernel as $\|x\| \to \infty$: thin-tailed kernels decay at least exponentially and fat-tailed kernels decay more slowly than any exponential function.

Equations of the form (7) with dispersal terms of the form (8) and regular kernels $\mathcal{J}$ of mass 1, null drift term $\mathcal{B}(x,t,[U]) = 0$ and $f(x,t,U) = -\kappa^2\, U(x,t)$ fit in the class of SPDEs considered in Carrizo Vergara et al. (2021), as $-D(\mathcal{J}\star U-U)+\kappa^2\, U = \mathcal{F}^{-1}(g\,\mathcal{F}(U))$, with symbol function $g = D(1-(2\pi)^{d/2}\mathcal{F}(\mathcal{J}))+\kappa^2$, which is a Hermitian-symmetric function (see Theorem 1 in Carrizo Vergara et al., 2021). We also consider singular kernels for which, even if $\mathcal{J}\star u$ is not properly defined, the right-hand side in (8) is still well-defined, for instance when $\mathcal{D}(x,t,[U]) = -(-\Delta)^{\alpha/2}U(x,t)$. We compare here, from a mechanistic viewpoint, the physical interpretation of these equations

$$\partial_t U(x,t) = D\,(\mathcal{J}\star U - U)(x,t) - \kappa^2\, U(x,t) + \sigma\, W_t(x), \tag{9}$$

compared to evolving Matérn equations

$$\partial_t U(x,t) = -L^{\alpha/2}U(x,t) + \sigma\, W_t. \tag{10}$$

The main findings of this section are summarised in Table 1.

3.1.1. Fractional Laplacian

We begin by recalling that the fractional Laplacian $-(-\Delta)^{\alpha/2}$ has a well-known pointwise representation when $\alpha \in (0,2)$ (e.g., Kwaśnick, 2017): for all $v \in \mathcal{L}^p(\mathbb{R}^d)$ ($p \in [1,\infty)$),

$$-(-\Delta)^{\alpha/2}v(x) = c_{d,\alpha}\int_{\mathbb{R}^d}\frac{v(y)-v(x)}{\|x-y\|^{d+\alpha}}\,dy, \tag{11}$$

for some constant $c_{d,\alpha} > 0$. Due to a singularity around 0, the above integral has to be understood in the Cauchy principal value sense when $\alpha \in [1,2)$. Eq. (11) shows that the fractional Laplacian is a nonlocal operator, and formally corresponds to a dispersal term of the form (8), although the

kernel $\mathcal{J}(x) = c_{d,\alpha}/\|x\|^{d+\alpha}$ is not integrable around 0. The kernel $\mathcal{J}$ belongs here to the family of (very) fat-tailed kernels, which is consistent with the microscopic interpretation of this operator as seen in Section 2, since Lévy flights (3) are continuous time random walks, with possible large jumps if $\alpha < 2$.

3.1.2. Fractional damped Laplacian

We recall that the solution to $-L^{\alpha/2}U(x) = W_S$ in dimension $d \geq 1$ is a Gaussian random field with Matérn covariance of parameter $\nu = \alpha - d/2 > 0$:

$$\rho_U^{W_S}(h) = \frac{1}{(2\pi)^{\frac{d}{2}} 2^{\alpha-1}\kappa^{2\alpha-d}\Gamma(\alpha)}(\kappa\,\|h\|)^{\alpha-\frac{d}{2}} K_{\alpha-\frac{d}{2}}(\kappa\,\|h\|),\ \ h \in \mathbb{R}^d, \tag{12}$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind with index $\nu > 0$. Notice that in the limit case $\nu = 0$, $\rho_U^{W_S}(h)$ is a generalised Matérn covariance which is well defined for $\|h\| > 0$.

In order to find kernel $\mathcal{J}$ associated to $-L^{\alpha/2}$, we use a semigroup formula (e.g., Stinga, 2019) (for $\alpha \in (0,2)$ and $v \in \mathcal{L}^p(\mathbb{R}^d)$, $p \in [1,\infty)$):

$$-L^{\alpha/2}v = \frac{1}{|\Gamma(-\alpha/2)|}\int_0^\infty \left(e^{-tL}v - v\right)\frac{1}{t^{1+\alpha/2}}\,dt, \tag{13}$$

with Γ the gamma function. The quantity $u(\cdot,x) = e^{-tL}v$ is the semigroup generated by L acting on v, i.e., the solution of the Cauchy problem:

$$\partial_t u = -L\,u = \Delta\,u - \kappa^2\,u, \quad t > 0,\ x \in \mathbb{R}^d,\ \text{ with } u(\cdot,0) = v.$$

More explicitly, $u(x,t) = e^{-\kappa^2 t}(G_t \star v)(x)$, with $G_t(x) = (4\pi t)^{-n/2}\,\exp(-\|x\|^2/(4\,t))$, the heat kernel. Replacing into (13), we have:

$$-L^{\alpha/2}v(x) = \frac{1}{|\Gamma(-\alpha/2)|}\int_0^\infty \left(\int_{\mathbb{R}^d} e^{-\kappa^2 t}G_t(x-y)v(y)\,dy - v(x)\right)\frac{1}{t^{1+\alpha/2}}\,dt.$$

Thus,

$$-L^{\alpha/2}v(x) = \frac{1}{|\Gamma(-\alpha/2)|}\int_0^\infty\int_{\mathbb{R}^d} e^{-\kappa^2 t}G_t(x-y)(v(y)-v(x))\frac{1}{t^{1+\alpha/2}}\,dy\,dt - v(x)\,h_\kappa, \tag{14}$$

with

$$h_\kappa := \frac{1}{|\Gamma(-\alpha/2)|}\int_0^\infty \left(1 - e^{-\kappa^2 t}\right)\frac{1}{t^{1+\alpha/2}}\,dt \in (0,\infty) \text{ for } \alpha \in (0,2). \tag{15}$$

Then, we note that:

$$\begin{aligned}\int_0^\infty e^{-\kappa^2 t}G_t(x-y)\frac{1}{t^{1+\alpha/2}}\,dt &= \frac{2\,(2\kappa)^{(\alpha+d)/2}}{(4\pi)^{d/2}}\frac{K_{(\alpha+d)/2}(\kappa\,\|x-y\|)}{\|x-y\|^{(\alpha+d)/2}}\\ &= \frac{2\,(2\kappa)^{(\alpha+d)/2}}{(4\pi)^{d/2}}\mathcal{J}_{(\alpha+d)/2,\kappa}(\|x-y\|),\end{aligned}$$

with

$$\mathcal{J}_{(\alpha+d)/2,\kappa}(s) := \frac{K_{(\alpha+d)/2}(\kappa\,|s|)}{|s|^{(\alpha+d)/2}},\ \ s \in \mathbb{R}^*.$$

Finally, using Fubini's theorem and assuming the convergence of the double integral in (14), we formally get the following pointwise expression for the operator $-L^{\alpha/2}$, in the Cauchy principal value sense:

$$-L^{\alpha/2}v(x) = \tilde{c}_{d,\alpha,\kappa}\int_{\mathbb{R}^d}\mathcal{J}_{(\alpha+d)/2,\kappa}(x-y)(v(y)-v(x))\,dy - v(x)\,h_\kappa,$$

with $\tilde{c}_{d,\alpha,\kappa} = 2\,(2\kappa)^{(\alpha+d)/2}/[(4\pi)^{d/2}\,|\Gamma(-\alpha/2)|]$.

This formula leads to a mechanistic interpretation of the operator $-L^{\alpha/2}$: it contains an absorption term (the reaction term $-v\,h_\kappa$) which becomes stronger as κ is increased, and a nonlocal dispersal term of the form (8). The kernel $\mathcal{J}_{(\alpha+d)/2,\kappa}(s)$ behaves as $s^{-(1+\alpha+d)/2}e^{-\kappa s}$ for large s: it is therefore a thin-tailed kernel, and its tail becomes thinner as κ is increased. Contrarily to the fractional Laplacian (with $\alpha \in (0,2)$), and similarly to the standard Laplacian, the operator $-L^{\alpha/2}$ therefore describes short-distance dispersal events. If we come back to the definition of the Matérn covariance (12), we note that $\rho_U^{W_S}(h)$ has an exponential decay (like $e^{-\kappa\|h\|}\,\|h\|^{\alpha-d/2-1/2}$) as $\|h\| \to \infty$, corresponding to short-distance dependencies.

3.1.3. Fractional Laplacian with linear reaction

The above analysis shows that the evolving Matérn equation. (10) can be written in the form (9), with $D = \tilde{c}_{d,\alpha,\kappa}$, $\mathcal{J} = \mathcal{J}_{(\alpha+d)/2,\kappa}$ and κ^2 replaced by h_κ. This emphasises the interest of using equations of the general form (9) with kernel operators of the form (8) as they should lead to additional flexibility, compared to evolving Matérn equations. We consider here Eq. (9) with $D(\mathcal{J}\star U - U)(x,t) = -(-\Delta)^{\alpha/2}U(x,t)$, which amounts to considering the fractional Laplacian operator with linear reaction term:

$$-M^{\alpha/2}U := -(-\Delta)^{\alpha/2}U - \kappa^2 U. \tag{16}$$

In this case, the (spatial) symbol function g in (2) is $g(\xi) = \|\xi\|^\alpha + \kappa^2$, to be compared with $g(\xi) = (\|\xi\|^2+\kappa^2)^{\alpha/2}$ for the damped Laplacian. The spectral measure is (see Theorem 1 in Carrizo Vergara et al., 2021):

$$d\mu_U^W(\xi,\omega) = \frac{1}{(2\pi)^{\frac{d+1}{2}}}\frac{d\xi\,d\omega}{\omega^2+(\kappa^2+\|\xi\|^\alpha)^2} \quad \text{and} \quad d\mu_U^{W_S}(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}}\frac{d\xi}{(\kappa^2+\|\xi\|^\alpha)^2},$$

in the time-dependent and time-independent cases (when U does not depend on t, and with a spatial white noise W_S), respectively. In the time-independent case, some explicit expressions for the stationary covariance function can be obtained. First, with $\alpha = 1$, and in dimension $d = 1$, we apply a Fourier transform to obtain the stationary covariance function:

$$\rho_U^{W_S}(h) = \frac{1}{\kappa^2\pi}\left[1-\kappa^2|h|f(\kappa^2|h|)\right],\ h\in\mathbb{R}, \tag{17}$$

with $f(z) = \int_0^\infty e^{-z\,t}/(t^2+1)\,dt$. We note that f is the so-called "first auxiliary function" associated with trigonometric integrals, which can also be written $f(z) = \mathrm{Ci}(z)\sin(z) + (\pi/2 - \mathrm{Si}(z))\cos(z)$ (Chapter 5 in Abramowitz and Stegun, 1964), with $\mathrm{Ci}(z)$ and $\mathrm{Si}(z)$ the Cosine and Sine integrals. Using the asymptotic expansions $\mathrm{Ci}(z) = \sin(z)/z - \cos(z)/z^2 - 2\sin(z)/z^3 + o(1/z^3)$ and $\pi/2 - \mathrm{Si}(z) = \cos(z)/z + \sin(z)/z^2 - 2\cos(z)/z^3 + o(1/z^3)$, we get $\rho_U^{W_S}(h) \sim 2/(\kappa^6\pi\,|h|^2)$ as $|h| \to \infty$. With the same parameter values ($d=1$, $\alpha=1$), recall that the damped Laplacian leads to the 1/2-Matérn covariance function, which decays exponentially for large $|h|$.

In dimension $d = 2$, still with $\alpha = 1$, we use the relationship between the Fourier transform of a radially symmetric function and the Hankel transform to compute the stationary covariance:

$$\begin{aligned}\rho_U^{W_S}(h) &= \frac{1}{\|h\|^{\frac{d}{2}-1}}\int_0^\infty r^{\frac{d}{2}-1}d\mu_U^{W_S}(r)J_{\frac{d}{2}-1}(\|h\|\,r)\,r\,dr = \frac{1}{2\pi}\int_0^\infty \frac{1}{(\kappa^2+r)^2}J_0(\|h\|\,r)\,r\,dr\\ &= \frac{1}{2\pi}\int_0^\infty \frac{\partial}{\partial(\kappa^2)}\left[\frac{-1}{(\kappa^2+r)}\right]J_0(\|h\|\,r)\,r\,dr = \frac{1}{2\pi}\frac{\partial}{\partial(\kappa^2)}\left[\int_0^\infty \frac{-1}{(\kappa^2+r)}J_0(\|h\|\,r)\,r\,dr\right],\end{aligned} \tag{18}$$

with J_ν the Bessel function of the first kind. The last integral in the above equality is a tabulated integral, and corresponds to the Hankel transform of order 0 of $1/(\kappa^2+r)$: from Prudnikov et al. (case 2.12.3.6 in 1986) we know that

$$\int_0^\infty \frac{1}{(\kappa^2+r)}J_0(\|h\|\,r)\,r\,dr = \frac{1}{\|h\|} - \frac{\pi\,\kappa^2}{2}\left[(H_0-Y_0)(\kappa^2\,\|h\|)\right],$$

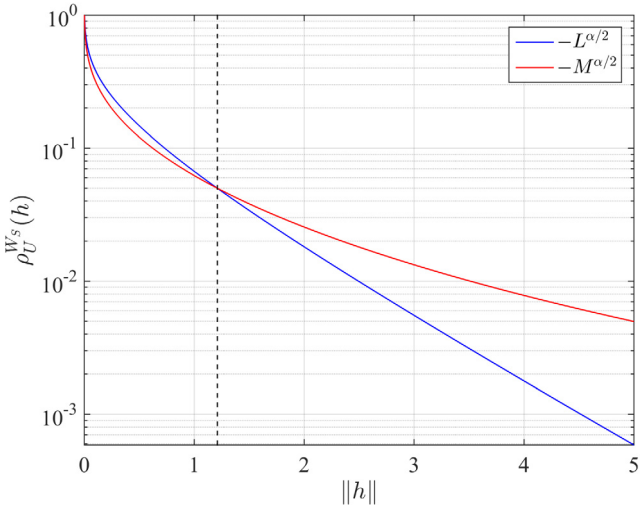


Fig. 1. Covariance functions: damped Laplacian $-L^{\alpha/2}$ vs fractional Laplacian with linear reaction $-M^{\alpha/2}$. The blue line corresponds to the generalised Matérn covariance function associated with the operator $-L^{\alpha/2}$, and given by (12). The red line corresponds to the covariance function associated with the operator (16) and given by (19). In both cases, we consider the time-independent problems in dimension $d = 2$ and with $\alpha = 1$. We set $\kappa = 1$ in the operator $-L^{\alpha/2}$, and $\kappa = 0.78$ in the operator $-M^{\alpha/2}$ so that the two covariance functions have same practical range, i.e. they intersect each other when they are equal to 0.05 (black dashed line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with $Y_\nu(\cdot)$ the Bessel function of the second kind and $H_\nu(\cdot)$ the Struve function of order ν. Differentiating this expression with respect to κ^2 and using (18), we get:

$$\rho_U^{W_S}(h) = \frac{\kappa^2 \|h\|}{4}\left[(Y_1 - H_1)(\kappa^2\|h\|) + \frac{2}{\pi}\right] - \frac{1}{4}\left[(Y_0 - H_0)(\kappa^2\|h\|)\right], \ h \in \mathbb{R}^2. \tag{19}$$

Here, $\rho_U^{W_S}(h) \sim 1/(\kappa^6 \pi \|h\|^3)$ as $\|h\| \to \infty$. Note that, with $d = 2$ and $\alpha = 1$, we have $\alpha = d/2$. In this case, the damped Laplacian does not lead to a Matérn covariance, but to a generalised covariance function proportional to $K_0(\kappa\|h\|) \sim e^{-\|h\|}/\sqrt{\|h\|}$ as $\|h\| \to \infty$ (Carrizo Vergara et al., 2021; Allard et al., 2021).

In both cases $d = 1$ and $d = 2$, we have thus observed that the covariance function $\rho_U^{W_S}(h)$ associated with the fractional Laplacian with absorption has an algebraic decay as $\|h\| \to \infty$. As expected from the mechanistic interpretation of the operator, the obtained covariance function therefore describes long-distance dependencies. See Fig. 1 for an illustration of the behaviour of the new covariance function (19), compared to the generalised covariance given by the damped fractional Laplacian in the case $d = 2$ and $\alpha = 1$.

3.1.4. Convolution kernels

We now move away from pseudo-differential operators and consider Eq. (9) with regular convolution kernels with mass 1, for which $\mathcal{J} \star u$ is well-defined in (8). The spectral measures are (Theorem 1 in Carrizo Vergara et al., 2021):

$$d\mu_U^W(\xi, \omega) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} \frac{d\xi \, d\omega}{\omega^2 + (D(1 - (2\pi)^{d/2}\mathcal{F}(\mathcal{J})(\xi)) + \kappa^2)^2},$$
$$d\mu_U^{W_S}(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{d\xi}{(D(1 - (2\pi)^{d/2}\mathcal{F}(\mathcal{J})(\xi)) + \kappa^2)^2},$$

in the time-dependent and time-independent cases, respectively. As $\mathcal{F}(\mathcal{J})(\xi) \to 0$ as $\|\xi\| \to \infty$, these measures are not finite, see Remark 1 in Carrizo Vergara et al. (2021). In the time-independent case, with an exponential kernel $\mathcal{J}(x) = e^{-\|x\|/\beta} / \left[(2\beta)^d \pi^{\frac{d-1}{2}} \Gamma(\frac{d+1}{2})\right]$, we compute the Fourier transform of $\mathcal{J}$ thanks to the Hankel transform:

$$\begin{aligned}(2\pi)^{d/2}\,\mathcal{F}(\mathcal{J})(\xi) &= \frac{(2\pi)^{d/2}}{\|\xi\|^{\frac{d}{2}-1}} \int_0^\infty r^{\frac{d}{2}-1} \mathcal{J}(r) J_{\frac{d}{2}-1}(\|\xi\| r)\, r\, dr \\ &= \frac{1}{(1+\beta^2\|\xi\|^2)^{\frac{d+1}{2}}},\end{aligned} \tag{20}$$

where the last equality follows from Prudnikov et al. (1986, case 2.8.12.4 with $\nu = d/2 - 1$ and $\alpha = \nu + 2$). In the case $d = 1$,

$$\begin{aligned}d\mu_U^{W_S}(\xi) =& \frac{1}{(2\pi)^{\frac{1}{2}}} \left(\frac{1+\beta^2\xi^2}{(\kappa^2+D)\beta^2\xi^2+\kappa^2}\right)^2 d\xi \\ =& \frac{1}{(2\pi)^{\frac{1}{2}}} \frac{d\xi}{(\kappa^2+D)^2} + \frac{2D}{\beta^2(\kappa^2+D)^3} \frac{1}{(2\pi)^{\frac{1}{2}}} \frac{d\xi}{\xi^2+\kappa^2/[\beta^2(\kappa^2+D)]} \\ &+ \frac{D^2}{\beta^4(\kappa^2+D)^4} \frac{1}{(2\pi)^{\frac{1}{2}}} \frac{d\xi}{(\xi^2+\kappa^2/[\beta^2(\kappa^2+D)])^2}.\end{aligned}$$

Since $\mathcal{F}[1/((2\pi)^{\frac{1}{2}}(\xi^2+a^2))](h) = e^{-a|h|}/(2a)$ and $\mathcal{F}[1/((2\pi)^{\frac{1}{2}}(\xi^2+a^2))^2](h) = (1+a|h|)e^{-a|h|}/(4a^3)$ (for $a > 0$), applying a Fourier transform, we obtain an expression for the stationary covariance function:

$$\rho_U^{W_S}(h) = \frac{1}{(\kappa^2+D)^2}\delta(h) + (c_{\kappa,D}|h| + \tilde{c}_{\kappa,D}) e^{-\frac{\kappa|h|}{\beta\sqrt{\kappa^2+D}}}, \quad h \in \mathbb{R}$$

for some positive constants $c_{\kappa,D}$, $\tilde{c}_{\kappa,D}$. In this case, we observe that the covariance function is comparable to a 3/2-Matérn covariance with an additional Dirac mass at 0, referred to as the "nugget effect" in geostatistics.

Let us now consider the more general case where the Fourier transform of $\mathcal{J}$ is $(2\pi)^{d/2}\,\mathcal{F}(\mathcal{J})(\xi) = (1+\beta^2\|\xi\|^2)^{-m}$, with $m \in \mathbb{N}^*$. This corresponds to an exponential kernel with an odd value of the dimension $d = 2m - 1$, see Eq. (20) or more generally to a dispersal kernel given by a Matérn covariance in $\mathbb{R}^d$ with parameter $\nu = m$. Then,

$$\begin{aligned}d\mu_U^{W_S}(\xi) =& \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{(1+\beta^2\|\xi\|^2)^{2m}}{\left[(\kappa^2+D)(1+\beta^2\|\xi\|^2)^m - D\right]^2} d\xi \\ =& \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{d\xi}{(\kappa^2+D)^2} + \frac{2D}{(\kappa^2+D)^2} \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{d\xi}{(1+\beta^2\|\xi\|^2)^m(\kappa^2+D) - D} \\ &+ \frac{D^2}{(\kappa^2+D)^2} \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{d\xi}{[(1+\beta^2\|\xi\|^2)^m(\kappa^2+D) - D]^2}.\end{aligned}$$

The spectral measure associated to a general Matérn dispersal kernel is thus the sum of a Lebesgue measure and two spectral densities which are the inverse of positive polynomials of $\|\xi\|^2$. Thanks to the Rozanov Theorem (Rozanov, 1977), it is known that the Gaussian random fields on $\mathbb{R}^d$ characterised by such spectral densities are Gaussian Markov random fields, and thus with thin-tailed covariances.

As the kernel $\mathcal{J}$ were thin-tailed here, the exponential decay of the covariance functions at infinity is not surprising. We expect a different behaviour for fat-tailed kernels (see Table 1 in Klein et al., 2006), leading to long-range dependencies as those observed with the singular kernel $\mathcal{J}(x) = c_{d,\alpha}/\|x\|^{d+\alpha}$ corresponding to the fractional Laplacian.

3.2. Spatially and temporally varying drift terms

The "heat equation" $\partial_t U(x,t) = D\,\Delta\,U(x,t)$ describes the probability density of a standard Brownian motion. Itô diffusion generalises Brownian motion by adding a drift coefficient $B(t,X_t) \in \mathbb{R}^d$ (see Section 2) which takes into account the presence of some external force, or bias. In environmental sciences, such drift may arise from multiple reasons, the effect of the wind on particle movements being one obvious example. The associated density follows an advection–diffusion equation with a drift term (Fokker–Planck equation; see e.g. Gardiner, 2009):

$$\mathcal{B}(x,t,[U]) = -\mathrm{div}(B(x,t)\,U(x,t)), \tag{21}$$

where $B = (B_1,\ldots,B_d)$ and $\mathrm{div}(B) = \sum_{i=1}^{d} \partial B_i/\partial x_i$. Plugging this drift term into (7), with a standard diffusion term $\mathcal{D}(x,t,[U]) = D\,\Delta U(x,t)$ and an absorption term $f(x,t,U) = -\kappa^2\,U(x,t)$, with $D,\kappa > 0$, we get the stochastic reaction–advection–diffusion equation

$$\partial_t U(x,t) = D\,\Delta U(x,t) - \mathrm{div}(B(x,t)\,U(x,t)) - \kappa^2\,U(x,t) + \sigma W_t(x),\ (x,t) \in \mathbb{R}^d \times \mathbb{R}^+. \tag{22}$$

Lindgren et al. (2011, Section 3.5) and Sigrist et al. (2015) have considered SPDEs of this form (but with a constant anisotropic diffusion matrix $\mathbf{D}$, which means that $D\,\Delta U(x,t)$ is replaced by $\mathrm{div}(\mathbf{D}\nabla U(x,t))$, see Remark 1) and constant drift vector B, with an application to the modelling of rainfall data. It is worth emphasising that the drift term (21) involves the divergence of the product between the varying drift vector $B(x,t)$ and $U(x,t)$, with $\mathrm{div}(B(x,t)\,U(x,t)) = U(x,t)\mathrm{div}(B(x,t)) + B(x,t)\cdot\nabla U(x,t)$. When B is constant, as in the above mentioned works, this reduces to $\mathrm{div}(B\,U(x,t)) = B\cdot\nabla U(x,t)$.

To the best of our knowledge, a drift vector that varies in space and time has never been considered. It is beyond the scope of this paper to mathematically characterise the associated random fields, as the operator does not fit in the class considered in Carrizo Vergara et al. (2021) when the vector field B is not constant. In Fig. 2, we compare the solution of Eq. (22) in a square domain of $\mathbb{R}^2$ at a fixed time T, with the solution of the standard equation $(\mathcal{S})$ without drift term. In both cases, we start from an initial condition $U(x,0) = 0$, and we approach the solution with a finite difference approximation over a bounded domain with Dirichlet boundary conditions, i.e. $U(x,t) = 0$ on the boundary of the domain. We use an explicit Euler scheme in time (Matlab codes are available in the Open Science Framework repository: https://osf.io/w5utd). The drift term is chosen as follows:

$$\begin{cases} B_1(x_1,x_2,t) = 2\,x_2\cos(2\,\pi\,t/10)/\sqrt{0.1+x_1^2+x_2^2}, \\ B_2(x_1,x_2,t) = 2\,x_1\sin(2\,\pi\,t/10)/\sqrt{0.1+x_1^2+x_2^2}. \end{cases} \tag{23}$$

The other parameter values are given in the legend of Fig. 2. The effect of the drift term is more visible in the video file https://osf.io/ydrqk/, to be compared with the dynamics in the absence of drift https://osf.io/krvug/.

3.3. Bistable reaction terms

In the absence of noise, that is when $\sigma = 0$, $u(x,t) = 0$ is the only solution of $(\mathcal{S})$ as $t \to \infty$ and it is asymptotically stable (notice that we use here the lower case u to denote the solution since it is not random). In other words, in the PDE framework, starting from any bounded initial condition, the solution $u(x,t)$ converges towards 0 uniformly in space and time. This easily follows from a comparison argument (Protter and Weinberger, 1967): $\min_{x\in\mathbb{R}^d} u(x,0)\,e^{-\kappa^2 t} \le u(x,t) \le \max_{x\in\mathbb{R}^d} u(x,0)\,e^{-\kappa^2 t}$ for all $t \ge 0$ and $x \in \mathbb{R}^d$. On the other hand, if the reaction term $-\kappa^2\,u(x,t)$ is replaced by a bistable reaction term:

$$f(x,t,u) = u(x,t)(K - u(x,t))(u(x,t) - \rho), \tag{24}$$

with $K > 0$ and $\rho \in (0,K)$, the equation exhibits much more complex dynamics. Ignoring dispersal and drift, the ordinary differential equation $v'(t) = f(v)$, $t > 0$ admits two stable stationary states:

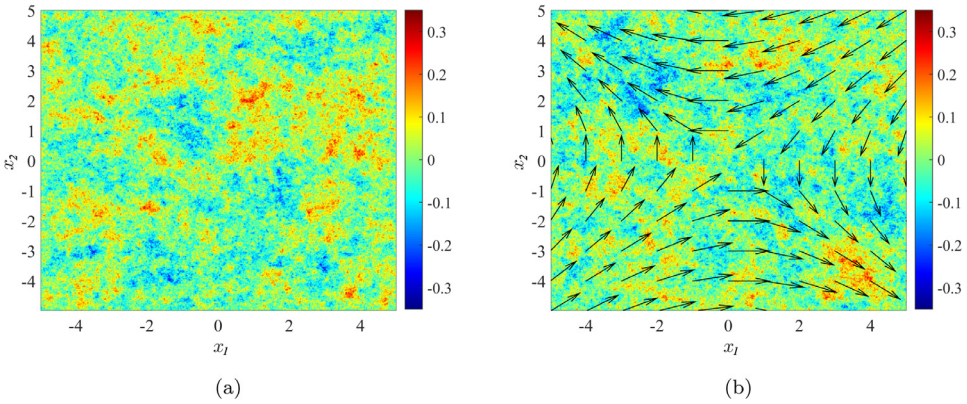


Fig. 2. Effect of a drift term. Solution of the SPDE (22) with null drift coefficient $B(x,t) \equiv 0$ in panel (a), and a drift coefficient $B(x,t)$ given by (23) in panel (b). In both cases, $D = 0.05$, $\kappa = 0.1$, $\sigma = 1$ and the solution is plotted at a fixed time $T = 6$. In panel (b), the arrows correspond to the directional components of the vector field B.

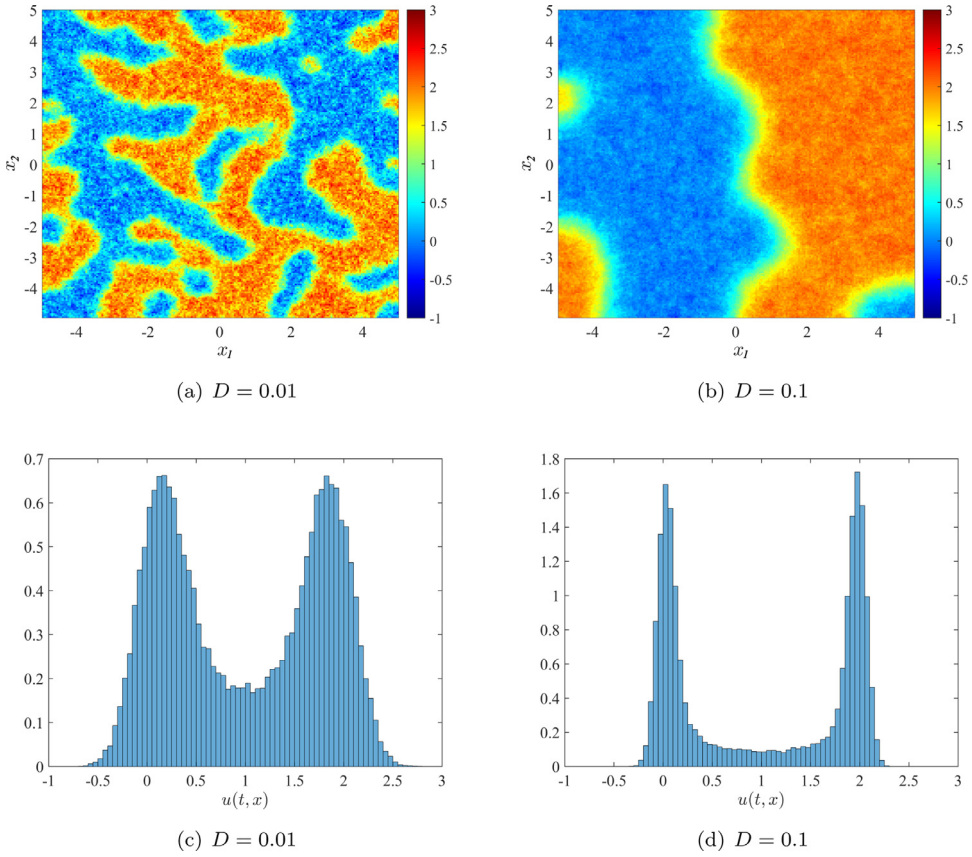
0 and K, and an unstable one: ρ. Thus, starting with an initial condition $v(0) \in (-\infty, \rho)$, $v(t)$ ultimately converges to 0, and if $v_0 \in (\rho, +\infty)$, $v(t)$ converges to K (see e.g. Roques, 2013). In the PDE framework, when $u(x,t)$ also depends explicitly on x, the behaviour of the solution strongly depends on the precise shape of the initial condition $u(\cdot, 0)$ and on the value of ρ. This is the consequence of the interplay between the dispersal and the reaction term. Consider for example $u(x,0) = K\,\chi_\Omega(x)$, where χ_Ω is the characteristic function of some set $\Omega \subset \mathbb{R}^d$. Then the solution tends to converge towards K for large and aggregated sets Ω and tends to converge towards 0 for small and fragmented sets Ω (Zlatoš, 2006; Matano and Du, 2010; Garnier et al., 2012). More generally, the sign of the integral of f over $(0, K)$ determines which state is more stable: 0 if the integral is nonpositive and K if the integral is positive. The "more stable" state tends to invade the other one for a larger class of initial conditions (e.g., Fife and McLeod, 1977). The particular case $\rho = K/2$ corresponds to a fully symmetric situation.

The reaction term (24) is often used in biology to take into account an "Allee effect" (Lewis and Kareiva, 1993; Keitt et al., 2001). In such case $f(x,t,u)$ corresponds to the growth rate of a population with local density $u(x,t) \geq 0$. The local population density must be above the "Allee threshold" ρ to reach a positive growth rate. This means that a form of cooperation between the individuals is required. When the population density is too large ($u(x,t) > K$) the growth rate becomes negative again, meaning that the population density tends to decay due to e.g. resource saturation when it reaches too large values. From a more empirical viewpoint, this reaction term reproduces the bistable behaviour that can be observed in some ecosystems. For instance, it is known that a positive feedback between the vegetation abundance and the local conditions for the growth of other plants leads to bistability and patchiness, especially in arid systems (Rietkerk and van de Koppel, 1997; Kéfi et al., 2007).

Here, we consider Eq. (7) with a diffusive dispersal term $\mathcal{D}(x,t,[U]) = D\,\Delta U$, a null drift term $\mathcal{B}(x,t,[U]) = 0$ and a reaction term (24) with $\rho = K/2$. This corresponds to the stochastic Allen–Cahn equation,

$$\partial_t U(x,t) = D\,\Delta U(x,t) - \kappa^2\,U(x,t) + U(x,t)(K - U(x,t))(U(x,t) - K/2) + \sigma\,W_t(x), \qquad (25)$$

for which existence and uniqueness are only established when the space dimension is $d = 1$. When $d \geq 2$, as in the examples that we consider in this work, it is conjectured that the equation is ill-posed, and it was shown that modelling with this equation is questionable, because the numerical solutions converge to the zero-distribution when the mesh grid is refined (Ryser et al., 2012). Several



(a) $D = 0.01$ (b) $D = 0.1$

(c) $D = 0.01$ (d) $D = 0.1$

Fig. 3. Effect of a bistable reaction term. Panels (a, b): finite difference solution of the SPDE (25). Panels (c, d): corresponding histograms describing the distribution of $U(x, T)$. In all cases, $K = 2$, $\sigma = 1$ and $T = 10$. The mesh size is $\delta x = 0.05$.

studies work around this problem by considering a coloured noise with a finite spatial correlation length λ (e.g. Kohn et al., 2007). In spite of these difficulties, and even though the theory in Carrizo Vergara et al. (2021) does not apply in this case due to the nonlinear nature of f, we foresee that the numerical solution of the SPDE may exhibit interesting properties, especially to model variables with bimodal distributions. As argued by Ryser et al. (2012), although the SPDE (25) may be ill-posed, it is still possible to give an interpretation to its "finite difference solution", with a fixed mesh size δx: it shall be interpreted as the solution of an equation driven by a noise with a finite correlation length λ which is much smaller than δx. We provide an example in Fig. 3. Here we start from an initial condition $U(x, 0) = K/2$ for $x \in [-5, 5] \times [5, 5]$. Using an explicit Euler scheme in time, we compute the finite difference approximation with a regular mesh size $\delta x = 0.05$, over a bounded domain with Neumann boundary conditions, i.e. $\nabla U(x, t) = 0$ when x is on the boundary of the domain. Matlab codes are available at https://osf.io/w5utd/. We observe in panels (a,b) that the corresponding random fields tends to be much more clustered than those generated by $(\mathcal{S})$ (compare with Fig. 2a). Moreover, increasing diffusion coefficients D tend to create larger clusters. As expected from the discussion above, at a fixed time T, $U(x, T)$ has a very clear bimodal distribution despite the Gaussian random noise W_t, with one mode close to 0 and the other one close to K (see Figs. 3c and 3d).

4. SDEs and point processes: the case of dispersing particle systems

We now describe a theoretical framework for modelling the spatio-temporal dynamics of particles and explore the point processes that such a framework generates. The point process viewpoint is then used to tackle the estimation of model parameters using spatio-temporal observations. This estimation task naturally fits into the mechanistic-statistical approach (e.g., Soubeyrand and Roques, 2014), which bridges the gap between mechanistic models (here a Lévy flight model describing particle dynamics; see Eq. (3)) and data, via a probabilistic observation model. Here, observations are derived from an inhomogeneous spatio-temporal point process whose intensity follows a PDE, namely the Fokker–Planck Eq. (6) associated with the stochastic dispersal process. Other studies rather use the INLA/SPDE approach to analyse point processes over space and time, but the literature is relatively scarce (see e.g., Simpson et al., 2016; Opitz et al., 2020; Yuan et al., 2017; Koh et al., 2021). In these studies, the point process is modelled as an inhomogeneous Poisson process conditional on a random log-Gaussian intensity.

A stochastic model for the dispersal and deposition of particles. The positions of the particles at any time $t \geq 0$ are described by a vector field $\mathbf{X}_t$ of size $d \times N_t$, where d is the space dimension, usually 2 or 3, and N_t is the number of particles in our system at time $t \geq 0$. The initial number of particles N_0 follows a Poisson distribution with parameter $E[N_0] = \eta_0$, and these particles are distributed in $\mathbb{R}^d$ according to a probability distribution function p_0 (hence, the initial pattern of particle locations is drawn from an inhomogeneous Poisson point process; Illian et al., 2008; Chiu et al., 2013). Between $t = 0$ and the deposition times, the particle dynamics are described by independent d-dimensional α-stable Lévy processes, for some $\alpha \in (0, 2]$, see Eq. (3).

As in Section 2, we assume that the expected duration of the dispersal before deposition (or removal) is $1/\kappa^2 > 0$. The deposition events are independent and identically distributed from an exponential distribution with parameter κ^2. Namely, for each $i \in \{1, \ldots, N_0\}$, the deposition time τ_i of the particle i satisfies $\tau_i \sim \text{Exp}(\kappa^2)$.

Fokker–Planck equations. As already mentioned in Section 2, the expected density of still dispersing particles $u(x, t)$ satisfies Eq. (6). Here, N_0 is a random variable, and the initial condition is $u_0(x) := u(x, 0) = \eta_0\, p_0(x)$. The only non-conservative term in Eq. (6) satisfied by u is $-\kappa^2\, u$. It corresponds to particle removal due to deposition. The expected density $w(x, t)$ of deposited particles thus satisfies $\partial_t w(x, t) = \kappa^2\, u(x, t),\ t > 0,\ x \in \mathbb{R}^d$, with initial condition $w(\cdot, 0) = 0$ (all of the particles are dispersing at the initial time $t = 0$).

Inhomogeneous spatial Poisson point process. At a fixed time t, the spatial point process formed by the dispersing particles is an inhomogeneous Poisson point process with intensity function $u(\cdot, t)$; see a more precise formulation of this rather straightforward result in Appendix A, Lemma 1. Indeed, let A be a bounded region in $\mathbb{R}^d$, and denote $N_t(A)$ the number of dispersing particles in A at time t. We have

$$\mathbb{P}(N_t(A) = m) = \sum_{k=m}^{\infty} \mathbb{P}(N_t(A) = m | N_0 = k)\mathbb{P}(N_0 = k).$$

As the particles follow independent movements,

$$\mathbb{P}(N_t(A) = m | N_0 = k) = \binom{k}{m} \rho^m (1 - \rho)^{k-m},$$

with $\rho = \int_A h(y, t)\, dy$ the probability that a given particle is still dispersing (alive) and belongs to A at time t. The function h was defined in Section 2; it satisfies the same equation as u but with the initial condition $h(\cdot, 0) = p_0$. Thus,

$$\mathbb{P}(N_t(A) = m) = \sum_{k=0}^{\infty} \binom{k}{m} \rho^m (1 - \rho)^{k-m} \frac{\eta_0^k}{k!} e^{-\eta_0} = e^{-\rho\, \eta_0} \frac{(\rho\, \eta_0)^m}{m!}.$$

As $\rho\, \eta_0 = \int_A u(y, t)\, dy$, this shows that, at each fixed time t, the distribution of dispersing particles follows an inhomogeneous Poisson point process with intensity $u(\cdot, t)$. The same reasoning implies

that, at each fixed time t, the distribution of deposited (dead) particles follows an inhomogeneous Poisson point process with intensity $w(\cdot, t)$.

Non-Poisson spatio-temporal point process. Despite the independence of particle trajectories, the spatio-temporal point process formed by the locations of still-dispersing particles is not a spatio-temporal Poisson point process because of the temporal dependencies of the successive locations of any particle (see Appendix A, Lemma 2 with $A = \mathbb{R}^d$). However, we can state that the point process $\mathbf{X}_t$ formed by the particle locations at time t given the particle locations $\mathbf{X}_s$ ($s < t$) is the union of N_s independent thinned binomial point processes with size 1, thinning probability $1-\exp(-(t-s)\kappa^2)$, and p.d.f. $p^s_{\delta(\cdot - X^i_s)}(\cdot, t)$, where X^i_s is the point of $\mathbf{X}_s$ with label i and $p^s_{\delta(\cdot-X^i_s)}$ satisfies Eq. (5) with initial condition $p^s_{\delta(\cdot-X^i_s)}(\cdot, s) = \delta(\cdot - X^i_s)$ at time s (see Appendix A, Lemma 3 with $A = \mathbb{R}^d$). More generally, based on Lemma 3, the point process $\mathbf{X}_t$ conditional on particle locations in a subset $A \subset \mathbb{R}^d$ at time $s < t$ (without being able to *recognise* the particles and *follow them* across sampling times) leads to a point process made of the union of an inhomogeneous Poisson point process and independent thinned binomial point processes with size 1.

Likelihood- and pseudo-likelihood-based estimation. Let us consider a family of observation windows $\{A_k\}_{k=1,\dots,K}$, and an increasing sequence $\{t_k\}_{k=1,\dots,K}$ of observation times. We focus on the computation of a likelihood based on the observation of dispersing particles located within the spatial$\times$time windows $\{A_k, t_k\}_{k=1,\dots,K}$ to estimate the model parameters; we could use observations of deposited particles as well. Let $\mathbf{X}^{(k)}$ denote the observation of $\mathbf{X}_{t_k}$ restricted to A_k. The likelihood satisfies

$$\mathcal{L}(\Theta) := \mathbb{P}(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(K)}) = \mathbb{P}(\mathbf{X}^{(1)}) \prod_{k=2}^{K} \mathbb{P}(\mathbf{X}^{(k)} \mid \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(k-1)}),$$

where Θ is the vector of model parameters to be estimated. The conditional probability $\mathbb{P}(\mathbf{X}^{(k)} \mid \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(k-1)})$ simplifies into $\mathbb{P}(\mathbf{X}^{(k)} \mid \mathbf{X}^{(k-1)})$ only if $A_{k-1} = \mathbb{R}^d$ (in this case, for any $k' \geq k$, $\mathbb{P}(\mathbf{X}^{(k')} \mid \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(k'-1)}) = \mathbb{P}(\mathbf{X}^{(k')} \mid \mathbf{X}^{(k-1)}, \dots, \mathbf{X}^{(k'-1)})$). If $A_{k-1} \neq \mathbb{R}^d$ for some k (a very standard situation), the likelihood can be written as a multiple integral with respect to the unobserved parts of $\mathbf{X}_{t_1}, \dots, \mathbf{X}_{t_{K-1}}$, say $\tilde{\mathbf{X}}^{(1)}, \dots, \tilde{\mathbf{X}}^{(K-1)}$ (with $\tilde{\mathbf{X}}^{(k)} = \mathbf{X}_{t_k} \backslash \mathbf{X}^{(k)}$). Namely, we show by induction (see Appendix A.3) that:

$$\mathcal{L}(\Theta) = \int_{\tilde{\mathbf{X}}^{(K-1)}, \dots, \tilde{\mathbf{X}}^{(1)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=2}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) d\mathbb{P}(\mathbf{X}_{t_1}). \tag{26}$$

$\mathbb{P}(\mathbf{X}_{t_1})$ is the probability measure of the point pattern $\mathbf{X}_{t_1}$, which is drawn from an inhomogeneous Poisson point process with intensity $u(\cdot, t_1)$ satisfying Eq. (6) with initial condition $\eta_0 p_0$ at $t = 0$. $\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}})$ can be written:

$$\begin{aligned} \mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) &= \sum_{\psi_k} \mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}, \psi_k) \mathbb{P}(\psi_k) \\ &= \sum_{\psi_k} \left(\prod_{i=1}^{N_{t_k}} p^{t_{k-1}}_{\delta(\cdot - X^{\psi_{ki}}_{t_{k-1}})}(X^i_{t_k}, t_k) \right) \mathbb{P}(\psi_k), \end{aligned} \tag{27}$$

where the sum is over all the ordered samplings ψ_k of N_{t_k} elements from $N_{t_{k-1}}$ elements (without replacement). Namely, $\psi_k = (\psi_{k1}, \dots, \psi_{kN_{t_k}})$ where ψ_{ki} indicates to which point in $\mathbf{X}_{t_{k-1}}$ corresponds the ith point $X^i_{t_k}$ in $\mathbf{X}_{t_k}$ and $p^{t_{k-1}}_{\delta(\cdot - X^{\psi_{ki}}_{t_{k-1}})}(\cdot, t_k)$ (defined in the notation paragraph of Appendix A) is the probability density function of $X^i_{t_k}$ given the particle i was at $X^{\psi_{ki}}_{t_{k-1}}$ at time t_{k-1}. Note that N_{t_k} is finite since it is less than N_0, which is Poisson-distributed with mean η_0. Similarly to $\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}})$, $\mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}})$ can be written:

$$\mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) = \sum_{\psi_K} \left(\prod_{i=1}^{N_{t_K}(A_K)} p^{t_{K-1}}_{\delta(\cdot - X^{\psi_{Ki}}_{t_{K-1}})}(X^{(K)i}, t_K) \right) \mathbb{P}(\psi_K), \tag{28}$$

where $X^{(K)i}$ is the location of the ith point of $\mathbf{X}^{(K)}$.

Maximum likelihood estimation or Bayesian estimation can be performed using the likelihood expression (26). The multiple integration (with respect to the latent point patterns $\tilde{\mathbf{X}}^{(1)}, \ldots, \tilde{\mathbf{X}}^{(K-1)}$ over $A_1, \ldots, A_{K-1}$ at times $t_1, \ldots, t_{K-1}$ and the latent vectors $\psi_2, \ldots, \psi_K$) may be handled with stochastic estimation algorithms such as SEM, MCEM, MCMC or adaptive importance sampling. However, such approaches likely require high computational cost.

Alternatively, we propose here to use a pseudo-likelihood function constructed by ignoring the temporal dependencies except the dependencies generated by the temporal evolution of the intensity function of $\mathbf{X}_t$, that is to say the intensity function u satisfying Eq. (6) with initial condition $\eta_0 p_0$; see Appendix A, Lemma 1 and Corollary 1. Thus, we ignore dependencies between observations at successive observation times and make as if $\{\mathbf{X}_t\}_{t\geq 0}$ is an inhomogeneous spatio-temporal Poisson point process with spatio-temporal intensity u (u is thereafter denoted by u_Θ to highlight the dependence of u with respect to the parameters). Consequently, the pseudo-likelihood satisfies (see e.g. Møller and Waagepetersen, 2007):

$$\tilde{\mathcal{L}}(\Theta) = \prod_{k=1}^{K} e^{|A_k| - \varphi_{\Theta,t_k}(A_k)} \prod_{X^i_{t_k} \in A_k} u_\Theta(t_k, X^i_{t_k}), \tag{29}$$

with $|A_k|$ the Lebesgue measure of A_k, and $\varphi_{\Theta,t_k}(A_k) = \int_{A_k} u_\Theta(t_k, Y)\, dY$ the intensity measure of A_k at time t_k.

Assume now as a benchmark that the observations only consist of counting data, i.e., numbers of particles in windows A_k at times t_k, $k = 1, \ldots, K$. Ignoring as above temporal dependencies except the dependencies generated by the temporal evolution of the intensity function, the numbers of particles in windows A_ks at observation times t_ks follow independent Poisson distributions with mean values $\varphi_{\Theta,t_k}(A_k)$ (intensity measure of A_k), and the pseudo-likelihood now reads:

$$\tilde{\mathcal{L}}_c(\Theta) = \prod_{k=1}^{K} [\varphi_{\Theta,t_k}(A_k)]^{N_{t_k}(A_k)} \frac{e^{-\varphi_{\Theta,t_k}(A_k)}}{N_{t_k}(A_k)!}. \tag{30}$$

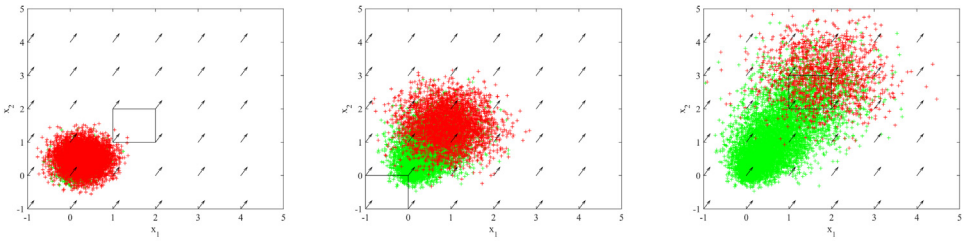
Thus, in the proposed mechanistic-statistical approach, we consider a unique mechanistic model (for which we aim to estimate parameters) but we consider two different observation processes, consisting of either observing particle locations in A_ks at t_ks (in which case $\tilde{\mathcal{L}}(\Theta)$ is maximised), or counting particles in A_ks at t_ks, (in which case $\tilde{\mathcal{L}}_c(\Theta)$ is maximised).

Numerical example. We ran 10^3 simulations of the discrete particle system in $\mathbb{R}^2$ with a Brownian diffusion ($\alpha = 2$, $\gamma = 1/2$), and spatially-homogeneous coefficients $B = (0.3, 0.5)$, $\sigma = 0.3$ and mean deposition time $1/\kappa^2 = 3$. The particles were released in a small disc around $x = 0$, with an initial number drawn in a Poisson distribution with parameter $\eta_0 = 10^4$. Assuming that the initial distribution is $p_0(x) = \eta_0\, \delta(x)$, the solution $u(x, t)$ of the Fokker–Planck equation has an analytic expression

$$\tilde{u}(x, t) = \frac{\eta_0\, e^{-\kappa^2 t}}{(\sigma\sqrt{2\pi t})^d}\, e^{-\frac{|x - Bt|^2}{2\sigma^2 t}}, \quad \text{for } t > 0. \tag{31}$$

Regarding the observations, we focused on the dispersing particles, and considered 3 observation times $t_1 = 1$, $t_2 = 3$, $t_3 = 6$ and 3 observation windows $A_1 = (0, 1) \times (0, 1)$, $A_2 = (-1, 0) \times (-1, 0)$ and $A_3 = (1, 2) \times (2, 3)$, see Fig. 4.

The unknown parameters are $\Theta = (B(1), B(2), \sigma, \kappa)$. The parameters are estimated by maximising the pseudo-likelihood $\tilde{\mathcal{L}}(\Theta)$ in (29) when using the full spatial observation data, and by using $\tilde{\mathcal{L}}_c(\Theta)$ in (30) when the information is restricted to the counting. We assumed the following constraints on the parameter values: $\Theta \in \Gamma := (-1, 1) \times (-1, 1) \times (0.01, 1) \times (0.1, 10)$. We used the Matlab© built-in gradient-based minimisation algorithm fmincon© with default options, and applied to $-\log(\tilde{\mathcal{L}})$ to compute the maximum pseudo-likelihood estimators (MPLE). The results are summarised in Table 2. As expected, the observation of the positions of the particles leads



(a) $t_1 = 1$ (b) $t_2 = 3$ (c) $t_3 = 6$

Fig. 4. Observation windows. Dispersing particles (red crosses) and deposited particles (green crosses) at successive times $t = 1, 3, 6$. The black arrows describe the vector field B, and the black rectangles are the observation windows A_1, A_2, A_3. The particles were initially uniformly distributed in a disc of radius 0.05, centred at 0. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Mean and standard deviation (based on 10^3 repetitions) of the maximum pseudo-likelihood estimators (MPLE) obtained from observations consisting of either particle locations (using $\tilde{\mathcal{L}}(\Theta)$ in (29)) or counting data (using $\tilde{\mathcal{L}}_c(\Theta)$ in (30)).

Parameter	True value	MPLE - locations	MPLE - counting
$B(1)$	0.3	0.30 (0.02)	0.36 (0.10)
$B(2)$	0.5	0.50 (0.02)	0.56 (0.07)
σ	0.3	0.30 (0.02)	0.27 (0.03)
$1/\kappa^2$	3	3.04 (0.23)	8.34 (2.28)

to a more accurate estimation compared to counting data. In addition, in this example, the use of the pseudo-likelihood (instead of the likelihood) and the simplifying approximation concerning the initial distribution of particles provides a quite satisfactory estimation accuracy when particle locations are observed.

5. Discussion

In this work, we have revisited the SPDE approach now commonly used in spatial statistics, as proposed in Lindgren et al. (2011) and further developed in Carrizo Vergara et al. (2021). We have adopted a mechanistic perspective based on the movement of microscopic particles, which led us to relate pseudo-differential operators to dispersal kernels. This original mechanistic approach led us to a new and deeper understanding of a certain class of models and paved the way to new models with original features that do not fit in the general framework of Carrizo Vergara et al. (2021) based on Fourier transforms, such as drifts and nonlinear reaction terms. We also showed on a synthetic case study how such mechanistic models, associated to a probabilistic observation model, can be used in a hierarchical setting to estimate the parameters of the particle dynamics. This approach has a large potential in ecology and epidemiology for the studies of the dynamics of populations that can be described at the resolution of individuals (animals, seeds, pollen grains, pathogens, spores; Tufto et al., 1997; Klein et al., 2003; Dragon et al., 2012; Soubeyrand et al., 2015).

Regarding stationary models on which a Fourier analysis is possible, we provided in Section 3.1 a detailed exposition of several dispersal operators, with the aim to bridge the gap between the mechanistic view (through dispersal kernels) and the statistical view (through covariance functions). Our main findings are summarised in Table 1. We first have shown that the solution of the evolving Matérn equation (10) is driven by three main forces: dispersal through a thin-tailed dispersal kernel; absorption through a negative linear reaction term; and additive noise. The spatial symbol function is $g(\xi) = (\|\xi\|^2 + \kappa^2)^{\alpha/2}$ and the associated covariance function is a Matérn function which decays exponentially fast as $\|h\| \to \infty$. We then turned our attention to the much less usual fractional Laplacian operator with a linear reaction term. In this case, the dispersal is driven by a fat-tailed

kernel, and the corresponding symbol function is $g(\xi) = \|\xi\|^\alpha + \kappa^2$. The associated covariance functions describe long-distance dependencies, see Eqs (17) and (19). The difference between these covariance functions and the usual Matérn covariance is illustrated in Fig. 1. The difference between the thin-tailed and fat-tailed long distance behaviours is clearly visible, even though both covariances are comparable at short positive distances. Notice that the two operators are equal when $\alpha = 2$, which corresponds to the usual Laplacian operator with linear reaction. Finally, we also considered integral dispersal operators which are directly defined through a convolution with a dispersal kernel. This makes it possible to consider very general dispersal tails, from very thin (e.g., Gaussian) to very fat (e.g., with algebraic decay). With such operators, the computations of the symbol function g and therefore of the spectral measure are straightforward. We derived explicit (spatial) covariance functions for exponential kernels which, as expected, are closely related to the Matérn family. When the kernel is a general Matérn function (with integer parameter ν), it is shown that the associated Gaussian random field is the sum of Markov random fields. As illustrated with the fractional Laplacian operator, other types of kernels should lead to other families of covariance functions. However, the derivation of explicit expressions is not straightforward in general. These new models can be useful in various geostatistical applications, in particular when a physical interpretation of the short or long range dependencies is available. In ecology and epidemiology, examples are aerial transport of particles, dissemination of spores and other propagules (Klein et al., 2006; Soubeyrand et al., 2015), and disease spread caused by human movement (Bjørnstad et al., 2019; Roques et al., 2020).

When a Fourier analysis is not possible, for instance with spatially heterogeneous coefficients or with a nonlinear reaction term, the framework of Carrizo Vergara et al. (2021) cannot be applied as such. Nevertheless, the mechanistic approach that we have proposed allows to construct spatio-temporal random fields, based on an intuitive approach. Of course, their mathematical characterisation should rely on much more work and is far beyond the scope of our work. Still, the examples in Sections 3.2 and 3.3 illustrate the interest of this intuitive approach. First, to construct spatio-temporal random fields which reflect the effect of external driving forces, such as the effect of wind on a concentration. It appears in the simulations of Section 3.2 that adding a spatio-temporally varying drift term in the equations leads to a particular form of spatio-temporal patterns which are consistent with the underlying physical assumptions. When such external constraints are known, it seems essential to take them into account for the realism of the model. Second, to tackle the case with several attracting states. For instance, as we explained in Section 3.3, many biological systems have bistable dynamics, with typically the steady state 0 (the absence of the species) and another positive steady state. We proposed here to construct spatio-temporal random fields which reflect these dynamics through stochastic Allen–Cahn equations. These equations must be used with care, as their well-posedness depends on the dimension and on the noise term. Nevertheless, our preliminary computations show that the characteristics of the corresponding random fields meet our expectations, especially regarding the bimodality of the distributions. This advocates for further mathematical studies of these equations.

In Section 3.2, we have considered spatio-temporally varying drift terms. Spatio-temporally varying diffusion coefficients may be considered as well. Two main types of diffusion operators are found in the reaction–diffusion literature: the Fokker–Planck diffusion operator $\mathcal{D}(t, x, [u]) = \Delta(D(x, t)u)$, and the Fickian diffusion operator $\mathcal{D}(t, x, [u]) = \mathrm{div}(D(x, t)\nabla u)$ (Roques, 2013), see also (Alfaro et al., 2021) for generalisations of the form $\mathcal{D}(t, x, [u]) = \mathrm{div}(D(x, t)^q \nabla (D(x, t)^{1-q} u))$ ($q \in (0, 1)$). Both of them have anisotropic versions (see Remark 1), but for simplicity we consider here the isotropic case. We have observed in Section 2 that the Fokker–Planck diffusion emerges as the macroscopic limit of a stochastic diffusion process with spatially-dependent diffusion term $\sigma(x, t) = \sqrt{2\,D(x, t)}$. On the other hand Fickian diffusion is usually derived from physical relationships between particle flux, concentration and concentration gradient, in which the driving force along time is that a system seeks equilibrium. If in some elementary space–time volume the density is higher than in the neighbouring area, particles will move from the denser area to the less dense ones. More precisely, the flux is assumed to be linearly linked to the particle concentration gradient (Fick's Law), with a coefficient $D(x, t)$ that accounts for the spatial variations of the environment. Thus, the Fokker–Planck diffusion operator is generally better suited to the description of particle movements (see

e.g., Hannunen and Ekbom, 2001; Roques et al., 2008), while the Fickian diffusion operator is more suitable for the description of physical phenomena, such as the dilution of a dye in a liquid or heat diffusion in a heterogeneous body (Fick, 1855). We can note that Fokker–Planck diffusion equation can also be written: $\Delta(D\,u) = \mathrm{div}(\nabla[D\,u]) = \mathrm{div}(D\,\nabla u) + \mathrm{div}(u\nabla D)$. We thus obtain a Fickian diffusion with a drift term oriented opposite to the gradient of D. Thus, these two operators, which are identical when the coefficient $D(t,x)$ is constant, lead to different behaviours of the solution in a heterogeneous environment. Fickian tends to homogenise the solution, while with the Fokker–Planck equation, the solution will tend to concentrate in the regions where D is small, due to the drift term $\mathrm{div}(u\nabla D)$. This is consistent with the above explanations: for particles, one can expect to find a greater concentration in regions of low mobility; conversely, for heat diffusion, a homogenisation phenomenon is expected, whatever the diffusion in the media considered. We briefly checked numerically whether some obvious differences could be observed on the solutions of (7), but did not observe significant differences, see Appendix B. This finding is in strong contrast with the deterministic case, where the solutions tend at large times to become proportional to $1/D(x)$ (if D does not depend on time) with a Fokker–Planck diffusion vs to a constant with a Fickian diffusion.

In our opinion, this preliminary work opens new avenues for research in an effort to further tie together spatial statistics and PDEs. We see at least three main lines of research. First, as already pointed out above, more effort should be made to propose new SPDEs with physical or biological interpretations: including drifts, non homogeneous parameters and nonlinear reactions are key for a flexible and useful modelling. However, the theoretical mathematical behaviour is unclear or unknown in some cases, but numerical solutions are possible. In this work, the exponent α was set to be less than or equal to 2. While the Fourier definition $\mathcal{F}(-(-\Delta)^{\alpha/2}u)(\xi) = -\|\xi\|^{\alpha}\mathcal{F}(u)(\xi)$ allows to consider values $\alpha > 2$, the mechanistic construction of the fractional Laplacian proposed in Section 2 and the convolution formula (11) cannot be used for such values of α. Another approach uses hypersingular integrals based on higher order finite differences (at least of order $\lceil\alpha\rceil$) instead of the first order finite difference in (11) (see Part 1, Chapter 3 in Samko, 2001). However, the interpretation of this operator in terms of particle motion is far from obvious.

As a second line of research, algorithmic developments are required for these models, in particular in a spatio-temporal context where datasets can be large or massive. Using the Galerkin method as in Lindgren et al. (2011) is one possible option, but care must be taken because in application of the Rozanov theorem (Rozanov, 1977) the Markov property is lost when the inverse of the spectral density is not a polynomial of $\|\xi\|^2$, which implies dense precision matrices. One way around could be to approximate the spectral density using the rational SPDE approach advocated in Bolin and Kirchner (2020) or the Galerkin–Chebyshev approximation proposed in Lang and Pereira (2021) for models involving the fractional Laplacian. At the cost of more intense computational effort, both of these approaches build precision matrices with an interesting amount of sparsity. A completely different approach is a direct estimation of the parameters of the SPDE via a mechanistic-statistical model based on a SPDE mechanistic core, or simply a least square minimisation argument. The numerical analysis community has developed several tools for solving PDEs, such as Freefem++ (Hecht, 2012), which can be adapted to handle SPDEs (e.g., Boulakia et al., 2015). However, the estimation of the parameters requires numerous calls to the solver and would be computationally expensive. Nevertheless, new approaches to the simulation and estimation of PDEs, such as physics-informed neural networks (e.g. Raissi et al., 2017) may offer promising prospects for the direct estimation of the parameters of SPDEs especially in the context of massive datasets.

As a third line of research, we advocate the use of hybrid models in the spirit of the application shown in Section 4 combining (presumably in a hierarchical setting) mechanistic, probabilistic and statistical compartments, on which inference of the parameters can be made in a frequentist or in a Bayesian context. Interestingly, the proposed approach leads to spatial and spatio-temporal point processes with a mechanistic foundation and an interpretable intensity function. These point processes, as ordinary point processes, can be fitted to point patterns, counting data, and even information about the presence or absence of points. However, due to their construction, we can also use trajectory data to estimate their parameters. Whatever the sampling density of the trajectories,

the crucial feature is the ability to follow certain points in time, that is to say to have (at least partial) information about the ψ_ks in Eqs (27)–(28). Such data are expected to bring accurate information about the parameters but, for fully exploiting their information content, the true likelihood (26) should be used, or at least a new pseudo-likelihood, yet to be proposed, accounting for full temporal dependencies in the locations of particles that are tracked across time. It would also be interesting to combine several types of data (locations, counting, trajectories) and even being able to advise how to mix them for maximising the information they bring for a limited total sampling effort. This task could be carried out with the study of the Fisher information, at least in simple cases where the expression of u is explicit like in Eq. (31). Finally, including direct dependencies between trajectories of particles (generated by interactions between particles such as attraction like in Soubeyrand et al., 2011 or repulsion) deserves to be studied to gain in modelling realism and to be able to handle more complex dynamics. In this case, the challenges would be to derive (i) the mechanistic equations satisfied by the intensity function (this could take a similar course to the Keller–Segel approach, possibly with fractional diffusion as in Bournaveas and Calvez, 2010), (ii) the nature of the resulting point processes, which are expected to be more complex than those obtained in the case of independent trajectories, and (iii) an estimation approach that would allow us to achieve accurate parameter estimators.

Appendix A. Point process generated by the SDE-based model of dispersing particles

A.1. Reminder of and supplementary notations

$\mathbf{X}_t$ is the process of locations of still dispersing particles at time $t \geq 0$, N_0 the initial number of particles that is Poisson-distributed with mean η_0, and τ_i the deposition time of particle i (τ_is are independently and identically distributed from the exponential distribution with mean $1/\kappa^2$). The initial locations (given by $\mathbf{X}_0$) are independently and identically distributed from the probability distribution function p_0, which can be either discrete or continuous on the space $\mathbb{R}^d$. When a particle has been deposited, it is no more included in the point process $\mathbf{X}_t$.

$A, A_1, \ldots, A_K$ are non-empty Lebesgue-measurable subsets of $\mathbb{R}^d$, $K \in \mathbb{N}^*$, $0 < s < t$ and $0 < t_1 < \cdots < t_K$. A^c is the complementary set of A, i.e., $A \cup A^c = \mathbb{R}^d$ and $A \cap A^c = \emptyset$. For $t \in [0, \tau_i)$, X_t^i is the location of particle i at time t. $\mathbf{X}_{tA}$ is the restriction of $\mathbf{X}_t$ over A. I_{tA} is the subset of indices in $\{1, \ldots, N_0\}$ corresponding to the points $\mathbf{X}_{tA}$. $N_t(A)$ is the number of particle located in A at t. We use the notation $\mathbb{P}$ to denote the probability measure of variables or sets of variables. Conditional probability measures are written $\mathbb{P}(\cdot \mid \cdot)$. We use the abbreviation p.d.f. for probability distribution function.

For any $s \geq 0$ and any distribution q on $\mathbb{R}^d$, we denote by $u_q^s(\cdot, t)$ the solution of Eq. (6) starting at time s with the condition $u(\cdot, s) = q(\cdot)$. Similarly, we denote by $p_q^s(\cdot, t)$ the solution of Eq. (5) starting at time s with the condition $p(\cdot, s) = q(\cdot)$. We note that the two quantities are related by the relationship $u_q^s(\cdot, t) = e^{-(t-s)\kappa^2} p_q^s(\cdot, t)$ for all $t \geq s$. When q is the sum of Dirac measures centred at the locations of the particles $\mathbf{X}_{sA}$, $q(x) = \sum_{i=1}^{N_s(A)} \delta(\cdot - X_{sA}^i)$, we simply write $u_{\mathbf{X}_{sA}}^s(\cdot, t)$ (resp. $p_{\mathbf{X}_{sA}}^s(\cdot, t)$).

A.2. Lemmas

Lemma 1. *For any fixed time $t > 0$, the spatial point process $\mathbf{X}_t$ is an inhomogeneous Poisson point process with intensity $u_{\eta_0 p_0}^0(\cdot, t)$.*

$\mathbf{X}_t$ is a spatial Poisson point process with intensity function ρ defined over $\mathbb{R}^d$ if for any bounded region $A \subset \mathbb{R}^d$, $N_t(A)$ is Poisson distributed with mean $\int_A \rho(x)dx$, and conditioned on $N_t(A) = n$, the n events in A are independent and identically distributed with density proportional to ρ restricted to A. The elements of proof for Lemma 1 are provided in the main text (Section 4).

Lemma 2. *For any fixed times $t > s > 0$, the spatial point process $\mathbf{X}_t$ conditional on the partial observation $\mathbf{X}_{sA}$ of $\mathbf{X}_s$ restricted to A, is not an inhomogeneous Poisson point process.*

Proof of Lemma 2. Let A be an open bounded set and assume that $N_s(A) > 0$. Let $B \subset A$ an open set such that $N_s(B) = 1$, and denote $X_t^1 \in \mathbb{R}^d$ the position of this single particle at times $t \geq s$. According to Section 2, the expected particle density at times $t \geq s$ is $u_q^s(\cdot, t)$, with the initial condition (here, at time s),

$$q(x) = \delta(x - X_s^1) + \tilde{u}(x), \text{ with } \int_B \tilde{u}(y)\, dy = 0,$$

where δ is a Dirac measure at 0. By continuity of $t \mapsto \int_B u(y,t)\, dy$ at $t = s$, this implies that $\lim_{t\to s^+} E[N_t(B)] = \int_B u(y,s)\, dy = 1$. Assume that the spatial point process $\mathbf{X}_t$ conditional on the partial observation $\mathbf{X}_{sA}$ of $\mathbf{X}_s$ restricted to A is an inhomogeneous Poisson point process for $t > s$. Then, as $\lim_{t\to s^+} E[N_t(B)] = 1$, we have

$$\lim_{t\to s^+} P(N_t(B) \geq 2) = 1 - 2\, e^{-1}. \tag{32}$$

On the other hand, the probability $P(X_t^1 \in B \text{ and } t < \tau_1)$ that X_t^1 is still dispersing and belongs to B at time $t \geq s$ is given by $\int_B h(y,s)\, dy$, with h the solution of Eq. (6), this time with initial condition $h(x,s) = \delta(x - X_s^1)$. By continuity, one has $\lim_{t\to s^+} P(X_t^1 \in B \text{ and } t < \tau_1) = 1$. Writing

$$E[N_t(B)] \geq P(X_t^1 \in B \text{ and } t < \tau_1) + P(\exists\, i \neq 1 \text{ s.t. } X_t^i \in B \text{ and } t < \tau_i),$$

and passing to the limit $t \to s^+$, we obtain that $\lim_{t\to s^+} P(\exists\, i \neq 1 \text{ s.t. } X_t^i \in B \text{ and } t < \tau_i) = 0$. This contradicts (32). □

Intuitively, the non-Poisson nature of $\mathbf{X}_t \mid \mathbf{X}_{sA}$ can be understood by looking at the variance of $N_{s^+}(A)$ when $t = s^+$ ($0 < s^+ - s \ll 1$). Assume that $N_s(A) \geq 1$. We have $V(N_{s^+}(A) \mid \mathbf{X}_{sA}) \approx V(N_s(A) \mid \mathbf{X}_{sA}) = 0$. On the other hand, if $\mathbf{X}_t \mid \mathbf{X}_{sA}$ was a Poisson process, we would have $V(N_{s^+}(A) \mid \mathbf{X}_{sA}) = E(N_{s^+}(A) \mid \mathbf{X}_{sA}) \approx N_s(A)$. Hence, given X_{sA}, the number of points in A at s^+ cannot be Poisson-distributed.

Lemma 3. *For any fixed times $t > s \geq 0$, the point process $\mathbf{X}_t$ conditional on the partial observation $\mathbf{X}_{sA}$ of $\mathbf{X}_s$ restricted to A, is the union of a Poisson point process with intensity $u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A^c}(\cdot)}(\cdot, t)$ and $N_s(A)$ independent thinned binomial point processes with size 1, thinning probability $1-\exp(-(t-s)\kappa^2)$, and p.d.f. $p^s_{\delta(\cdot - X^i_{sA})}(\cdot, t)$, where X^i_{sA} is the point of $\mathbf{X}_{sA}$ with label i and $\mathbb{1}_v$, $v \subset \mathbb{R}^d$, is the indicator function of v.*

Proof of Lemma 3. Each point in $\mathbf{X}_{sA}$ independently generates at time t either zero point or one point. It generates zero point if the particle is deposited between s and t, which occurs with probability $1 - \exp(-(t-s)\kappa^2)$. Otherwise, it generates one point whose probability distribution function over $\mathbb{R}^d$ is $p^s_{\delta(\cdot - X^i_{sA})}(\cdot, t)$.

From Lemma 1, $\mathbf{X}_s$ is an inhomogeneous Poisson point process with intensity $u^0_{\eta_0 p_0}(\cdot, s)$. Setting $\mathbf{X}_s = \mathbf{X}_{sA} + \mathbf{X}_{sA^c}$ and because of the Poisson nature of $\mathbf{X}_s$, $\mathbf{X}_{sA}$ and $\mathbf{X}_{sA^c}$ are independent and $(\mathbf{X}_{sA^c} \mid \mathbf{X}_{sA}) \equiv \mathbf{X}_{sA^c}$ in distribution. Hence, $\mathbf{X}_{sA^c} \mid \mathbf{X}_{sA}$ is an inhomogeneous Poisson point process with intensity $u^0_{\eta_0 p_0}(\cdot, s)\mathbb{1}_{A^c}(\cdot)$. It follows that the number of dispersing particles in A^c at s is drawn from a Poisson distribution with mean $\int_{A^c} u^0_{\eta_0 p_0}(\cdot, s)$, and the particle depositions are i.i.d. from the p.d.f. $u^0_{\eta_0 p_0}(\cdot, s)\mathbb{1}_{A^c}(\cdot) / \int_{A^c} u^0_{\eta_0 p_0}(\cdot, s)$. Hence, we can apply Lemma 1 between times s and t to determine the distribution of the part of $\mathbf{X}_t$ that was restricted to A^c at time s. This part of $\mathbf{X}_t$ is an inhomogeneous Poisson point process with intensity $u^s_{E(\mathbf{X}_{sA^c})}(\cdot, t) = u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A^c}(\cdot)}(\cdot, t)$. □

Corollary 1. *For any fixed time $t > 0$, the intensity function of the point process $\mathbf{X}_t$ conditional on the partial observation $\mathbf{X}_{sA}$ of $\mathbf{X}_s$ restricted to A is $x \mapsto u^s_{\mathbf{X}_{sA}}(x, t) + u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A^c}(\cdot)}(x, t)$ whose expectation is $u^0_{\eta_0 p_0}(\cdot, t)$.*

Proof of Corollary 1. The first part of the proof of Lemma 3 yields that the sub-point process $\mathbf{X}_{sA}$ generates at t a (non-Poisson) point process with conditional intensity $\sum_{i=1}^{N_s(A)} u^s_{\delta(\cdot - X^i_{sA})}(x,t) = u^s_{\mathbf{X}_{sA}}(x,t)$ given $\mathbf{X}_{sA}$. Moreover, from Lemma 3, the remaining points in $\mathbf{X}_t$ (i.e., the locations at t of particles that were in A^c at s) have intensity $u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A^c}(\cdot)}(x,t)$. The sum of the two intensities gives the intensity of $\mathbf{X}_t$ given $\mathbf{X}_{sA}$. This is clear with the following reasoning: Let Λ denote the intensity of $\mathbf{X}_t$ conditional on $\mathbf{X}_s$: $\Lambda(x) = u^s_{\mathbf{X}_s}(x,t) = u^s_{\mathbf{X}_{sA}}(x,t) + u^s_{\mathbf{X}_{sA^c}}(x,t)$. The expectation of Λ given $\mathbf{X}_{sA}$ is the intensity of $\mathbf{X}_t \mid \mathbf{X}_{sA}$ and satisfies, because of the linearity of Eq. (6):

$$\begin{aligned}
E(\Lambda(x) \mid \mathbf{X}_{sA}) &= u^s_{\mathbf{X}_{sA}}(x,t) + E(u^s_{\mathbf{X}_{sA^c}}(x,t)) \\
&= u^s_{\mathbf{X}_{sA}}(x,t) + u^s_{E(\mathbf{X}_{sA^c})}(x,t) \\
&= u^s_{\mathbf{X}_{sA}}(x,t) + u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A^c}(\cdot)}(x,t) \\
&= u^s_{\sum_{i=1}^{N_s(A)} \delta(\cdot - X^i_{sA}) + u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A^c}(\cdot)}(x,t) \\
&= u^0_{\eta_0 p_0}(x,t) + \left(u^s_{\mathbf{X}_{sA}}(x,t) - u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A}(\cdot)}(x,t) \right).
\end{aligned}$$

The third line above gives the expression of the intensity mentioned in the corollary. From the fifth line and the linearity of Eq. (6) we get:

$$\begin{aligned}
E(\Lambda(x)) &= u^0_{\eta_0 p_0}(x,t) + \left(E(u^s_{\mathbf{X}_{sA}}(x,t)) - u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A}(\cdot)}(x,t) \right) \\
&= u^0_{\eta_0 p_0}(x,t) + \left(u^s_{E(\mathbf{X}_{sA})}(x,t) - u^s_{u^0_{\eta_0 p_0}(\cdot,s)\mathbb{1}_{A}(\cdot)}(x,t) \right) \\
&= u^0_{\eta_0 p_0}(x,t),
\end{aligned}$$

which is consistent with Lemma 1 that states that $u^0_{\eta_0 p_0}(\cdot,t)$ is the unconditional intensity of $\mathbf{X}_t$. □

A.3. Likelihood function

Our goal is to prove the formula (26). We begin by writing:

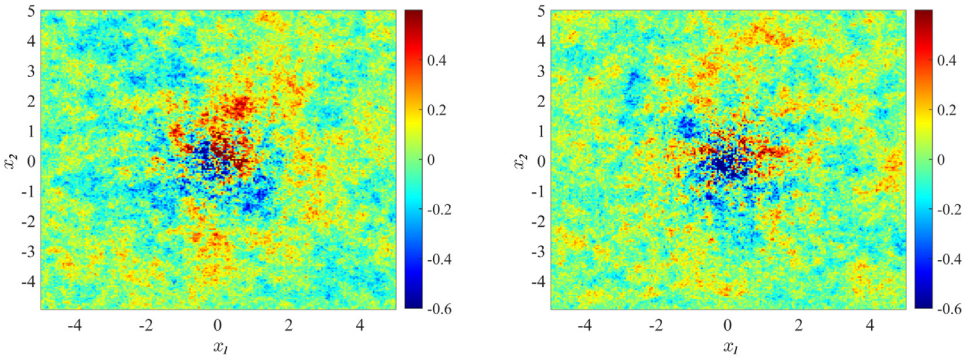
$$\begin{aligned}
\mathcal{L}(\Theta) &= \int_{\tilde{\mathbf{X}}^{(K-1)}} d\mathbb{P}(\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K)}, \tilde{\mathbf{X}}^{(K-1)}) \\
&= \int_{\tilde{\mathbf{X}}^{(K-1)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K-1)}, \tilde{\mathbf{X}}^{(K-1)}) d\mathbb{P}(\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K-1)}, \tilde{\mathbf{X}}^{(K-1)}) \\
&= \int_{\tilde{\mathbf{X}}^{(K-1)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K-2)}, \mathbf{X}_{t_{K-1}}) d\mathbb{P}(\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K-2)}, \mathbf{X}_{t_{K-1}}) \\
&= \int_{\tilde{\mathbf{X}}^{(K-1)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) d\mathbb{P}(\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K-2)}, \mathbf{X}_{t_{K-1}}).
\end{aligned}$$

Then, consider the property $(\mathcal{P}_j)$:

$$\mathcal{L}(\Theta) = \int_{\tilde{\mathbf{X}}^{(K-1)}, \ldots, \tilde{\mathbf{X}}^{(K-1-j)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=K-j}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) d\mathbb{P}(\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K-2-j)}, \mathbf{X}_{t_{K-1-j}}).$$

The above computations show that the property $(\mathcal{P}_0)$ is true (with the convention that the product equals 1 if $j = 0$). Assume that $(\mathcal{P}_j)$ is satisfied for some integer $j \in \{0, \ldots, K-4\}$. Then

$$\begin{aligned}
\mathcal{L}(\Theta) = &\int_{\tilde{\mathbf{X}}^{(K-1)}, \ldots, \tilde{\mathbf{X}}^{(K-2-j)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=K-j}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) \\
&d\mathbb{P}(\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K-2-j)}, \tilde{\mathbf{X}}^{(K-2-j)}, \mathbf{X}_{t_{K-1-j}}),
\end{aligned}$$



(a) Fokker-Planck (b) Fick

Fig. 5. Fokker–Planck vs Fickian diffusion. Solution of the SPDE (7) with dispersal term $\mathcal{D}(t,x,[u]) = \Delta(D(x,t)u)$ (panel a) vs $\mathcal{D}(t,x,[u]) = \mathrm{div}(D(x,t)\nabla u)$ (panel b), reaction term $f(t,x,u) = -\kappa^2\, u(x,t)$ (absorption) and drift term $\mathcal{B}(t,x,[u]) = 0$. The diffusion coefficient is $D(x,t) = D(x) = D_0 + D_1(1 - e^{-\|x\|^2/(2\sigma_D^2)})$, with $D_0 = 10^{-3}$ and $D_1 = 10^{-1}$. In both cases $\kappa = 0.1$, $\sigma = 1$ and $\sigma_D = 2$. The solution is plotted at a fixed time $T = 10$.

$$
\begin{aligned}
&= \int_{\tilde{\mathbf{X}}^{(K-1)},\dots,\tilde{\mathbf{X}}^{(K-2-j)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=K-j}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) \\
&\quad d\mathbb{P}(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(K-3-j)}, \mathbf{X}_{t_{K-2-j}}, \mathbf{X}_{t_{K-1-j}}), \\
&= \int_{\tilde{\mathbf{X}}^{(K-1)},\dots,\tilde{\mathbf{X}}^{(K-2-j)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=K-j}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) \\
&\quad d\mathbb{P}(\mathbf{X}_{t_{K-1-j}} \mid \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(K-3-j)}, \mathbf{X}_{t_{K-2-j}}) d\mathbb{P}(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(K-3-j)}, \mathbf{X}_{t_{K-2-j}}) \\
&= \int_{\tilde{\mathbf{X}}^{(K-1)},\dots,\tilde{\mathbf{X}}^{(K-2-j)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=K-j-1}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) \\
&\quad d\mathbb{P}(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(K-3-j)}, \mathbf{X}_{t_{K-2-j}}).
\end{aligned}
$$

Thus, property $(\mathcal{P}_{j+1})$ is true. By induction, this shows that $(\mathcal{P}_{K-3})$ is true, i.e.:

$$
\mathcal{L}(\Theta) = \int_{\tilde{\mathbf{X}}^{(K-1)},\dots,\tilde{\mathbf{X}}^{(2)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=3}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) d\mathbb{P}(\mathbf{X}^{(1)}, \mathbf{X}_{t_2}).
$$

Finally, integrating the right-hand-side of $(\mathcal{P}_{K-3})$ with respect to $\tilde{\mathbf{X}}^{(1)}$ leads to:

$$
\begin{aligned}
\mathcal{L}(\Theta) &= \int_{\tilde{\mathbf{X}}^{(K-1)},\dots,\tilde{\mathbf{X}}^{(1)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=3}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) d\mathbb{P}(\mathbf{X}^{(1)}, \tilde{\mathbf{X}}^{(1)}, \mathbf{X}_{t_2}) \\
&= \int_{\tilde{\mathbf{X}}^{(K-1)},\dots,\tilde{\mathbf{X}}^{(1)}} \mathbb{P}(\mathbf{X}^{(K)} \mid \mathbf{X}_{t_{K-1}}) \left(\prod_{k=3}^{K-1} d\mathbb{P}(\mathbf{X}_{t_k} \mid \mathbf{X}_{t_{k-1}}) \right) d\mathbb{P}(\mathbf{X}_{t_1}, \mathbf{X}_{t_2})
\end{aligned}
$$

and, by writing $d\mathbb{P}(\mathbf{X}_{t_1}, \mathbf{X}_{t_2}) = d\mathbb{P}(\mathbf{X}_{t_2} \mid \mathbf{X}_{t_1}) d\mathbb{P}(\mathbf{X}_{t_1})$, we get formula (26).

Appendix B. Fokker–Planck and Fick's law

We computed the solution of (7) with a null drift term $\mathcal{B}(t, x, [u]) = 0$ and an absorption term $f(t, x, u) = -\kappa^2 u(x, t)$, depending on the type of dispersal operator $\mathcal{D}$ of either Fokker–Planck or Fickian form. The results in Fig. 5, and other preliminary computations (Matlab codes: https://osf.io/w5utd/) did not show significant differences between the solutions obtained with the two operators. Contrarily to the deterministic case ($\sigma = 0$), Fickian diffusion does not seem to induce a spatial homogenisation. In both cases, the regions with lower diffusion (i.e., where $D(x, t)$ is small) are associated with more contrasted values of the solution. Of course, one may expect differences between the covariance functions associated with these two operators, and some parametrisations may also lead to graphically visible differences, but further theoretical studies are needed to clarify these aspects.

References

Abramowitz, M., Stegun, I.A., 1964. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Vol. 55. US Government printing office.

Alfaro, M., Giletti, Y.-J., Kim, G., Peltier, T., Seo, H., 2021. On the modelling of spatially heterogeneous nonlocal diffusion: deciding factors and preferential position of individuals. arXiv preprint arXiv:2104.00904.

Allard, D., Hristopulos, D.T., Opitz, T., 2021. Linking physics and spatial statistics: A new family of Boltzmann-gibbs random fields. Electron. J. Stat. 15 (2), 4085–4116.

Applebaum, D., 2009. Lévy Processes and Stochastic Calculus. Cambridge university press..

Arnold, L., 1974. Stochastic Differential Equations: Theory and Applications. New York.

Bakka, H., Krainski, D., Bolin, H., Rue, E., Lindgren, F., 2020. The diffusion-based extension of the Matérn field to space–time. arXiv preprint arXiv:2006.04917.

Bjørnstad, O.N., Grenfell, B.T., King, A.A., 2019. Comparison of alternative models of human movement and the spread of disease. bioRxiv.

Bolin, D., Kirchner, K., 2020. The rational SPDE approach for gaussian random fields with general smoothness. J. Comput. Graph. Statist. 29 (2), 274–285.

Bolin, D., Wallin, J., 2020. Multivariate type G Matérn stochastic partial differential equation random fields. J. R. Stat. Soc. Ser. B Stat. Methodol. 82 (1), 215–239.

Boulakia, M., Genadot, A., Thieullen, M., 2015. Simulation of SPDEs for excitable media using finite elements. J. Sci. Comput. 65 (1), 171–195.

Bournaveas, N., Calvez, V., 2010. The one-dimensional Keller–Segel model with fractional diffusion of cells. Nonlinearity 23 (4), 923.

Brown, P.E., Roberts, G.O., Tonellato, S., 2000. Blur-generated non-separable space–time models. J. R. Stat. Soc. Ser. B Stat. Methodol. 62 (4), 847–860.

Carrizo Vergara, R., Allard, D., Desassis, N., 2021. A general framework for SPDE-based stationary random fields. Bernoulli 28 (1), 1–32.

Chiu, S.N., Stoyan, D., Mecke, J., 2013. Stochastic Geometry and Its Applications. John Wiley & Sons.

Cont, R., Tankov, P., 2003. Financial Modelling with Jump Processes. Chapman and Hall/CRC.

Da Prato, G., Zabczyk, J., 2014. Stochastic Equations in Infinite Dimensions, second ed. In: Encyclopedia of Mathematics and its Applications, vol. 152, Cambridge University Press.

Dragon, A.-C., Bar-Hen, A., Monestiez, P., Guinet, C., 2012. Comparative analysis of methods for inferring successful foraging areas from argos and GPS tracking data. Mar. Ecol. Prog. Ser. 452, 253–267.

Einstein, A., 1905. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. Annalen Der Physik 322 (8), 549–560.

Fick, A., 1855. On liquid diffusion. Lond. Edinb. Dublin Philoso. Mag. J. Sci. 10 (63), 30–39.

Fife, P.C., McLeod, J., 1977. The approach of solutions of nonlinear diffusion equations to traveling front solutions. Arch. Ration. Mech. Anal. 65 (1), 335–361.

Fuglstad, G.-A., Lindgren, F., Simpson, D., Rue, H., 2015. Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. Statist. Sinica 11, 5–133.

Gardiner, C., 2009. Stochastic Methods. Springer, Berlin.

Garnier, J., 2011. Accelerating solutions in integro-differential equations. SIAM J. Math. Anal. 43, 1955–1974.

Garnier, J., Roques, L., Hamel, F., 2012. Success rate of a biological invasion in terms of the spatial distribution of the founding population. Bull. Math. Biol. 74, 453–473.

Gelfand, I.M., Shilov, G.E., 1964. Generalized Functions, 4: Applications of Harmonic Analysis. Academic Press.

Hairer, M., Ryser, M., Weber, H., 2012. Triviality of the 2D stochastic Allen–Cahn equation. Electron. J. Probab. 17, 1–14.

Hannunen, S., Ekbom, B., 2001. Host plant influence on movement patterns and subsequent distribution of the polyphagous herbivore Lygus rugulipennis (Heteroptera: Miridae). Environ. Entomol. 30 (3), 517–523.

Hecht, F., 2012. New development in Freefem++. J. Numer. Math. 20 (3–4), 251–265.

Horsthemke, W., Lefever, R., 2006. Noise-induced transitions in physics, chemistry, and biology. In: Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology. pp. 164–200.

Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns, Vol. 70. John Wiley & Sons.
Itô, K., 1954. Stationary random distributions. Mem. Coll. Sci. Univ. Kyoto. Ser. A Math. 28 (3), 209–223.
Kéfi, S., Rietkerk, M., Van Baalen, M., Loreau, M., 2007. Local facilitation, bistability and transitions in arid ecosystems. Theor. Popul. Biol. 71 (3), 367–379.
Keitt, T.H., Lewis, M.A., Holt, R.D., 2001. Allee effects, invasion pinning, and species' borders. Am. Nat. 157, 203–216.
Klein, E.K., Lavigne, C., Foueillassar, X., Gouyon, P.-H., Larédo, C., 2003. Corn pollen dispersal: Quasi-mechanistic models and field experiments. Ecol. Monograph 73, 131–150.
Klein, E.K., Lavigne, C., Gouyon, P.H., 2006. Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. BMC Ecol. 6 (3).
Koh, J., Pimont, F., Dupuy, J.-L., Opitz, T., 2021. Spatiotemporal wildfire modeling through point processes with moderate and extreme marks. arXiv preprint arXiv:2105.08004.
Kohn, R.V., Otto, F., Reznikoff, M.G., Vanden-Eijnden, E., 2007. Action minimization and sharp-interface limits for the stochastic Allen–Cahn equation. Commun. Pure Appl. Math. J. Issued Courant Inst. Math. Sci. 60 (3), 393–438.
Kwaśnick, M., 2017. Ten equivalent definitions of the fractional Laplace operator. Fract. Calc. Appl. Anal. 20 (1), 7–51.
Lang, A., Pereira, M., 2021. Galerkin–Chebyshev approximation of Gaussian random fields on compact Riemannian manifolds. arXiv preprint arXiv:2107.02667.
Lewis, M.A., Kareiva, P., 1993. Allee dynamics and the spread of invading organisms. Theor. Popul. Biol. 43, 141–158.
Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 73 (4), 423–498.
Matano, H., Du, Y., 2010. Convergence and sharp thresholds for propagation in nonlinear diffusion problems. J. Eur. Math. Soc. 12 (2), 279–312.
Møller, J., Waagepetersen, R.P., 2007. Modern statistics for spatial point processes. Scand. J. Stat. 34 (4), 643–684.
Øksendal, B., 2003. Stochastic Differential Equations. Springer.
Opitz, T., Bonneu, F., Gabriel, E., 2020. Point-process based Bayesian modeling of space–time structures of forest fire occurrences in Mediterranean France. Spatial Stat. 40, 100429.
Protter, M.H., Weinberger, H.F., 1967. Maximum Principles in Differential Equations. Prentice-Hall, Englewood Cliffs, NJ.
Prudnikov, A.P., Brychkov, I.A., Marichev, O.I., 1986. Integrals and Series: Special Functions, Vol. 2. CRC Press.
Raissi, M., Perdikaris, P., Karniadakis, G.E., 2017. Physics informed deep learning (part I): Data-driven solutions of nonlinear partial differential equations. arXiv preprint arXiv:1711.10561.
Rietkerk, M., van de Koppel, J., 1997. Alternate stable states and threshold effects in semi-arid grazing systems. Oikos 6, 9–76.
Roques, L., 2013. Modèles de Réaction-Diffusion Pour L'éCologie Spatiale, Quae.
Roques, L., Auger-Rozenberg, M.-A., Roques, A., 2008. Modelling the impact of an invasive insect via reaction–diffusion. Math. Biosci. 216 (1), 47–55.
Roques, L., Bonnefon, O., Baudrot, V., Soubeyrand, S., Berestycki, H., 2020. A parsimonious approach for spatial transmission and heterogeneity in the COVID-19 propagation. Royal Soc. Open Sci. 7 (12), 201382.
Rozanov, J.A., 1977. Markov Random fields and stochastic partial differential equations. Math. USSR-Sbornik 32 (4), 515.
Ryser, M.D., Nigam, N., Tupper, P.F., 2012. On the well-posedness of the stochastic Allen–Cahn equation in two dimensions. J. Comput. Phys. 231 (6), 2537–2550.
Samko, S., 2001. Hypersingular Integrals and their Applications. Taylor & Francis.
Schertzer, D., Larchevêque, M., Duan, J., Yanovsky, V., Lovejoy, S., 2001. Fractional Fokker–Planck equation for nonlinear stochastic differential equations driven by non-Gaussian Lévy stable noises. J. Math. Phys. 42 (1), 200–212.
Sigrist, F., Künsch, H.R., Stahel, W.A., 2015. Stochastic partial differential equation based modelling of large space–time data sets. J. R. Stat. Soc. Ser. B Stat. Methodol. 3–33.
Simpson, D., Illian, J.B., Lindgren, F., Sørbye, S.H., Rue, H., 2016. Going off grid: Computationally efficient inference for log-Gaussian Cox processes. Biometrika 103 (1), 49–70.
Soubeyrand, S., Roques, L., 2014. Parameter estimation for reaction–diffusion models of biological invasions. Popul. Ecol. 56 (2), 427–434.
Soubeyrand, S., Roques, L., Coville, J., Fayard, J., 2011. Patchy patterns due to group dispersal. J. Theoret. Biol. 271, 87–99.
Soubeyrand, S., Sache, I., Hamelin, F., Klein, E.K., 2015. Evolution of dispersal in asexual populations: to be independent, clumped or grouped?. Evol. Ecol. 29, 947–963.
Stinga, P.R., 2019. User's guide to the fractional Laplacian and the method of semigroups. In: Fractional Differential Equations. De Gruyter, pp. 235–266.
Tufto, J., Engen, S., Hindar, K., 1997. Stochastic dispersal processes in plant populations. Theor. Popul. Biol. 52 (1).
Whittle, P., 1954. On stationary processes in the plane. Biometrika 43, 4–449.
Whittle, P., 1963. Stochastic-processes in several dimensions. Bull. Int. Stat. Inst. 40 (2), 974–994.
Wikle, C.K., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84, 1382–1394.
Yuan, Y., Bachl, F.E., Lindgren, F., Borchers, D.L., Illian, J.B., Buckland, S.T., Rue, H., Gerrodette, T., 2017. Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales. Ann. Appl. Stat. 11 (4), 2270–2297.
Zlatoš, A., 2006. Sharp transition between extinction and propagation of reaction. J. Amer. Math. Soc. 19, 251–263.