



Universiteit
Leiden

The Netherlands

Ecology meets human health: studies on human gut microbiota in health and disease

Pinto, S.

Citation

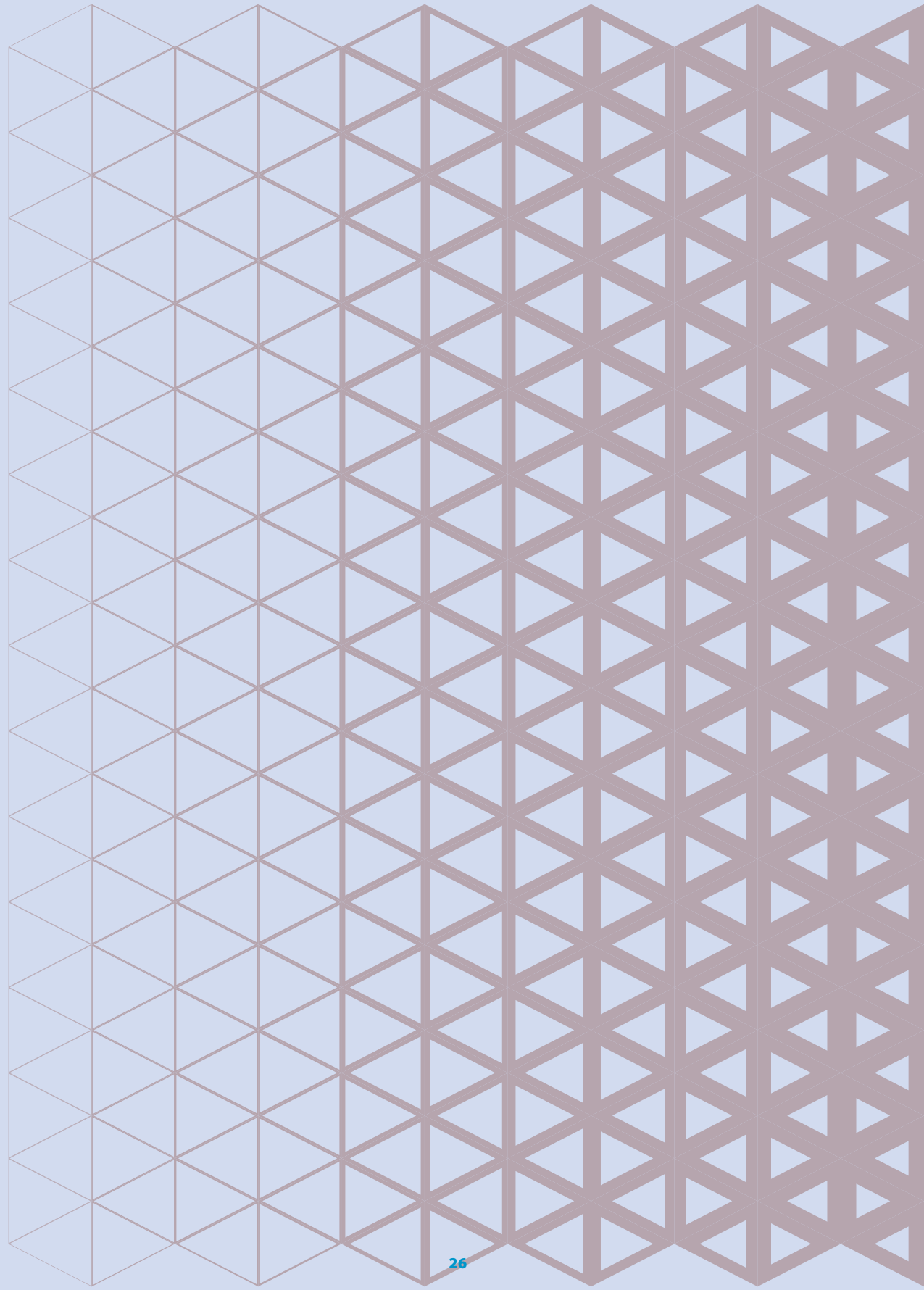
Pinto, S. (2025, November 20). *Ecology meets human health: studies on human gut microbiota in health and disease*. Retrieved from <https://hdl.handle.net/1887/4283645>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

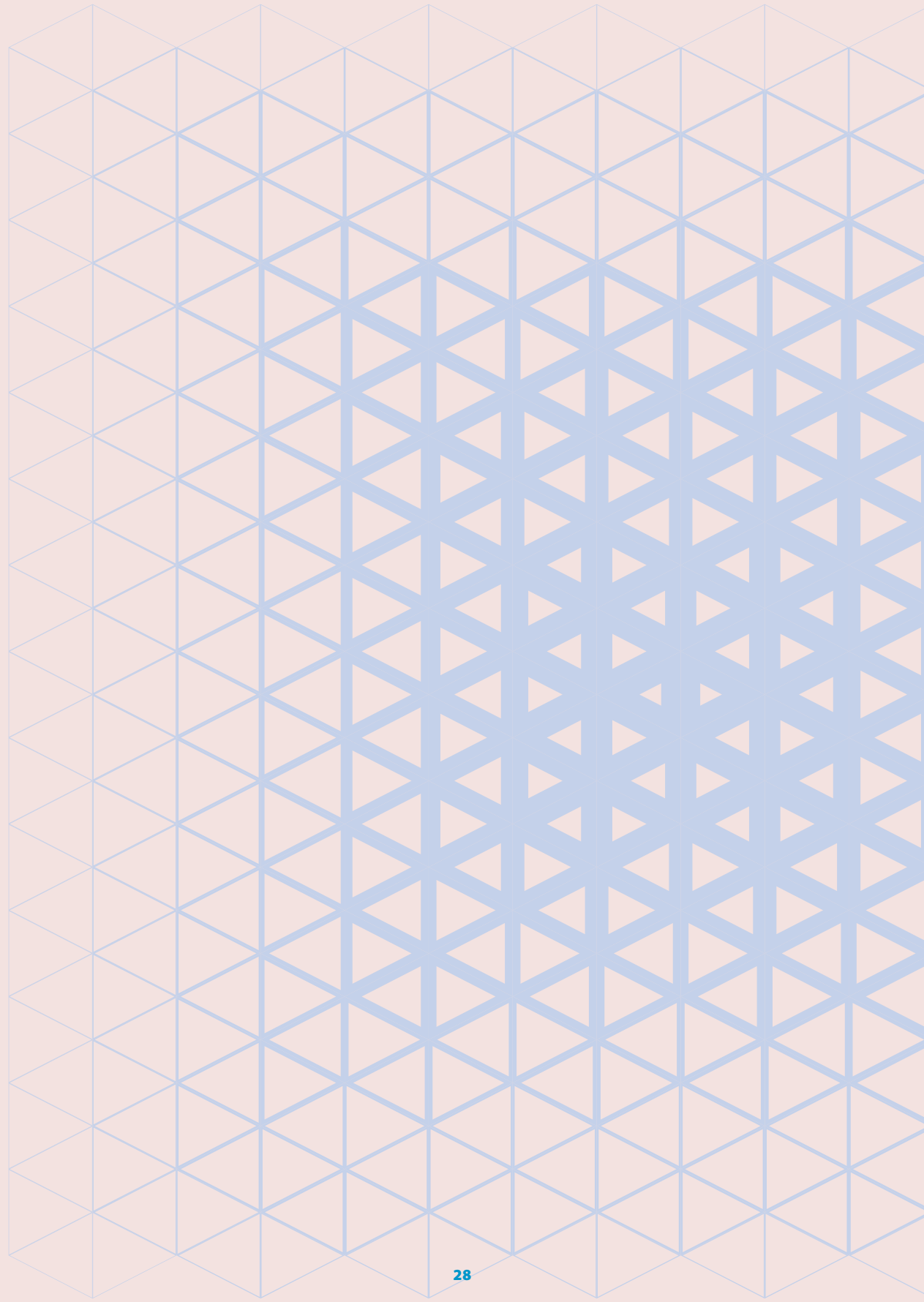
Downloaded from: <https://hdl.handle.net/1887/4283645>

Note: To cite this publication please use the final published version (if applicable).



Part I

**Ecological structure in
the human gut microbiota**



Chapter 2

Correlations carry limited information about network interactions

Susanne Pinto¹, Elisa Benincà², Egbert H. van Nes³, Marten Scheffer³, Johannes A. Bogaards⁴

- 1 Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands
- 2 Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
- 3 Department of Environmental Sciences, Wageningen University and Research, Wageningen, the Netherlands
- 4 Department of Epidemiology and Data Science, Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

PLOS Comput Biol. 2022 Sep 9;18(9):e1010491. doi: 10.1371/journal.pcbi.1010491.

Species abundance correlations carry limited information about microbial network interactions

Abstract

Unraveling the network of interactions in ecological communities is a daunting task. Common methods to infer interspecific interactions from cross-sectional data are based on co-occurrence measures. For instance, interactions in the human microbiome are often inferred from correlations between the abundances of bacterial phylogenetic groups across subjects. We tested whether such correlation-based methods are indeed reliable for inferring interaction networks. For this purpose, we simulated bacterial communities by means of the generalized Lotka-Volterra model, with variation in model parameters representing variability among hosts. Our results show that correlations can be indicative of the presence of bacterial interactions, but only when measurement noise is low relative to the variation in interaction strengths between hosts. Indication of interaction was affected by type of interaction network, process noise, and sampling under non-equilibrium conditions. The sign of a correlation mostly coincided with the nature of the strongest pairwise interaction, but this is not necessarily the case. For instance, under rare conditions of identical interaction strength, we found that competitive and exploitative interactions can result in positive as well as negative correlations. Thus, cross-sectional abundance data carry limited information on specific interaction types. Correlations in abundance may hint at interactions but require independent validation.

Introduction

The human body harbors an exceptional bacterial diversity.²¹ The composition of these bacterial communities is generally shaped by characteristics of the host and by the ecological dependencies among bacterial species themselves.^{8, 13, 245} These dependencies often occur through competitive or synergistic interactions, which may lead to a (mutual) decrease or increase in the abundance of interacting species.¹²³ For instance, it is known that bacteria can interact with each other through excreted metabolites, which can function as an antimicrobial or as a food source.^{8, 246} Among other mechanisms, for example negative interactions take place when toxic compounds produced by one species harm other bacteria, whereas positive interactions occur when bacteria feed on the nutrients that are produced by others. Besides, many different forms of interactions exist, depending on the effects experienced by the species involved. Knowledge of interspecific interactions in the human microbiome is paramount to understand ecological processes and compositional changes in relation to health and disease.^{49, 190}

Most human microbiome studies are limited to only a few samples in time, presenting mere ‘snapshots’ of the microbial ecosystem, even if these samples are derived from hundreds of human hosts.

A common way to infer microbial networks from such cross-sectional data is by quantifying co-occurrence, e.g., through (partial) correlations, between bacterial phylogenetic groups. Several different conclusions have been derived from such endeavors, for example on species associations that reflect shared or overlapping niche preferences,²⁴⁷ microbial community structure,^{248, 249} the resilience of microbial communities to perturbations,¹⁰⁰ and keystone species in microbial networks.¹¹⁴ Currently there are several correlation-based network tools available that can deal with the difficulties of microbiome data, such as compositionality.^{230, 250, 251} The potential of correlation-based approaches for uncovering microbial networks has been highlighted in previous research.²⁵²

Whether correlation-based networks represent meaningful ecological structure in microbial communities is, however, debated. Carr et al. (2019) showed that spurious correlations may occur due to the use of sequencing methods, data transformations, and the large number of unmeasured variables.²¹⁶ Berry & Widder (2014) and Hirano & Takemoto (2019) assessed the performance of different co-occurrence methods for inferring interaction structure and found that their performance strongly depends on the underlying network properties, such as network size and density, and the number of samples used to construct the network.^{114, 229} Apart from the challenges of metagenomic-based abundance data and disagreement between various network tools, here we question whether correlations themselves are at all useful to distinguish between different ecological interaction types. Resource competition and metabolic cooperation have been successfully inferred within environmental microbiomes, by linking ecological distribution data to multi-species metabolic models and subsequent verification of putative interactions by means of experimental co-growth analysis.¹¹⁹ However, host-associated microbiomes often include non-culturable organisms, without information on nutrient requirements or metabolic function. Likewise, performance of correlation analysis in relation to alternative interaction types in the human microbiota is not well understood and deserves further investigation.

Correspondence of correlations with ecological interactions needs to be studied against a known ground truth, which can be achieved by means of simulation. Mathematical models have been used as ground truth in assessment of correlation network techniques before,²⁵³ but correlation networks have not been systematically investigated against distinct interaction types in dynamic models. This requires elucidation especially as the 'true' ecological networks governing microbiome dynamics are still unknown. For this purpose, we assessed the performance of correlation-based network reconstruction by simulating abundance data based on the generalized Lotka-Volterra (gLV) model. The gLV model describes the collective dynamics of multiple species by means of an interaction matrix that can modulate different types of interactions.²³² The model is commonly used in microbiome studies for different aims: to simulate microbial communities under various interaction structures,²³² to infer interaction structure from time series data,¹⁰⁰ to forecast population dynamics after a perturbation,²⁵⁴ to infer the network topology from steady state samples,²⁵⁵ and to identify the efficiency of intervention protocols in altering the state of a system via the addition or subtraction of microbial species.²⁵⁶ In ecology, gLV-type models have been questioned for their reliance on pairwise additive interactions, as well as for the strictly linear effects imposed on interspecific interactions. Nonetheless, from the perspective of network inference, it makes sense to first investigate gLV-type models, as their first-order description of ecological dependencies, specified through a pairwise interaction matrix, resembles the objective of correlation analysis and most network models.⁸

In addressing how gLV-type interactions can be inferred from cross-sectional data, we mainly focus on the correspondence between the obtained correlation-based networks and the underlying network of ecological interactions. We specifically investigate how the inference of microbial interaction types is enabled by interindividual variation in population-dynamic parameters (e.g., species-specific carrying capacities, intrinsic growth rates, and strength of interspecific interactions) and how network reconstruction is affected by gLV model assumptions. We highlight several situations where correlations cannot distinguish microbial interaction types and therefore recommend careful interpretation and validation when inferring networks from cross-sectional abundance data.

Methods

Two-species Lotka-Volterra model with self-limitation

First, we investigated how interactions between two species of microbial populations are displayed in terms of correlations of abundances in the Lotka-Volterra model. For the sake of convenience, we use the term ‘species’, although in studies with real microbiota data it is often not possible to characterize the taxonomic abundances at species-level and therefore genera or higher taxonomic levels are often used instead.

The two-species Lotka-Volterra model is given by the following set of ordinary differential equations:

$$\frac{dN_1}{dt} = r_1 N_1 (1 - K_1^{-1} N_1 + \alpha_{12} N_2) \quad \text{Eq. 2.1}$$

$$\frac{dN_2}{dt} = r_2 N_2 (1 - K_2^{-1} N_2 + \alpha_{21} N_1) \quad \text{Eq. 2.2}$$

Here, N_i is the abundance of either species 1 or species 2 (with $i = 1$ or $i = 2$). The term r_i is the intrinsic growth rate of each species, here normalized to 1 and 2 per time unit for species 1 and 2, respectively. The effect of each species’ abundance on its own growth is defined in terms of the species-specific carrying capacities K_i , with $\alpha_{ii} = -K_i^{-1}$ denoting intraspecific competition. We arbitrarily chose the carrying capacity for the first species to be higher than the carrying capacity for the second species ($K_1 = 1.5$; $K_2 = 1.1$), meaning intraspecific competition is less strong for species 1 compared to species 2. Furthermore, α_{ij} ($i = 1, 2$; $j = 1, 2$; $i \neq j$) indicates the interspecific interactions (the effect of one species abundance on the growth of the other species). A positive α_{ij} (e.g., as in the case of mutualism) denotes a positive effect of species j on the growth of species i , a negative α_{ij} (e.g., as in the case of competition) means a negative effect of species j on the growth of species i (Appendix Figure 2.1). We assessed the effect of variation in the interspecific interaction parameters on correlation in equilibrium abundance between both species. For this purpose, the interspecific interaction strengths (α_{12} and α_{21}) were drawn randomly from two normal distributions with similar or different mean and similar or different standard deviations (σ_α). Moreover, we also investigated the situation where $|\alpha_{12}| = |\alpha_{21}|$. Note that it was not possible to achieve stable coexistence for every combination of α_{12} and α_{21} . More information on the conditions for coexistence can be found in Box 2.1.

Box 2.1 - Coexistence in a two-species Lotka-Volterra model with self-limitation.

The conditions for coexistence in the two-species Lotka-Volterra model with self-limitation can be derived by setting both growth equations to zero and investigating what parameter combinations yield $\bar{N}_1 > 0$ as well as $\bar{N}_2 > 0$. Here, $\bar{N}_1$ denotes the equilibrium abundance of species 1 and $\bar{N}_2$ denotes the equilibrium abundance of species 2. Writing these conditions in terms of N_1 as functions of N_2 gives the following:

$$\frac{dN_1}{dt} = 0 \wedge N_1 > 0 \rightarrow N_1 = f_1(N_2) = K_1 + \alpha_{12}K_1N_2 \quad \text{Eq. 2.A}$$

$$\frac{dN_2}{dt} = 0 \wedge N_2 > 0 \rightarrow N_1 = f_2(N_2) = -\frac{1}{\alpha_{21}} + \frac{1}{(\alpha_{21}K_2)}N_2 \quad \text{Eq. 2.B}$$

The joint equilibrium abundance of both species ($\bar{N}_1, \bar{N}_2$) is determined by $f_1(N_2) = f_2(N_2)$. Equation 2.A shows that species 1 grows to its carrying capacity K_1 in the absence of interspecific interactions, i.e., if $\alpha_{12} = 0$. Likewise, $\alpha_{12} > 0$ allows species 1 to grow to higher abundance in the presence of species 2 than determined by its own carrying capacity, whereas $\alpha_{12} < 0$ leads to a reduced abundance of species 1 in the presence of species 2. Similar relations hold for the abundance of species 2 in the presence of species 1, depending on α_{21} . From Equation 2.B, it can also be derived that $\bar{N}_1 > 0$ is only compatible with $\bar{N}_2$ being above its carrying capacity K_2 if at the same time $\alpha_{21} > 0$, whereas $\bar{N}_2$ being below K_2 requires $\alpha_{21} < 0$.

Joint inspection of Equations 2.A and 2.B also establishes the following, more subtle, conditions for coexistence:

- 1 If $\alpha_{12} > 0$ and $\alpha_{21} > 0$, e.g., in case of mutualism, f_2 has a negative intercept in the Cartesian (N_2, N_1) coordinate system (Figure 2.A - panel A). As both functions have a positive slope in this situation, and f_1 always has a positive intercept, f_2 must have a stronger slope than f_1 for both to intersect in the positive quadrant. This boils down to $\frac{1}{\alpha_{21}K_2} > \alpha_{12}K_1$, or equivalently $\alpha_{21}\alpha_{12} < \alpha_{11}\alpha_{22}$, as $K_i = -\frac{1}{\alpha_i}$ by definition. This means that the product of interspecific mutualism needs to be smaller than the product of intraspecific competition for both species to co-exist, otherwise there is no control of population growth.
- 2 If $\alpha_{12} < 0$ and $\alpha_{21} < 0$, e.g., in case of competition, both functions have positive intercept and negative slope (Figure 2.A - panel B). Intersection in the positive quadrant requires the function with the larger intercept to intersect the abscissa, i.e., the N_2 axis where $N_1 = 0$, at a smaller value than the function with the smaller intercept. Thus, this requires $|\alpha_{21}| > \frac{1}{K_1}$ and $|\alpha_{12}| > \frac{1}{K_2}$, with f_2 having the larger intercept, or alternatively, $|\alpha_{21}| < \frac{1}{K_1}$ and $|\alpha_{12}| < \frac{1}{K_2}$, with f_2 having the larger intercept. In the first instance, interspecific competition is stronger than intraspecific competition, whereas in the second instance, interspecific competition is less strong than intraspecific competition. It turns out that only the last of these conditions yields a stable equilibrium, meaning that the abundances of both species return to equilibrium after small displacements.

- 3 If $a_{12} < 0$ and $a_{21} > 0$, e.g., in case of exploitation of species 1 by species 2, f_1 has a positive intercept and negative slope, whereas f_2 still has a negative intercept and positive slope (Figure 2.A - panel C). Intersection in the positive quadrant requires f_1 to intersect the abscissa at a larger value than K_2 , the point where f_2 intersects the abscissa. The condition for coexistence thus becomes $|a_{12}| < \frac{1}{K_2}$, or equivalently $a_{12} < a_{22}$, meaning that the parasite should exert stronger inhibitory effect on its own growth than on that of the exploited species.
- 4 Conversely, in case of exploitation of species 2 by species 1, i.e., if $a_{12} > 0$ and $a_{21} < 0$, both f_1 and f_2 have a positive intercept, but f_1 now has a positive slope whereas f_2 has a negative slope (Figure 2.A - panel D). Intersection in the positive quadrant then requires f_1 to have a smaller intercept than f_2 . The condition for coexistence thus becomes $|a_{21}| < \frac{1}{K_1}$, or equivalently $a_{21} < a_{11}$, again meaning that the parasite should exert stronger inhibitory effect on its own growth than on that of the exploited species.

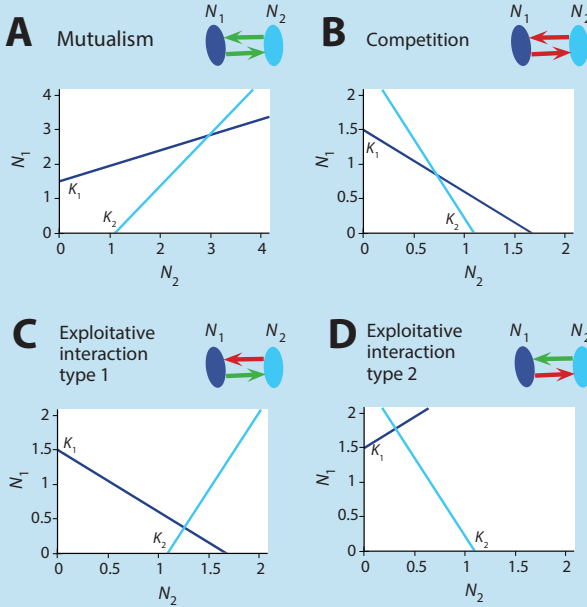


Figure 2.A - Zero-growth isoclines ('null-clines') in the two-species Lotka-Volterra model. Visualization of the effect of species 1 and 2 abundances on each other in the Cartesian (N_2, N_1) coordinate system. Here, f_1 denotes the isocline of zero growth of species 1, i.e., $f_1(N_1)$ (in dark blue), and f_2 denotes the isocline of zero growth of species 2, i.e., $f_2(N_1)$ (in light blue). Their point of intersection represents the joint equilibrium abundance of both species, i.e., $(\bar{N}_1, \bar{N}_2)$. Throughout $K_1 = 1.5$ and $K_2 = 1.1$. Parameters for the various scenarios: $a_{12} = 0.3$ and $a_{21} = 0.6$ under mutualism; $a_{12} = -0.6$ and $a_{21} = -0.4$ under competition; $a_{12} = -0.6$ and $a_{21} = 0.4$ under exploitative interaction type 1; and $a_{12} = 0.6$ and $a_{21} = -0.4$ under exploitative interaction type 2.

The additional requirement for stable coexistence is that the two-species system should be locally stable around the equilibria $(\bar{N}_1, \bar{N}_2)$, which can be formalized in terms of the Jacobian matrix of the Lotka-Volterra model evaluated at $(\bar{N}_1, \bar{N}_2)$. This amounts to determining the trace and determinant of the matrix of the partial derivatives of the growth equations regarding either species, i.e.,

$$\begin{bmatrix} r_1 - 2r_1\bar{N}_1/K_1 + r_1a_{12}\bar{N}_2 & r_1a_{12}\bar{N}_1 \\ r_2a_{21}\bar{N}_2 & r_2 - 2r_2\bar{N}_2/K_2 + r_2a_{21}\bar{N}_1 \end{bmatrix} \quad \text{Eq. 2.C}$$

It can be verified that the conditions for coexistence stated under mutualism and exploitative interactions yield equilibria that are locally stable, just as the last of the conditions under competition. We will not derive these conditions here, as these are covered by textbooks on theoretical ecology.²⁵⁷

In summary, the two-species Lotka-Volterra model with self-limitation has the following possibilities for stable coexistence (Table 2.A):

Table 2.A - Conditions for stable coexistence in the two-species Lotka-Volterra model.		
Type of interaction	Condition	Outcome
Mutualism $a_{12} > 0 \wedge a_{21} > 0$	$a_{12}a_{21} < \frac{1}{(K_1K_2)}$	$\bar{N}_1 > K_1 \wedge \bar{N}_2 > K_2$
Competition $a_{12} < 0 \wedge a_{21} < 0$	$ a_{12} < \frac{1}{K_2} \wedge a_{21} < \frac{1}{K_1}$	$\bar{N}_1 < K_1 \wedge \bar{N}_2 < K_2$
Exploitative interaction type 1 ^a $a_{12} < 0 \wedge a_{21} > 0$	$ a_{12} < \frac{1}{K_2}$	$\bar{N}_1 < K_1 \wedge \bar{N}_2 > K_2$
Exploitative interaction type 2 ^b $a_{12} > 0 \wedge a_{21} < 0$	$ a_{21} < \frac{1}{K_1}$	$\bar{N}_1 > K_1 \wedge \bar{N}_2 < K_2$

^a Exploitative interaction type 1: species 1 is being exploited by species 2

^b Exploitative interaction type 2: species 2 is being exploited by species 1

The condition for stable coexistence of competitors requires both species to have less effect on the growth of the other species than on itself. In case of an unstable equilibrium, either species will eventually outcompete the other; the species with an initial advantage will drive the other species to extinction, a condition referred to as competitive exclusion.^{258, 259} This will occur, for instance, when each species produces a substance which is toxic to the other species but relatively harmless to itself.

Generalized host-specific Lotka-Volterra model

Microbial abundance is not only shaped by intra- and interspecific interactions, but also by host characteristics, for example lifestyle, diet, and age.²⁶⁰ Therefore, we investigated the performance of correlation-based network inference of microbial networks for a host-specific version of the gLV model. The host-specific gLV model is given by:

$$\frac{dN_{i,m}}{dt} = r_{i,m} N_{i,m} \left(1 - K_{i,m}^{-1} N_{i,m} + \sum_{\substack{j=1 \\ j \neq i}}^S a_{ij,m} \frac{N_{j,m}}{K_{j,m}} \right) \quad \text{Eq. 2.3}$$

Here, $N_{i,m}$ is the abundance of each species i in host m , with $i = 1, \dots, s$ (s being the total number of bacterial species) and $m = 1, \dots, 300$ (the total number of hosts). The terms $r_{i,m}$ and $K_{i,m}$ are the intrinsic growth rates and the carrying capacities of each species i in host m . The carrying capacities are kept separated from the interaction matrix A which only contains interspecific interactions (namely, the pairwise terms a_{ij}), facilitating a one-to-one comparison with the correlation matrix.

Parameterization of the base case simulations

We started with a base case, and we added step by step variation to this case. Note that the base case parametrization does not reflect any particular real-world system. Rather, parameters were chosen in such a way to facilitate computation and promote coexistence among species. Variations to the base case parameters are shown later on, but also here, findings should be appreciated from a qualitative rather than quantitative viewpoint. In the base case, the number of bacteria equals ten. The species-specific growth rate r_i and the species-specific carrying capacity K_i were randomly drawn from uniform distributions, respectively $U(0.05, 0.1)$ and $U(0, 1)$. The density of the interaction matrix A in the base case was chosen such that both sparsity of the interaction network and coexistence of the species was promoted in all simulations; in the base case, density was $\frac{1}{4}$ meaning that three out of four possible interactions were set to zero. Moreover, to ensure coexistence between species in the model we chose stronger intraspecific interactions than pairwise interspecific interactions. The species-specific parameters a_{ij} were drawn from a Gaussian mixture distribution, as follows. Half of the interactions were drawn from a negative normal distribution: $a_{ij} \sim N(-0.25, 0.1)$; and the other half of the interactions were drawn from a positive normal distribution: $a_{ij} \sim N(0.25, 0.1)$. All interactions were restricted to lie between -0.5 and 0.5 , i.e., the normal distributions were truncated at -0.5 and 0.5 . The parameters r_i , K_i , and the interaction matrix A were randomly drawn 1000 times from the aforementioned distributions to obtain 1000 different parameter combinations. Hereafter, host-specific parameters were drawn from log-normal distributions around species-specific parameters, as follows:

$$\begin{cases} \ln(|a_{ij,m}|) \sim N(\ln(|a_{ij}|), \sigma_a) \\ \ln(r_{i,m}) \sim N(\ln(r_i), \sigma_r) \\ \ln(K_{i,m}) \sim N(\ln(K_i), \sigma_K) \end{cases} \quad \text{Eq. 2.4}$$

Here, σ_a denotes the interindividual variability in interspecific interactions among the 300 hosts (with $\sigma_a = 0.25$ in the base case), and $|a_{ij,m}|$ denotes the absolute strength of interaction from species j on the growth of species i for each host m . Note that, for the sake of simplicity, the use of log-normal distributions was adopted to induce fold-changes around population means, where both the presence and the sign of interspecific interactions are kept constant

across hosts. However, this may be untrue in real microbiota as many microbes can change metabolic pathways and therefore may switch between interaction types and interaction partners. In the base case model, the carrying capacities and growth rates were kept constant across hosts, meaning σ_r and σ_k were set equal to 0.

The simulation process yielded 300,000 time series (300 host-specific time series for each of the 1000 ten species networks). The running time of the model was chosen such that all species reached their equilibrium abundance. If at least one species did not survive (i.e., when its abundance dropped below 0.001), we rejected the simulation in favor of another randomly drawn parameter set. After sampling the abundances at equilibrium, we added independent and identically distributed noise v to mimic uncertainty in measurements (with $v \sim U(-0.01, 0.01)$ in the base case). This measurement noise can be thought of as representing, for example, sampling errors, environmental contamination, batch effects during sequencing, or annotation errors in reference genomes.²⁶¹ Simulations were performed in R (R version 3.6.0; www.r-project.org). The gLV model was solved with the 'lsoda' function from the 'deSolve' R package (version 1.24) which uses a FORTRAN ODE solver written by Petzold & Hindmarsh (1995).^{262, 263} R code is available on the GitHub repository ([susanepinto/gLV_microbiome](https://github.com/susanepinto/gLV_microbiome)). A general overview of the base case simulation design is given in Figure 2.1.

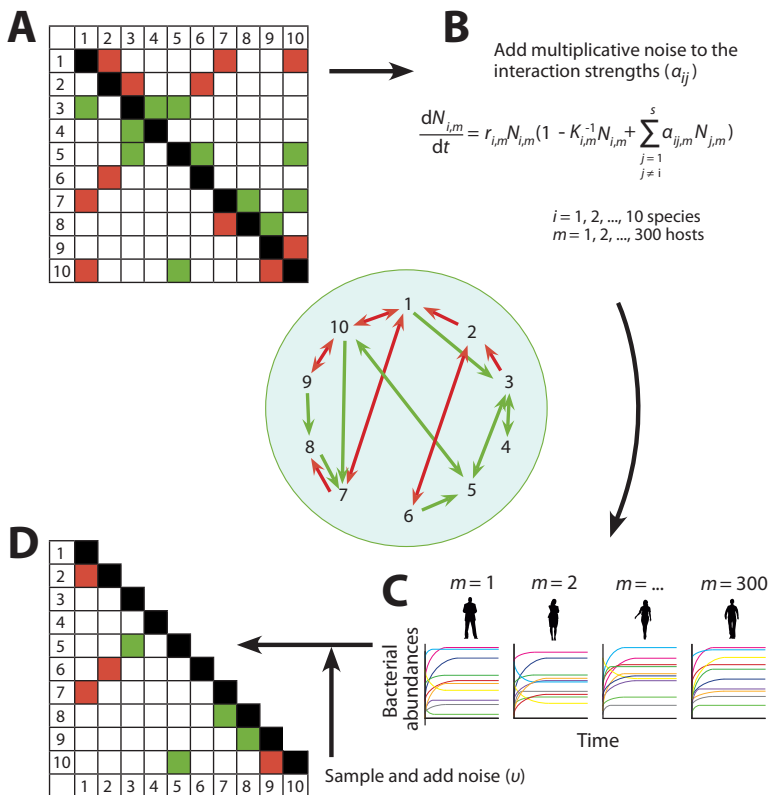


Figure 2.1 - Representation of the workflow. In an interaction network, singular green and red arrows represent a commensalistic interaction and an amensalistic interaction, respectively, whereas double green arrows represent mutualism and double red arrows competition. A combination of a green and red arrow signifies an exploitative interaction.

See Appendix Figure 2.1 for more details. A) A random interaction matrix i . This interaction matrix is implemented in the gLV model, B) together with the intrinsic growth rates and carrying capacities of the species. C) All time series are (slightly) different due to the variation in the interaction strengths. D) The partial correlations are calculated from the abundances per species sampled from the 300 different hosts at equilibrium. Only the significant correlations and the lower part of the matrix are used for the comparison with the original interaction matrix i . Variations to the workflow were studied by adding for example a perturbation or process noise.

Variations to the base case model

We studied multiple variations to the base case model. Similar to the base case simulations, we did 1000 simulations per variation. As a first variation, we added host-specific variability to the species-specific parameters r_i and K_i using Equation 2.4, with $\sigma_r = 0.25$ and $\sigma_K = 0.25$.

Second, we varied the amount of measurement noise, from $v \sim U(-0.01, 0.01)$ (medium noise in the base case) to $v \sim U(-0.001, 0.001)$ (low noise), and to $v \sim U(-0.1, 0.1)$ (high noise). We also simulated time series with a different type of noise, namely varying magnitudes of process noise W (Appendix Figure 2.2). In contrast to measurement noise, which was added only to the sampled abundances, process noise was added to the gLV model such that within-host population dynamics were perturbed at discrete time intervals Δt ($\Delta t = 1$ time unit). The time-varying process noise was drawn from a log-normal distribution to prevent the abundances from dropping below zero, i.e., $\Delta W_i = \ln(N_{i,m(\Delta t)}) - \ln(N_{i,m(t)}) \sim N(\ln(N_{i,t}), \sigma_W)$ (with $\sigma_W \sim N(0, 1)$ for high process noise and $\sigma_W \sim N(0, 0.1)$ for low process noise).

Further, we simulated data with interaction strengths drawn from a uniform ($\alpha_{ij} \sim U(-0.5, 0.5)$) or unimodal ($\alpha_{ij} \sim N(0, 0.15)$) distribution. As in the base case, the interaction strengths were restricted to lie between -0.5 and 0.5 (Appendix Figure 2.3).

We also analysed three different structures of microbial networks. First, we increased the number of species s from 10 to 30. To promote coexistence, we also reduced the density of the interaction matrix to $\frac{1}{6}$. Secondly, we simulated a network based on a producer-consumer relation between the species (Appendix Figure 2.4). Instead of random interaction networks (Appendix Figure 2.4A), the producer-consumer networks are based on a cross-feeding structure between producers and consumers (with equal numbers of producers and consumers) (Appendix Figure 2.4B). Producers excrete metabolites which are consumed by the consumers. Because consumers remove the 'waste' from the producers, the presence of a consumer can also be beneficial for the producers. Therefore, between producers and consumers positive interactions are more likely to occur than negative interactions. For this purpose, we drew the consumer-producer interactions from the positive side of the Gaussian mixture distribution ($\alpha_{ij} \sim N(0.25, 0.1)$). In contrast, among producers and consumers themselves, the interactions are predominantly negative as these species are more likely to compete for similar resources. For this purpose, we drew the interactions among producers and among consumers from the negative side of the Gaussian mixture distribution ($\alpha_{ij} \sim N(-0.25, 0.1)$). Third, we simulated a microbial network with interaction hubs, i.e., a network containing species with unusually high numbers of ecological interactions compared to other species in the network (Appendix Figure 2.4C).²⁶⁴ Hub-species networks were created according to the Barabási-Albert model²⁶⁵ and implemented with the 'barabasi.game' function from the 'igraph' R package (version 1.2.11). In the network-generating algorithm, interactions are distributed according to a mechanism of preferential attachment.

Thus, species with interactions obtain a higher chance of getting more interactions, resulting in a few 'hub-species' with many interactions. We constructed two scale-free directed graphs (with power = 2), denoting 'incoming' and 'outgoing' interactions, and combined these to obtain a bidirected graph. Density was kept similar to the base case model (1/4).

Next, we also investigated how network inference is affected by sample size by considering a scenario with 3000 instead of 300 hosts. We did this for the base case model with random interaction networks, as well as for the producer-consumer and hub-species networks described before.

Last, we investigated the effect of a perturbation on the performance of network inference. The populations were perturbed after 175 time units, with a perturbation that lasted for 50 time units. The perturbation was modelled by taking a new set of random carrying capacities per species per sample. Due to the simulated perturbation, the equilibrium distribution shifted. After the perturbation, the species grew back to their original equilibrium. Sampling occurred before, during, or after the perturbation.

Assessment of correlation-based network inference

With the simulated data at hand, we created a dataset with the abundances of the model species sampled at equilibrium for each host m . After adding measurement noise to the data, we inferred the correlations between species by calculating the Pearson's partial correlation coefficients ρ between all abundances N_i across the m different hosts (Figure 2.1). We did not use plain correlations, because partial correlations have the advantage of controlling for confounding interactions (e.g., interactions between bacterial species affecting the abundance of a third species).²²⁷ Agreement between the partial correlation matrix and the interaction matrix A from the gLV model was assessed qualitatively, i.e., we only considered whether significant entries in the partial correlation matrix agreed with the interaction matrix in terms of nonzero entries with the same sign. We used the Benjamini-Hochberg procedure to control for the expected proportion of 'false discoveries' after calculating partial correlations between each pair of species.²⁶⁶ The results (true positives, true negatives, false positives, and false negatives) were stored in a confusion matrix (Table 2.1). Because a correlation matrix is symmetric and an interaction matrix A is not, we only used half of the partial correlation matrix (Figure 2.1D) to construct the confusion matrix. For a correctly classified interaction, either one or both interactions in the upper and lower part of the A matrix must have the same sign as in the lower part of the partial correlation matrix. This can produce a bias, because asymmetric interactions can result in a true positive result for correspondence of the correlation coefficient (ρ) with either interaction. For example, for exploitative interactions, both negative and positive correlations were classified as true positive results. Therefore, we tested the effect of this bias on the success of network inference by specifying the intended sign in correlation analysis, as the sign of the strongest interaction in each pair of species. Hence, for an exploitative interaction, only a positive or a negative correlation is correct, depending on the weights of the asymmetric interactions. We also tested the effect of this bias on the success of network inference by setting the rule that the sign of both interactions must be matched by the inferred correlation coefficient. Therefore, only mutualism and competition can be inferred correctly, as amensalism, commensalism, and exploitative interactions are asymmetric.

Table 2.1. The confusion matrix as used in this study.

The inferred partial correlation coefficient ρ (from the lower part of the partial correlation matrix) must have the same sign as one of the interactions in the interaction matrix A to be considered as a true positive finding in base case analysis.

Interaction in the A matrix from the model		Inferred partial correlation		
		Negative ^a	Not significant	Positive ^a
No interaction	0, 0	false positive	true negative	false positive
Mutualism	+, +	false positive	false negative	true positive
Competition	-, -	true positive	false negative	false positive
Commensalism	+, 0 0, +	false positive	false negative	true positive
Amensalism	-, 0 0, -	true positive	false negative	false positive
Exploitative interaction	+, - -, +	true positive	false negative	true positive

^a Only significant partial correlations (with p -value < 0.05) are considered after correction for multiple testing with Benjamini-Hochberg procedure.

Performance of network inference was evaluated using precision and recall, as well as a combination of both measures, called the F_1 -score.²⁶⁷ The precision is the fraction of correctly classified interactions among the total number of significantly predicted interactions (i.e., significant partial correlations) and the recall is the fraction of correctly classified interactions among the total number of non-zero interactions in the interaction matrix A . The F_1 -score (on a scale from 0 (no agreement) to 1 (perfect agreement)) is obtained as the harmonic mean of precision and recall, weighted equally, as given in the following equation:

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad \text{Eq. 2.5}$$

Results

Inference of asymmetric and symmetric interactions in a two-species system

Correlations in abundances of the species in a two-species Lotka-Volterra model are shaped by the type of interaction involved. Figure 2.2 shows scatterplots of the abundances of two bacterial species for different interaction mechanisms over a range of different combinations of α_{12} and α_{21} . Mutualistic interactions clearly yielded a positive correlation in abundance between the two species involved (Figures 2.2A and Appendix Figure 2.5). Competitive interactions generally yielded negative correlations (Figures 2.2B and Appendix Figure 2.5). However, under perfectly symmetric competition (when $\alpha_{12} = \alpha_{21}$), we did find a positive correlation depending on interaction strength and carrying capacities of the species involved (Appendix Figure 2.5D - second panel). In the situation where one of the two species does not experience any benefits or limitations in growth from the other species, as is the case with commensalism and amensalism (i.e., $\alpha_{12} = 0$ or $\alpha_{21} = 0$), correlations are zero because one of the species will grow to its carrying capacity irrespective of the abundance of the other species (Figure 2.2C and 2.2D).

Correlations under exploitative interactions among bacteria, benefitting one but harming the other species, generally yielded positive correlations (Figures 2.2E and 2.2F, and Appendix Figure 2.5), but negative correlations were also found.

This happened when the exploitative benefit was of equal magnitude as the harm done to the other species (Appendix Figure 2.5D), or of similar mean magnitude but with more variation (e.g., species 1 is exploited by species 2; $-a_{12} = a_{21}$ and $\sigma_{a12} \ll \sigma_{a21}$ (exploitative interaction type 1) or species 2 is exploited by species 1; $a_{12} = -a_{21}$ and $\sigma_{a21} \ll \sigma_{a12}$ (exploitative interaction type 2) (Appendix Figure 2.5B). However, if the exploitative benefit outweighs the harm done to the other species, exploitative interactions will generally yield positive correlations. It should also be noted that the two species were not exchangeable, because species 1 was given a weaker intraspecific interaction strength than species 2. Thus, in the absence of interspecific interactions, species 1 can reach a higher abundance at equilibrium. This means that, for the same interspecific interaction strength, the species with the higher carrying capacity exerts a stronger (negative) effect on the growth of the other species.

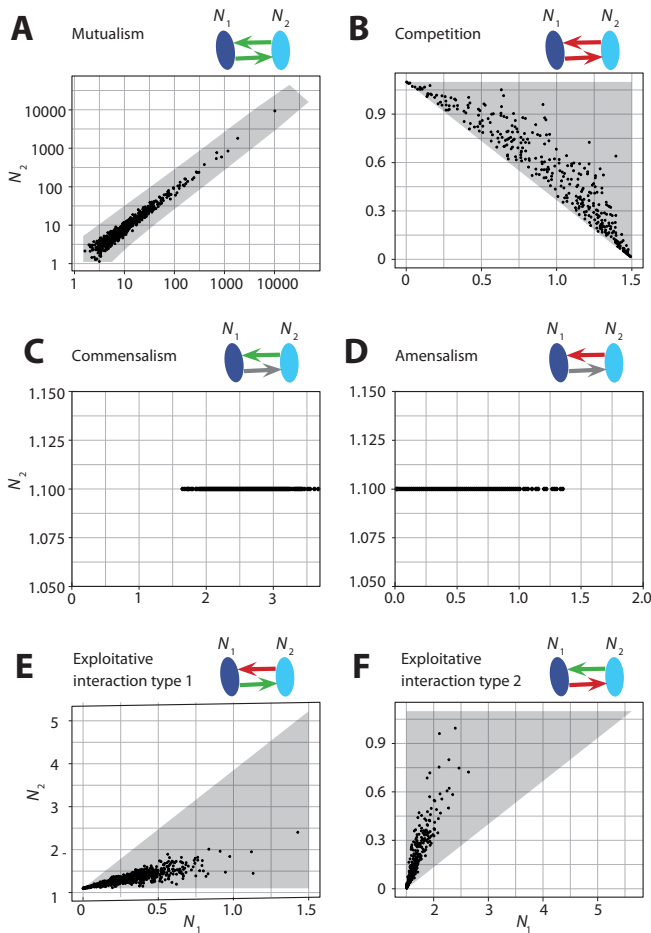


Figure 2.2 - Scatter plots between the abundances of two bacterial species for different interaction mechanisms. A) mutualism, B) competition, C) commensalism, D) amensalism, and E, F) exploitative interactions. The abundances of the two species N_1 and N_2 at equilibrium are shown as scatterplots and have been obtained by running the two-species Lotka-Volterra model, with $K_1 = 1.5$; $K_2 = 1.1$; $r_1 = 1$; $r_2 = 2$ and α_{ij} drawn randomly from normal

distributions with identical means and standard deviations ($\alpha_{12} \sim N([0.7], 0.2)$, $\alpha_{21} \sim N([0.7], 0.2)$). In the case of commensalism and amensalism: $\alpha_{12} \sim N([0.7], 0.2)$ and $\alpha_{21} = 0$. The two species can co-exist under certain combinations of α_{ij} (Box 2.1). The grey polygon indicates the area where coexistence is possible. Note that the axes have different ranges in each subplot. Because the two species have different carrying capacities, the two situations of exploitative interactions are different; i.e., in case of exploitative interaction type 1 (species 1 is exploited by species 2) and in case of exploitative interaction type 2 (species 2 is exploited by species 1).

Network inference under various interaction types

Here, we used the base case model to assess the success rate of recovering a particular interaction type between pairs of species: amensalism, commensalism, exploitative interactions, mutualism, and competition (Appendix Figure 2.1). Figure 2.3A shows that correlations were more often found in mutualistic and competitive interactions, where interacting species experience the same qualitative effects from each other, than in amensalistic and commensalistic interactions, where only one species experiences an effect from the presence of another species. For exploitative interactions among bacteria, either a positive or negative correlation coefficient ρ could be found, with a success rate comparable to amensalistic and commensalistic interactions. Contrary to the results that included symmetric interactions, there was no difference between the successful inference of positive interactions over negative interactions in any interaction type (Figure 2.3B). For all interaction types, the sign of the significant correlation coefficient ρ found, mostly agreed with the sign of the interaction type (Figure 2.3). However, with the inferred correlations neither the type nor direction of the original interaction could be recovered.

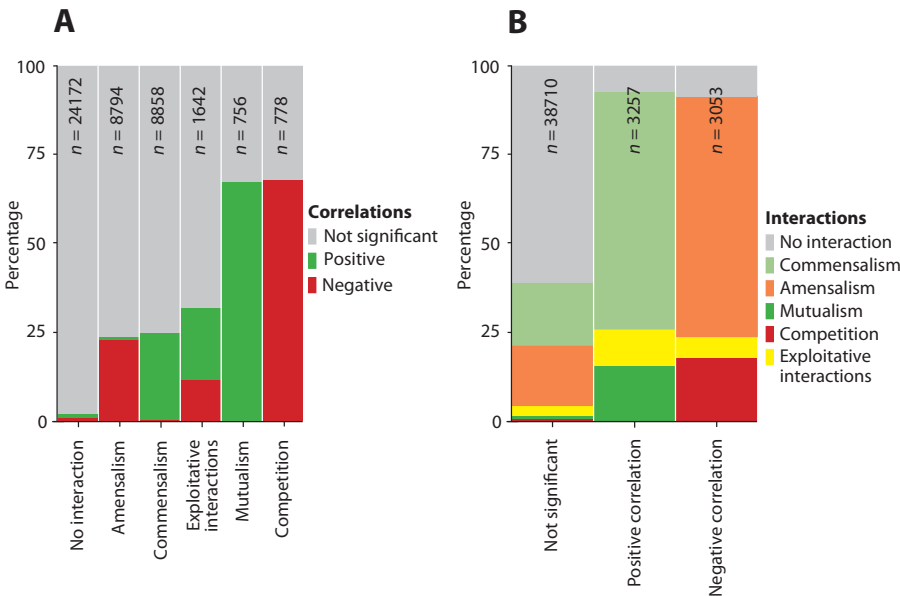


Figure 2.3 - The percentage of significant partial correlations (with sign matching interaction in either direction), as recovered from the base case model.

A) For different types of pairwise interactions and B) for the different correlations.

Network inference under various sources of process variability

Next, we investigated how correct network inference was affected by several variations to the base case model (Figure 2.4 and Appendix Table 2.1). In all cases considered, interactions were recovered with precision exceeding recall. This means that the likelihood of missing an interaction (i.e., $1 - \text{recall}$) was higher than the likelihood of finding a false interaction (i.e., $1 - \text{precision}$), illustrating the effect of false discovery rate control.

Partial correlations corresponded to non-zero entries in the interaction matrix only when interindividual variation existed in the interaction parameters (α_{ij}) and/or carrying capacities (K_i) (Figure 2.4A and 2.4B). These parameters directly influence microbial abundance patterns, as interspecific interactions and carrying capacities determine the equilibrium of the gLV model. The intrinsic growth rate only determines the speed at which species reach their equilibrium, and this parameter is not informative for the equilibrium abundances. In fact, performance under interindividual variation in growth rates was just as bad as the performance under pure measurement noise with no variation in model parameters (Figure 2.4B).

Performance of correlation-based network inference was robust to measurement noise, if measurement noise was small compared to interindividual variation in process parameters (Figure 2.4C). When measurement noise became of the same magnitude as the variation in interspecific interactions, the F_1 -score deteriorated, and it was no longer possible to use correlations as a proxy for interactions (Figure 2.4C). We also checked whether adding process noise would affect the inference. We did observe a significant improvement of the inference from a model with process noise relative to only measurement noise (Figure 2.4C and Appendix Table 2.1).

Hereafter, we investigated the effect of drawing the interaction strengths from different types of distributions (Figure 2.4D and Appendix Figure 2.3). We did not observe a difference between the success rate of network inference under a Gaussian mixture distribution or uniform distribution, which were conditioned to have similar variances (Appendix Table 2.1). However, successful inference deteriorates with reduced interaction strength; success rates were better under a Gaussian mixture distribution or uniform distribution compared to a unimodal distribution around zero (with smaller variance) (Figure 2.4D). The weaker interactions have a smaller effect on equilibrium abundances of other species, which makes them harder to detect with correlation analysis.

Figure 2.4E shows the results for different network types. Increasing the number of species from 10 to 30 had a significant negative effect on the success of the inference (Appendix Table 2.1), which was mainly due to reduced precision. Conversely, F_1 -scores were improved as compared to the base case when assuming a producer-consumer based network (Appendix Figure 2.4 and Appendix Table 2.1), on account of an improved recall. Inference in a network with interaction hubs (as explained in Appendix Figure 2.4) was significantly worse than in a random network, which could be attributed to a somewhat reduced recall.

Note that problems may arise with asymmetric relationships. When using the rule that pairwise correlations should match the strongest interaction between both species involved as the intended sign, we found only a slight non-significant reduction in F_1 -score as compared to the base case scenario (Figure 2.4F and Appendix Table 2.1).



Figure 2.4 - Inference under various sources of process variability. For the different scenarios we show the precision, recall, and the F_1 -score. A) The base case model. B) Host-specific variation in the carrying capacities and intrinsic growth rates. C) Decreased and increased amount of measurement noise (ν) and the effect of process noise (W) (Appendix Figure 2.2). D) Interaction strengths drawn from a uniform and unimodal distribution (Appendix Figure 2.3). E) The results for a 30-species system, a network based on a producer-consumer structure and a network with interaction hubs (Appendix Figure 2.4). F) The effect of network inference when specifying the intended sign in correlation analysis, as the sign of the strongest interaction in each pair of species, or by setting the rule that the sign of both interactions must be matched by the inferred correlation coefficient (strict inference). G) Three scenarios with 3000 hosts, for the base case with random interaction networks as well as for the scenarios with structured (i.e., producer-consumer and hub-species) networks. Network inference was assessed by the F_1 -score, which measures agreement between the interaction matrix in the gLV model and the inferred partial correlation matrix on a scale from 0 (no agreement) to 1 (perfect agreement) (according to the rules of Table 2.1). The dashed line indicates the median result from the base case model. The bars of the boxplots indicate the variability of the data outside the middle 50% (i.e., the lower 25% of scores and the upper 25% of scores). All p -values are given in Appendix Table 2.1.

Thus, pairwise interactions wherein the net effect on population growth is positive or negative are mostly picked up as such in correlation analysis. However, under the rule that mutual interactions must both be reflected in the sign of the correlations, asymmetric interactions cannot be recovered as correlations are symmetric. We indeed found much lower F_1 -scores when detection of asymmetric interactions was no longer considered as a true positive result after inferring a significant correlation coefficient ρ (either positive or negative) (Figure 2.4F).

Finally, we verified that network inference improved with increasing sample size. This applied to models with random as well as structured interaction networks (Figure 2.4G). In the base case, precision was somewhat reduced at increased sample size notwithstanding Benjamini-Hochberg control. However, this was compensated by substantially improved recall, resulting in significantly increased F_1 -scores. Interestingly, precision stayed more or less constant at increased sample size in producer-consumer and hub-species networks, whereas recall improved but remained somewhat behind that of random networks.

Network inference under non-equilibrium conditions

Figure 2.5 shows that the equilibrium assumption is not necessary for successful correlation-based network inference. In fact, our results even suggest that a perturbation can positively affect the performance of network inference. Variation in the growth rates becomes significantly informative outside the equilibrium (Appendix Table 2.2). Also, variation in the interactions becomes even more informative when the population is still growing towards the equilibrium. Network inference is impaired only right after the start of a perturbation, when the population is still far from a new equilibrium, unless the interindividual variation is in the carrying capacities (Figure 2.5B). We also assessed the success of correlation-based inference when the sampling occurred randomly in time in relation to the perturbation. We found that the F_1 -score resembled an average of F_1 -scores across various sampling time points.

Discussion

Correlation-based network inference has been used in many studies and for many different types of human and environmental microbial communities.²²⁷ The reliability of the results with regards to true ecological dependencies has been criticized, to the extent that correlation analysis has been suggested to almost never reveal anything substantive about the biotic relationships between bacteria.²¹⁶ However, the theoretical basis that enables ecological interactions to be inferred from cross-sectional abundance data remains poorly understood. Most of the previous research has focused on the reconstructed network properties or the difficulties pertaining to metagenomics-based abundance patterns, e.g., the compositionality of the data and the high proportion of zeros.^{216, 227, 268} While these difficulties are pervasive and merit further consideration, here, we question whether correlations are at all useful in distinguishing different interaction types in microbial networks.

We demonstrated multiple pitfalls when using correlation-based methods for inferring interactions. Some of those pitfalls are well known, as they relate to the inherent symmetry of correlation-based metrics and the frequent asymmetry of ecological interactions.²¹⁶ As a result, asymmetric interaction types (commensalism, amensalism, and exploitative interactions) cannot be recovered with an indication of the direction of interaction, which agrees with prior work done by Weiss et al. (2016).²⁵³

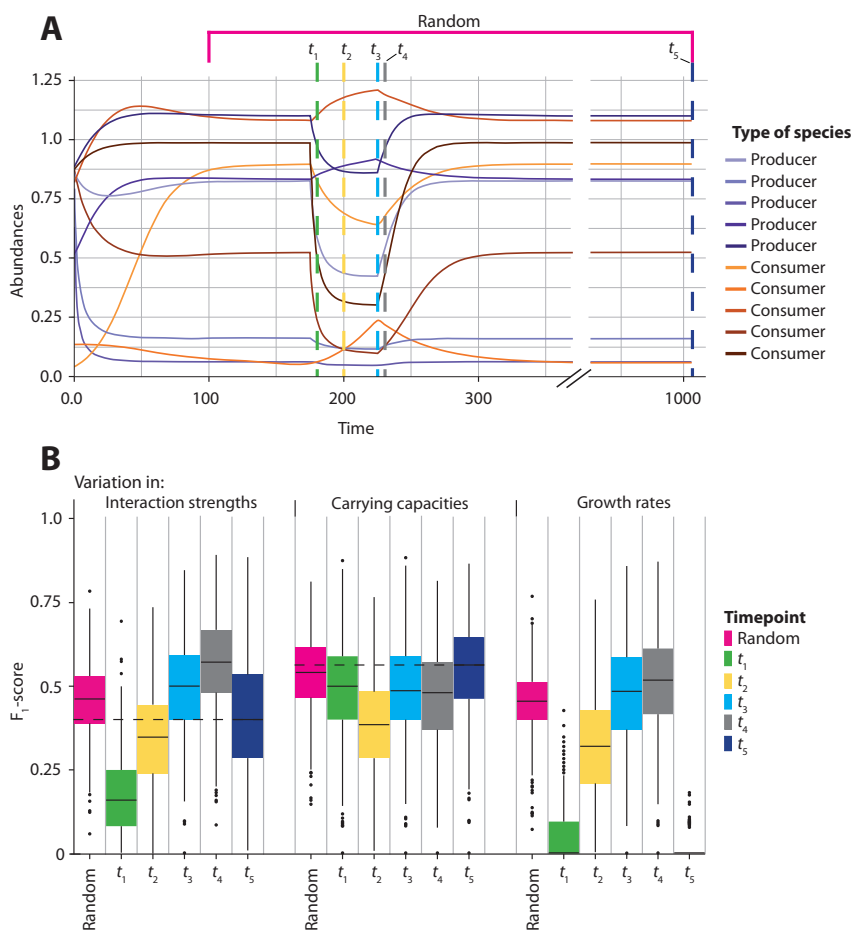


Figure 2.5 - The effect of a perturbation on correlation-based network inference.

A) Example of a time series. Dashed lines represent sampling time points. Sampling was performed during the perturbation (t_1 = green, t_2 = yellow, t_3 = blue, and t_4 = grey) and at equilibrium (t_5 = dark blue). Alternatively, sampling was performed randomly between $t = 100$ and $t = 1000$ (random = pink). **B)** Results (F_1 -scores) of network inference for sampling at various time points. After a perturbation all species grow back to their original equilibrium. The bars of the boxplots indicate the variability outside the middle 50% (i.e., the lower 25% of scores and the upper 25% of scores). Dashed lines represent median results of sampling during equilibrium. All p -values are given in Appendix Table 2.2.

Symmetric interaction types, where species involved affect each other's growth in a qualitatively similar way (competition and mutualism) can be recovered, although competitive interactions may also result in positive correlations, albeit in very rare cases where species have identical competitive strength. Likewise, we found that exploitative interactions generally induce positive correlations, especially in the likely circumstance where the exploitative benefit outweighs the harm to the exploited species. These findings might explain why empirical correlation-based networks have a relative shortage of negative correlations.^{119, 268, 269} It remains to be investigated whether the high frequency of positive edges in reconstructed networks is caused by methodologic limitations or whether the interspecific interactions in host-associated microbiota are primarily mutualistic.^{112, 136, 270, 271}

Still, as illustrated by our analysis, correlations in microbial abundance across independently sampled hosts can be indicative of underlying ecological interactions under host-specific variation in microbial population dynamics. That is, if microbial groups of interest are omnipresent and their interactions are appropriately captured by generalized Lotka-Volterra (gLV) dynamics, the variation in population abundances should be driven by interindividual variability in population-dynamic parameters. In the context of the gLV model, the informative parameters are primarily related to intrinsic growth rates, carrying capacities, and strength of between-species interactions of microbial groups considered. A change in species abundances can be informative for the interactions among those species, as was also previously shown by Stone and Roberts (1991).²⁷² It remains to be determined how much variability across individual hosts is driven by external forcing and by gradual differences in process related parameters relative to measurement noise. On one hand, it is well known that microbes adapt to host-specific environments, shaped by, among others, diet, lifestyle, hormonal regulation, and the immune system.²⁶⁰ As an example, increased abundance of a particular bacterial species at increased glucose intake levels might be reflective of increased resource availability (affecting carrying capacity and growth rate) or superior competitive strength (affecting interactions with other species).²⁴⁶ On the other hand, environmental drivers of bacterial growth can operate over different spatial and temporal scales and correlations in abundance can be reflective of shared environmental niches that have no meaning in terms of direct biotic interactions.²¹

Therefore, a correlation between the abundance of two species does not imply that those species are interacting.¹¹¹ Many of the detected correlations may be caused by shared environmental preferences rather than species interactions.²⁷³ Such environmental filtering can mask putative between-species interactions as well as induce spurious correlations.²¹⁶ Also, co-occurring species may appear to be dependent on each other, while their co-occurrence can be explained by them actually sharing a similar dependency on a third species so that co-occurrence, and hence apparent dependencies drawn from that, may also be explained by higher-order interactions.²⁷⁴ Berry and Widder (2014) suggested that network interpretation is only possible if samples are derived from similar environments.¹¹⁴ Our analysis suggests that network inference partially depends on a degree of heterogeneity in population-dynamic parameters. If differences in bacterial abundances between hosts are mainly due to measurement noise, their correlations are not informative of underlying interactions. In our simulations, with relative standard deviation in process-related parameters between hosts of about 25%, inference performed well as long as measurement noise had coefficients of variation well below 10% of the mean bacterial abundances. Strikingly, the inference of interactions was even improved when process noise was added. More research is needed to delineate the extent to which correlation analyses can be confounded by latent

environmental drivers of microbial population dynamics, and how strongly one should condition on environmental or host homogeneity.

Our results have been obtained using the gLV model. While the gLV model has been very popular in microbiome research because of its manageability, it has several drawbacks. In ecology, the gLV model has been criticized for the absence of trophic levels within the model.²⁷⁵ This is in contrast to most classical ecological (e.g., plant-herbivore or predator-prey) systems, where direct consumption and predation offer more opportunity for top-down regulation, possibly obscuring interactions in co-occurrence patterns.²⁷⁶ But trophic levels are probably not so relevant in the human microbiome as bacteria mainly interact with each other through excreted metabolites.⁸ Furthermore, the interactions between bacteria might be much more complex than the additive and pairwise interactions that the gLV model assumes. Momeni et al. (2017) claimed that pairwise modeling will often fail to predict microbial dynamics, as many interactions occur through chemical production pathways (such as cross-feeding and nutrient competition) involving more than two species.²⁴⁰ Correlation analysis fails to capture the resulting higher-order interactions, for which more advanced techniques, e.g., graphical models, might be more appropriate.²⁷⁷ It is unclear how well directed links predicted by these methods recover true ecological interaction types. Often, they require more prior knowledge of the network of microbial interactions, time series, or more fine-grained data on the pathways of interaction. Moreover, microbial networks can be bidirected and cyclic,¹¹⁹ which poses problems for inference of directionality and type of interactions from mere cross-sectional data. More classical methods of separating direct from indirect interactions, e.g., path analysis,²⁷⁸ rely on testing of specific alternative causal hypotheses, which can only be considered as a next step in network inference. To shed more light on causal pathways, there is a need in microbial ecology for models that can describe the full set of metabolite concentrations, metabolic fluxes, and species abundances within a community.¹²⁰ Based on metabolic modeling, Freilich et al. (2011) concluded that cooperative interactions are relatively rare among free-living bacteria and, if present, are often unidirectional. Machado et al. (2021) suggested that mutualistic interactions are much more common among host-associated bacteria, that often form highly cooperative communities and have smaller genomes and fewer metabolic genes compared to other species. Cooperative communities are resilient to nutrient change and adaptable to a wide variety of different environments, including the human body.^{119, 274} Metabolic modeling is still challenging and heavily based on a priori assumptions, but is also a rapidly developing field that may prove useful for computational validation of correlation-based interaction networks.²⁷⁹

In addition, the gLV model disregards important biological processes, such as adaptation (for instance, switching of mutualistic partners due to for example horizontal gene transfer²⁸⁰), that may affect the topology of ecological networks, rather than the strength of ecological interactions in a network. Furthermore, the gLV model displays dynamics that are characterized by strong equilibrium attractors. Many studies have shown the occurrence of complex dynamics as alternative stable states,¹⁰⁵ oscillations, and chaos in experimental,^{6, 281, 282} but also in field studies,⁶ with ecological communities. Whether this also applies to the bacterial communities inhabiting the human body is still unknown, due to the paucity of long-term human microbiome studies. However, a study among a thousand Western individuals has suggested the existence of tipping elements in the intestinal microbiome¹⁰² indicating the possible presence of alternative attractors in the dynamics of gut microbiome communities.^{97, 283}

As a general critique, the use of simulated data based on gLV dynamics raises the question to what extent the necessary model assumptions (and therefore the results) are representative for the human microbiome. Of course, real data are much more complex than simulated data. To reiterate, our base case parametrization does not reflect any particular real-world system, and findings should be appreciated from a qualitative rather than quantitative viewpoint. Even so, while models can only serve as very crude approximations, the main features of model-based analysis might still hold, as demonstrated by Freilich et al. (2018).²⁷³ They compared a well-resolved, empirically defined interaction network of species in the rocky intertidal zone in central Chile to a reconstructed network based on the co-occurrence of those species. There are similarities in their findings to our results. For example, they found that weak interactions are missed more often than interactions above a certain threshold. They also concluded that the ability to correctly detect a true link varies across different interaction types, and that positive interactions are better detected than negative interactions. Interestingly, in line with our results, they also found that negative interactions are misclassified as positive interactions more often than vice versa.

In our simulation studies, the chance of finding false interactions was well under control using partial correlations with adjustment for multiple testing. It should be noted that application of correlation-based network reconstruction to real-world high-throughput microbial abundance data typically requires additional constraints for control of false discovery rates. Real-world microbiome data have some specific challenges which may negatively affect the success of correlation-based network inference. The compositionality of the data, the diversity of species (with many rare species) and the density of interactions make these networks harder to predict and apparent correlations more likely to appear.^{229, 250} Various correlation-based methods, often free of charge and provided as pre-programmed packages are available to handle these challenges. However, Weiss et al. (2016) showed that with the same data, there is much disagreement between the inferred networks generated by different tools.²⁵³ Thus, even if correlations are a useful proxy of microbial interactions, performance of network inference in high-dimensional settings will also strongly depend on the specific network modeling approach taken.

To summarize, correlation-based methods are particularly insensitive for the detection of asymmetric interactions (such as exploitative interactions, amensalism, or commensalism), as direction of interaction cannot be recovered from co-occurrence data. Still, they may perform well when applied to networks that are dominated by mutualistic and competitive interactions, as in producer-consumer systems. Applicability of correlation-based network inference to readily available microbiome data thus depends on the type of interactions that govern microbiome dynamics, which likely depends on each application. To conclude, our study suggests that hypotheses about microbial interactions, generated with correlation-based methods, should be questioned with domain-specific knowledge. We highlight again the careful interpretation and validation that is required.

Appendices of Chapter 2

Appendix Table 2.1 - Mann-Whitney U test results for the F_1 -scores of the base case model and for the F_1 -scores of the model with different sources of process variability.

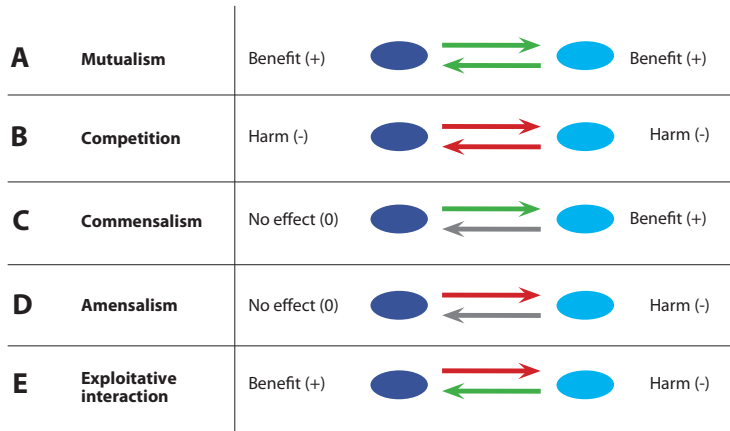
Significant results are highlighted in bold and blue.

Scenario	<i>p</i> -value
Variation in all parameters	< 0.001
Variation in carrying capacities	< 0.001
Variation in growth rates	< 0.001
No variation in parameters	< 0.001
Low measurement noise	< 0.001
High measurement noise	< 0.001
Low process noise	< 0.001
High process noise	< 0.001
Uniform distribution	> 0.05
Unimodal distribution	< 0.001
30 species system	< 0.001
Producer-consumer network	< 0.001
Hub-species network	< 0.001
Inference of intended sign	> 0.05
Strict inference	< 0.001
Base case (3000 hosts)	< 0.001
Producer-consumer network (3000 hosts)	< 0.001
Hub-species network (3000 hosts)	< 0.001

Appendix Table 2.2 - Mann-Whitney U test results for the F_1 -scores of the samples taken outside equilibrium relative to those taken at equilibrium (t_5 in Figure 2.5) and for the F_1 -scores of the samples taken randomly.

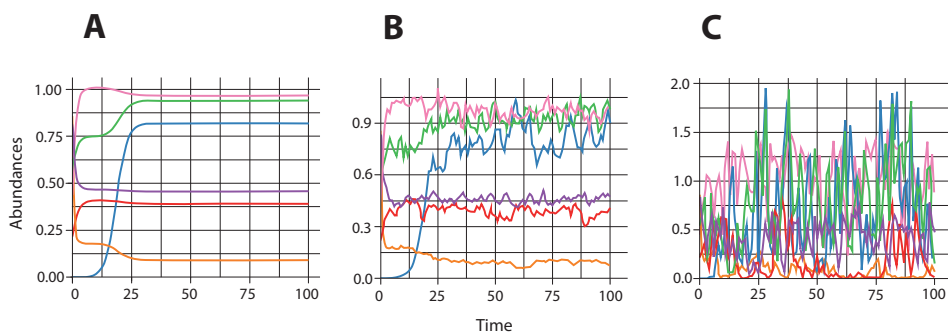
Significant results are highlighted in bold and blue.

Variation in interactions	
Time point	p-value
Random	< 0.001
t_1	< 0.001
t_2	< 0.001
t_3	< 0.001
t_4	< 0.001
Variation in carrying capacities	
Time point	p-value
Random	> 0.05
t_1	< 0.001
t_2	< 0.001
t_3	< 0.001
t_4	< 0.001
Variation in growth rates	
Time point	p-value
Random	< 0.001
t_1	< 0.001
t_2	< 0.001
t_3	< 0.001
t_4	< 0.001



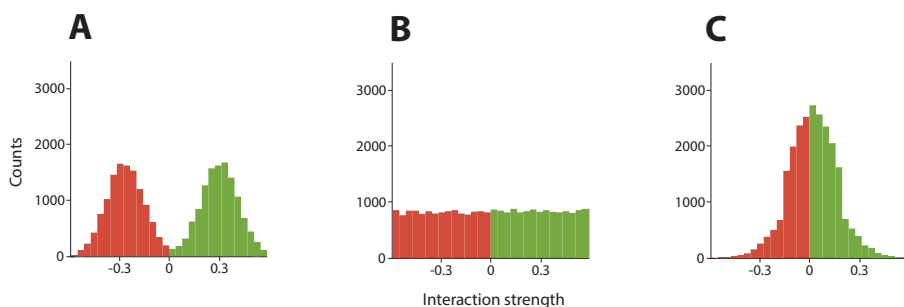
Appendix Figure 2.1 - Cartoon illustrating the different interaction mechanisms.

A) In mutualistic interactions, both species experience a positive effect. An example is when a species feeds on the metabolites excreted by the other species. B) In competitive interactions both species experience a negative effect. An example is when both species produce toxic compounds that are harmful to the other species as well as to themselves. C) Commensalism is a one-sided positive interaction. This type of interaction occurs when one species is beneficial to another species, without benefit or harm to itself. D) Amensalism is a one-sided negative interaction. Amensalism occurs when a species causes harm to another species, without benefit or harm to itself. E) Exploitative interactions occur when one species derives a benefit from another species at the expense of the latter, such as when one species kills and subsequently consumes the other. Red arrows represent negative interactions, green arrows represent positive interactions, and grey arrows indicate no interactions.

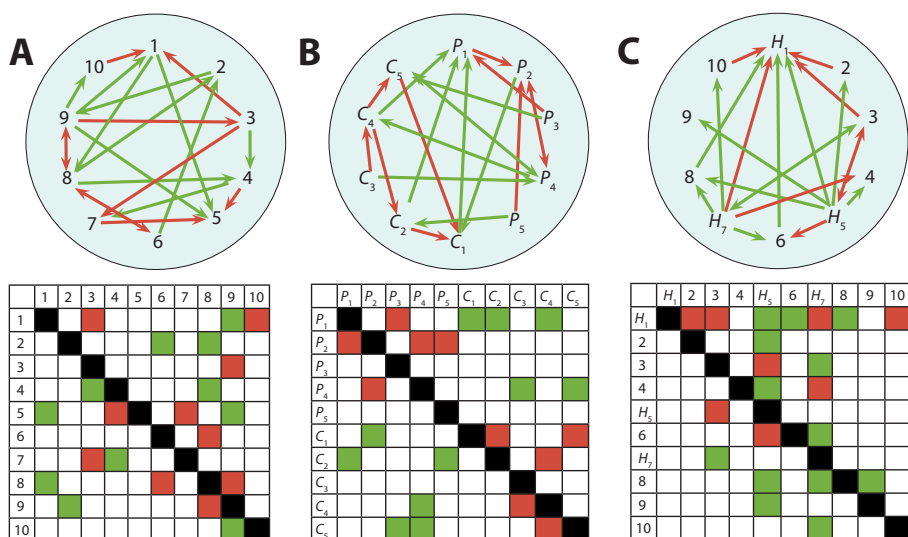


Appendix Figure 2.2 - The effect of process noise (W) on the within host population dynamics.

Process noise was added by means of the 'events' function from the 'deSolve' R package.²⁶³ The time-varying noise was drawn from a log-normal distribution to prevent the abundances from dropping below zero, i.e., $\Delta W_i = \ln(N_{i,m(\Delta t)}) - \ln(N_{i,m(t)}) \sim N(\ln(N_{i,t}), \sigma_W)$ at every timestep, $\Delta t = 1$. A) Simulated time series without process noise, B) with low process noise ($\sigma_W \sim N(0, 0.1)$), and C) high process noise ($\sigma_W \sim N(0, 1)$).

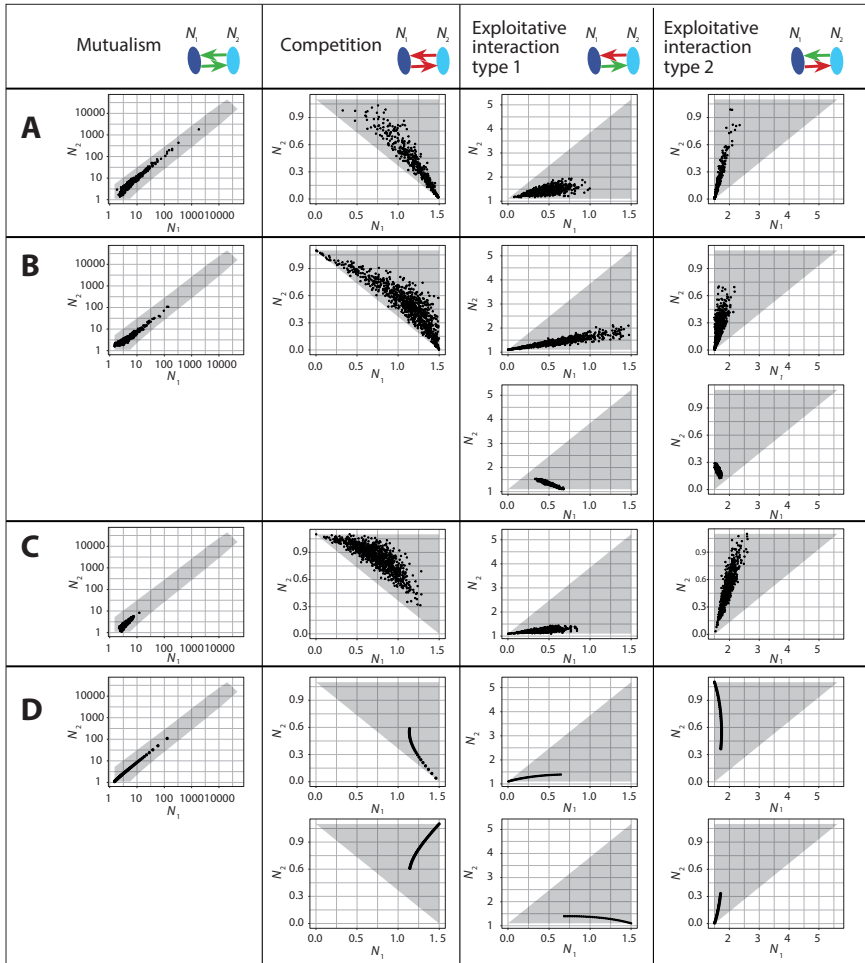


Appendix Figure 2.3 - Distributions of interaction strengths in three different scenarios. A) The interaction strengths in the base case follow a Gaussian mixture distribution. Half of the interactions were drawn from a negative normal distribution: $a_{ij} \sim N(-0.25, 0.1)$; and the other half of the interactions were drawn from a positive normal distribution: $a_{ij} \sim N(0.25, 0.1)$. B) The interaction strengths in Figure 2.4D-1 follow a uniform distribution ($a_{ij} \sim U(-0.5, 0.5)$). C) The interaction strengths in Figure 2.4D-2 follow a unimodal distribution ($a_{ij} \sim N(0, 0.15)$). All interactions were restricted to lie between -0.5 and 0.5 , i.e., the normal distributions were truncated at -0.5 and 0.5 .



Appendix Figure 2.4 - Network structures used in the different case studies.

A) An example of a random network and its corresponding interaction matrix. B) An example of a structured network with interaction modules and its corresponding interaction matrix. The modular networks are based on a cross-feeding structure between producers and consumers (with equal numbers of producers and consumers). Between producers ($P_i; i = 1:5$) and consumers ($C_j; j = 1:5$), positive interactions (indicated in green) are more likely to occur, because metabolites excreted by the producers are consumed by the consumer species. Among producers or among consumers, the interactions are predominantly negative (indicated in red) as these species are more likely to compete for similar resources. C) An example of a structured network with interaction hubs and its corresponding interaction matrix. The hub-species network contains species (H) with unusually high numbers of ecological interactions compared to other species in the network. This can occur when some species perform a central role in the microbial ecosystem, for example when a hub-species produces a metabolite that is required for growth by many other species.



Appendix Figure 2.5 - The effect of a_{ij} on the correlations between the abundances of two bacterial species for different interaction mechanisms.

The two species can co-exist under certain combinations of a_{ij} (Box 2.1). The abundances of the two species N_1 and N_2 at equilibrium are shown as scatterplots and were obtained by running the two-species Lotka-Volterra model with $K_1 = 1.5$, $K_2 = 1.1$, $r_1 = 1$, $r_2 = 2$, and a_{ij} drawn randomly from normal distributions with different combinations of means and standard deviations.

In A) the two distributions have different means and standard deviations: $a_{12} \sim N([0.5], 0.1)$ and $a_{21} \sim N([0.7], 0.2)$. In B) the distributions have identical means, but different standard deviations: $a_{12} \sim N([0.5], 0.2)$ and $a_{21} \sim N([0.5], 0.1)$. For exploitative interactions we also show the situations that negative correlations can occur when the exploitative benefit displays much more variation than the harm to the other species, i.e., $a_{12} \sim N(-0.5, 0.01)$ and $a_{21} \sim N(0.5, 0.2)$ for exploitative interaction type 1, and $a_{12} \sim N(0.5, 0.2)$ and $a_{21} \sim N(-0.5, 0.01)$ for exploitative interaction type 2. In C) interactions are randomly drawn from distributions with different means and identical standard deviations: $a_{12} \sim N([0.6], 0.1)$ and $a_{21} \sim N([0.3], 0.1)$. In D) the interactions have identical strengths for the two species, namely $|a_{12}| = |a_{21}|$. The mutualistic interactions are drawn from the distribution $a_{12} = a_{21} \sim U(0, 2.5)$, for competition and exploitative interactions we show two different scenarios, namely $|a_{12}| = |a_{21}| \sim U([0.4], [2.5])$ (upper graph) and $|a_{12}| = |a_{21}| \sim U(0, [0.4])$ (lower graph). Because the two species have different carrying capacities, the two situations of exploitative interactions are different. The grey polygon indicates the area where coexistence is possible. Note that the ranges of the axes are different in each subplot.