Modeling intra-host dynamics within the microbiota

Auguste Caen, Béatrice Laroche, Didier Guillemot, Jean-Louis Herrmann, Lulla Opatowski, Elisabeta Vergu

To cite this version:

Auguste Caen, Béatrice Laroche, Didier Guillemot, Jean-Louis Herrmann, Lulla Opatowski, et al.. Modeling intra-host dynamics within the microbiota. Epidemics: 9th International Conference on Infectious Disease Dynamics, Nov 2023, Bologna (ITALY), Italy. . hal-04495219

HAL Id: hal-04495219

https://hal.inrae.fr/hal-04495219

Submitted on 8 Mar 2024

Modeling intra-host dynamics within the microbiota.

Auguste Caen[1], Béatrice Laroche[1], Didier Guillemot[2,3], Jean Louis Herrman[4,5], Lulla Opatowski[2,3], and Elisabeta Vergu[1]

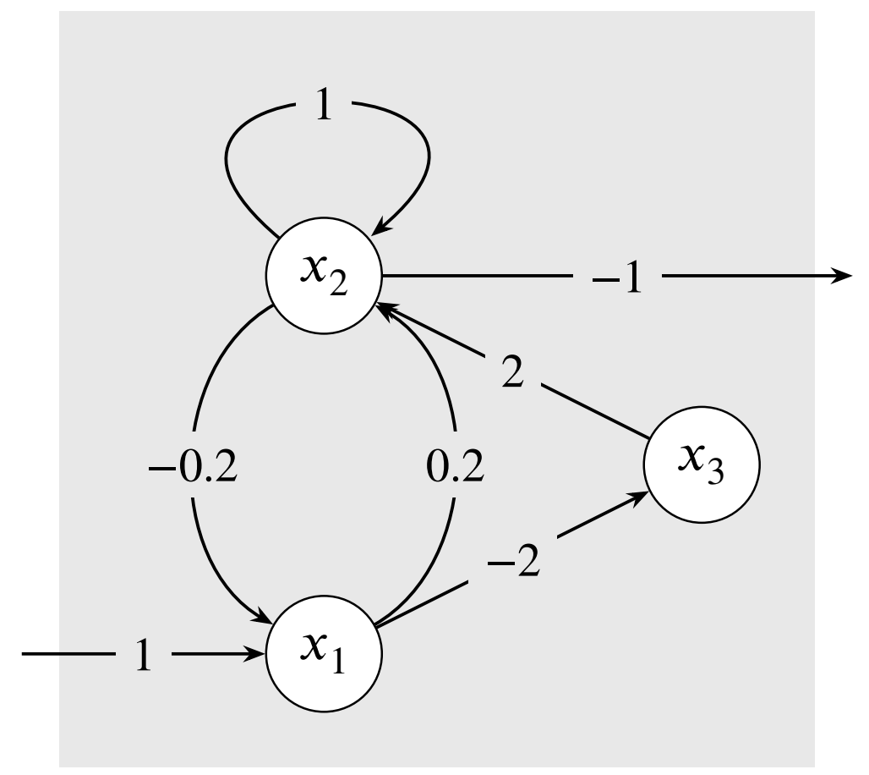
[1]MaIAGE, INRAE, Université Paris-Saclay, 78350 Jouy-en-Josas, France
[2]Institut Pasteur, Université Paris Cité, Epidemiology and Modelling of Antibiotic Evasion (EMAE), Paris, France
[3]Université Paris-Saclay, Université de Versailles Saint-Quentin-en-Yvelines, Inserm U1018, CESP, Anti-infective evasion and pharmacoepidemiology team, Montigny-Le-Bretonneux, France
[4]Université Paris-Saclay, Université de Versailles Saint-Quentin-en-Yvelines, Inserm U1173, Infection and Inflammation, Montigny-le-Bretonneux, France
[5]AP-HP, Groupe Hospitalo-Universitaire Paris-Saclay, Hôpital Raymond Poincaré, Garches, France

Abstract

Microbiota is a major reservoir of bacteria in the humans and animals organism. It is home to numerous commensal species, some of which can be sources of infection, such as staphylococcus aureus[1]. While the composition and properties of microbiota are increasingly understood, their dynamics remain difficult to model, due to the large number of species and their interactions. **Generalized Lotka Volterra** (LV) model is particularly interesting, since it allows simulations of a large number of interacting microbial populations. However, calibrating this model requires abundance data, while classical metagenomic analyses, which quantify the composition of microbiota, only provide **"frequency" data**, i.e. the proportion of each population among those present. Currently, to address this issue, either imprecise proxies of total microbiota abundance are used [2], or strong assumptions are made about the system, for example by assuming that total abundance is fixed [3]. Applying this model to microbiota data without using such assumptions is therefore a key challenge:

- We **characterize analytically** the **identifiability** conditions of the LV model on frequency data.
- We demonstrate analytically that such **identifiability is possible in the general case**, without requiring strong assumptions.
- We validate this result by **numerical analysis of simulations** of microbiotic dynamics.

The Lotka-Volterra model



N populations

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Environment impact:

$$\boldsymbol{\mu} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Interaction matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & -0.2 & 0 \\ 0.2 & 1 & 2 \\ -2 & 0 & 0 \end{bmatrix}$$

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = x_i(t)\left(\boldsymbol{\mu}_i + \sum_{j=1}^{3} \mathbf{A}_{i,j} x_j\right)$$

The identifiability problem

Is it possible to calibrate the Lotka-Volterra model using frequency data?

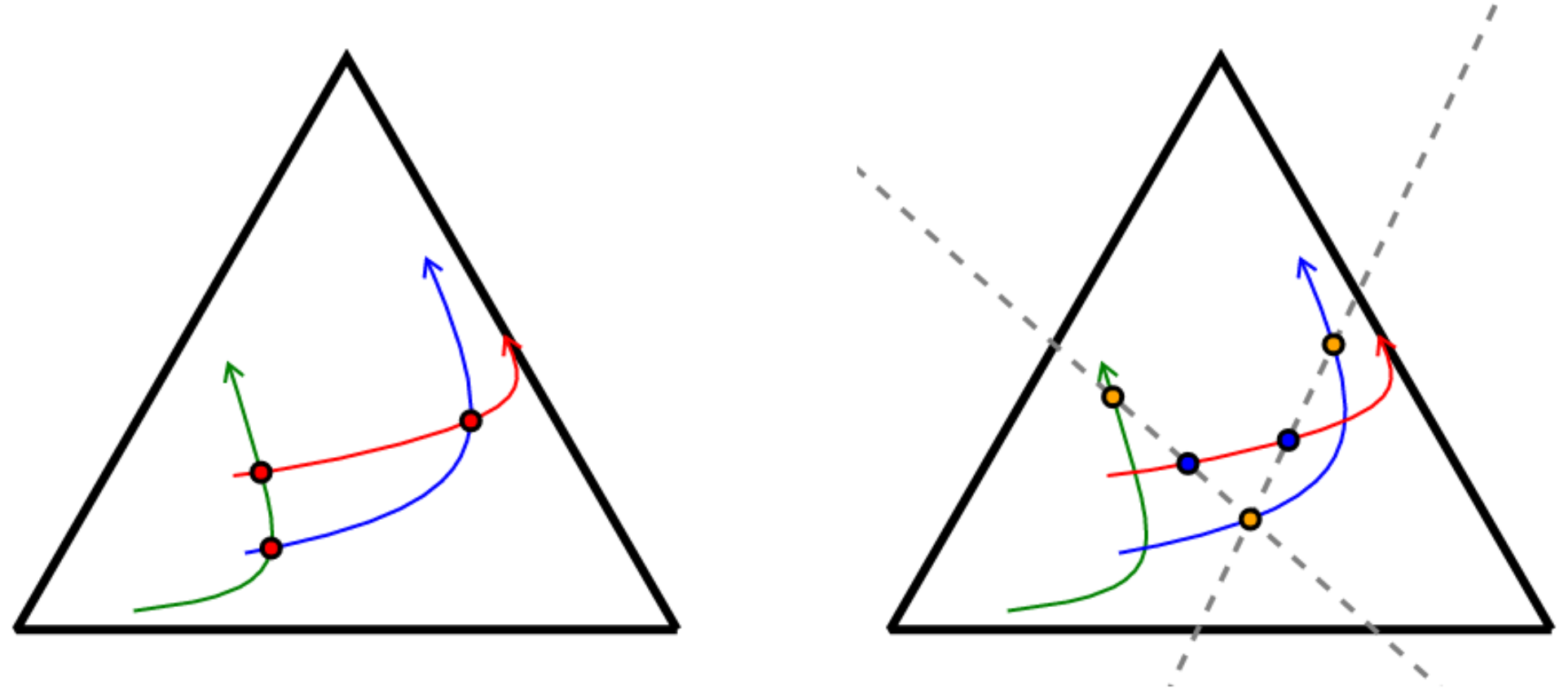
$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = x_i(t)\left(\boldsymbol{\mu}_i + \sum_{j=1}^{N} \mathbf{A}_{i,j} x_j\right) \qquad \frac{\mathrm{d}x'_i(t)}{\mathrm{d}t} = x'_i(t)\left(\boldsymbol{\mu}'_i + \sum_{j=1}^{N} \mathbf{A}'_{i,j} x'_j\right)$$

$$\frac{1}{\sum_{j=1}^{N} x_j(t)} x(t) = \frac{1}{\sum_{j=1}^{N} x'_j(t)} x'(t) \overset{?}{\Rightarrow} \boldsymbol{\mu}' = \boldsymbol{\mu} \;\&\; \mathbf{A}' = \mathbf{A}$$

Identifiability of the generalised Lotka-Volterra model

Theorem:

If the frequency data is rich enough:



And

$$\frac{1}{\sum_{j=1}^{N} x_j(t)} x(t) = \frac{1}{\sum_{j=1}^{N} x'_j(t)} x'(t)$$

Then

1) either $rank(\mathbf{A}) > 2$

$\Rightarrow \boldsymbol{\mu}' = \boldsymbol{\mu} \;\&\; \mathbf{A}' = \rho \mathbf{A}$ with $\rho > 0$

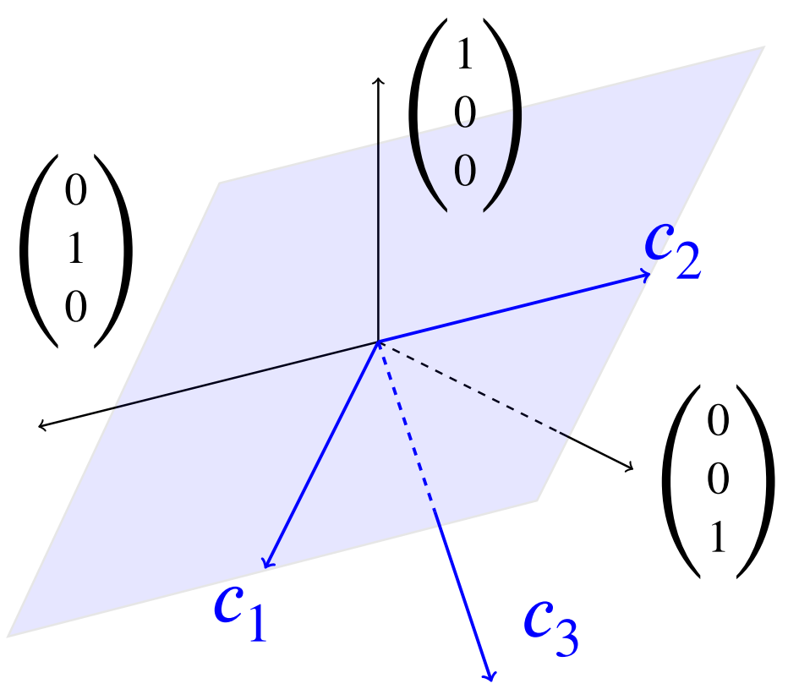
Identifiability !

2) or $rank(\mathbf{A}) \leq 2$

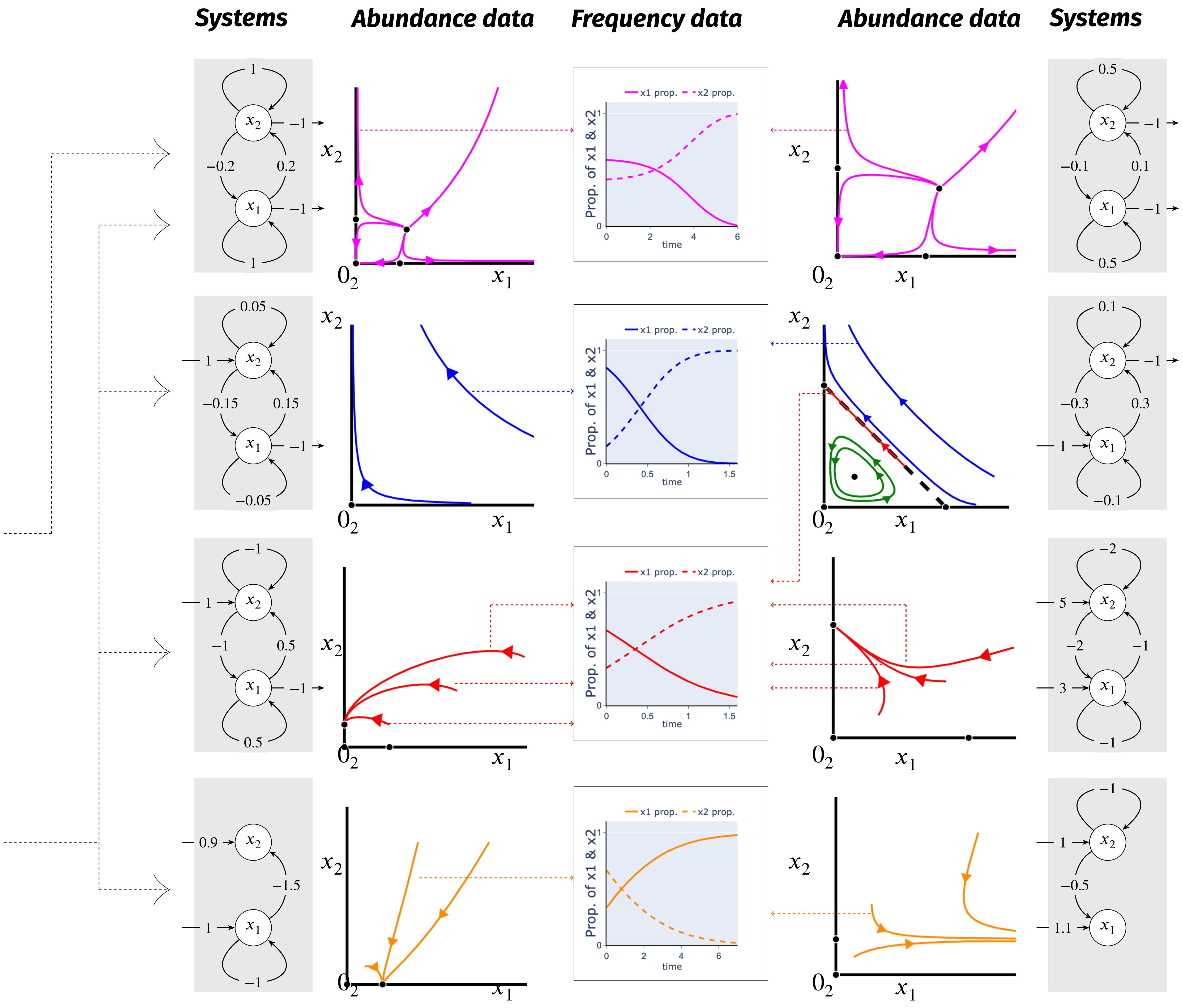
If $n \geq 3$:

$rank(\mathbf{A}) \leq 2 \iff c_1$, c_2 and c_3 are in the same plane

$$\begin{array}{cc} & \begin{array}{ccc} c_1 & c_2 & c_3 \\ \downarrow & \downarrow & \downarrow \end{array} \\ \mathbf{A} = & \begin{pmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \mathbf{A}_{1,3} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \mathbf{A}_{2,3} \\ \mathbf{A}_{3,1} & \mathbf{A}_{3,2} & \mathbf{A}_{3,3} \end{pmatrix} \end{array}$$



$$\Rightarrow \mathbb{P}\left(rank(\mathbf{A}) \leq 2\right) = 0$$



Conclusion and perspectives

- It is possible to **reconstruct** the **total abundance** of a microbiota **using only its frequency data**.
- It is possible to use the **Lotka-Volterra model** on frequency data of the **microbiota**.

Acknowledgements

This work is supported by an INCEPTION INTERDISCIPLINARY PROJECT (P2I) funding: MICROMOD – Modelling the role of microbiota in the acquisition and transmission of antibiotic resistant pathogenic bacteria

References

[1] Krismer, Bernhard, et al. "The commensal lifestyle of Staphylococcus aureus and its interactions with the nasal microbiota." Nature reviews microbiology 15.11 (2017): 675-687.
[2] Gloor, Gregory B., et al. "Microbiome datasets are compositional: and this is not optional." Frontiers in microbiology 8 (2017): 2224.
[3] Gao, Xuefeng, et al. "Inference of significant microbial interactions from longitudinal metagenomics data." Frontiers in microbiology 9 (2018): 2319.

oster