

Sparse species interactions reproduce abundance correlation patterns in microbial communities

Jose Camacho-Mateu¹, Aniello Lampo¹, Matteo Sireci^{2,3}, MA Muñoz^{2,3} and JA Cuesta^{1,4}

¹Grupo Interdisciplinar de Sistemas Complejos (GISC), Departamento de Matemáticas, Universidad Carlos III de Madrid, Leganés, Spain

²Departamento de Electromagnetismo y Física de la Materia, Universidad de Granada, Granada, Spain

³Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, Granada, Spain

⁴Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, Zaragoza, Spain

The study of the microscopic living world has been revolutionized by metagenomics, which revealed a vast microbial diversity not captured by traditional laboratory cultures. This, coupled with the ability to collect genetic material directly from natural environments, has led to the generation of massive datasets for studying microbial communities at a larger scale and with more detail than ever before.

To analyze this wealth of data, macroecology has emerged as a promising approach for understanding broad-scale patterns in microbial communities. Statistical analysis has unveiled distinct distributions of abundance and diversity, which can often be explained by fundamental ecological forces. Recent advancements have shed light on statistical regularities in temporal and spatial variations of taxa populations, as well as variations in species numbers.

In a significant step forward, Jacopo Grilli [1] characterized abundance variation patterns in nine different biomes and identified three macroecological laws: the gamma distribution of abundance fluctuations across samples, Taylor's law linking variances to means, and the lognormal distribution of mean abundances across species. However, the current approach of treating species as independent entities overlooks the existence of correlations between abundance fluctuations of species pairs, as revealed by empirical data.

The lack of attention given to abundance correlations in empirical studies reflects a weakness in current population models. The Stochastic Logistic Model (SLM), a well-known theoretical benchmark, treats the temporal evolution of microbial abundances independently. While the SLM captures the three macroecological laws and reproduces single-species abundance time series, it is unable to predict the distribution of abundance correlations between pairs of species.

In our research [2], we specifically address the challenge of reproducing correlation distributions in microbial communities by incorporating species interactions into population models. This approach is motivated by experimental studies that have consistently shown the widespread presence and significance of microbial interactions in maintaining community stability, influencing human health, and controlling medical disorders. The inclusion of interactions in population models is also supported by network inference approaches, which aim to construct species-coupling networks based on empirical observations. There is a growing consensus in the field about the crucial role of interactions in microbial ecosystems.

Interactions between species can be implemented in different ways. One approach is to assume that the various environmental noise terms are correlated with each other, namely, that fluctuations in environmental conditions affect different species in a similar or opposite manner. Another approach is to directly introduce a coupling between species abundances, inspired by the classical work of Lotka and

Volterra in ecology. This approach considers the interactions among species, such as competition, predation, mutualism, or other types of ecological relationships, as explicit factors influencing the dynamics of species abundances. An important result of the current work is to recognize this latter approach as successful.

Along this line we propose a dynamical model of populations - the Stochastic Lotka-Volterra model (SLVM) - combining a multiplicative environmental noise with direct Lotka-Volterra interactions. We design a Monte Carlo Markov Chain (MCMC) algorithm to obtain sets of interaction constants able to reproduce the experimental correlation distributions much better than the state-of-the-art attempts. Importantly, the model still reproduces single-species, experimental, macroecological patterns previously detected in [1], concerning the abundance fluctuations across both species and communities. Endorsed by the agreement with the observed phenomenology, our analysis provides insights on the properties of microbial interactions, and suggests their sparsity as a necessary feature to balance the emergence of different patterns.

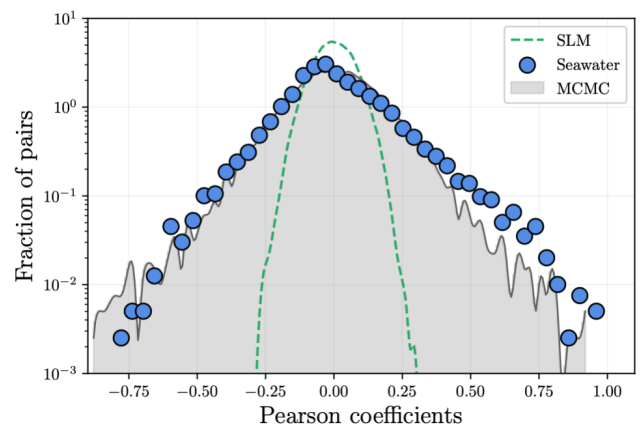


Fig. 1. Abundance correlation distributions for real and simulated communities. Colored bullets correspond to real *Seawater* biome. Green dashed line portrays the distribution of Pearson's coefficients for the abundance correlation of all pairs of species resulting from the SLM. Gray curves show the same distributions as obtained from the SLVM, with the Lotka-Volterra interaction constants inferred using the MCMC algorithm.

[1] J. Grilli, *Macroecological laws describe variation and diversity in microbial communities*, *Nature Communications* **11**, 4743 (2020).

[2] J. Camacho, A. Lampo, M. Sireci, M.A. Muñoz, J.A. Cuesta, *Species interactions reproduce abundance correlation patterns in microbial communities*, <https://doi.org/10.48550/arXiv.2305.19154> (2023).