Syntrophic interaction in microbial communities

Léo Buchenel

April 3, 2020

Contents

Chapter 1

Introduction

1.1 Consumer Resource Models in microbial ecology

Biology and Physics have always been tightly intertwined. Especially the years following the end of World War II saw many famous physicists getting interested in the blooming field of Biology [1], Leo Szilard or Erwin Schroedinger and his *What is Life? The Physical Aspect of the Living Cell* [2] among others. That exodus is no surprise, many biological phenomena at different scales are well modelled with Physics weaponry: from the use of Statistical Physics to solve protein folding problems [3] and find phase transitions in ecological communities [4] to the application of Hamiltonian dynamics to describe the movement of starling flocks [5].

However, Physics has not solved every problem yet: the study of microbial communities remains one of the biggest and most interesting challenges of contemporal microbiology. Indeed microbes and their complex interactions have a substantial, non trivial and very large impact on humans and their environment in various ways: we only start to understand the role of microbiological interactions in verterbrates' guts [6], or how they shape our soils [7] and oceans [8].

Population dynamics in ecological communities are often approximated by variations of the Lotka-Volterra model [9]. This approach works well when the mediators of the competitive interaction between species reach a steady state fast enough such that their own dynamics can be eliminated [10]. However, such an assumption is not always true and one must in general always ask themselves whether it may be applied [11]. For microbial communities, previous literature shows that the population dynamics are not always well captured by a Lotka-Volterra model [10], which explains the need of a more mechanistic approach, where the dynamics of both the microbes and their resources are explicitly modelled. Robert MacArthur is one of the first ecologists to establish and study such a *Consumer Resource Model* (CRM) [12], launching a field still active today [13].

In the light of recent developments in the microbiology literature [14], we propose here a CRM[1] which explicitly takes into account syntrophy. This process, which is largely ob-

[1] One could argue that Flux Balance Analysis (FBA) [15] would be well suited for such a study. We ruled it out because it is known to scale badly [16] with system size and we do not want to be hindered by this limitation.

served in microbial communities [14], by definition occurs when microbes release, through a metabolic process, byproducts that are consumed by some members of the microbial community. In short, we want to know what happens when consumers are also allowed to release resources.

1.2 Establishing the model and goals

We write down a CRM which describes the coupled evolution of the biomass of N_S different species and their N_R resources in a chemostat[2]. Resources are labelled $\mu = 1, \dots, N_R$ and consumers $i = 1, \dots, N_S$. The coupled time evolution of their respective abundances $\{R_\mu, S_i\}$ is given by:

$$\boxed{\begin{aligned} \frac{dR_\mu}{dt} &= l_\mu - m_\mu R_\mu - \sum_j \gamma_{j\mu} R_\mu S_j + \sum_j \alpha_{\mu j} S_j && \text{(1.1a)} \\ \frac{dS_i}{dt} &= \sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu S_i - d_i S_i - \sum_\nu \alpha_{\nu i} S_i && \text{(1.1b)} \end{aligned}}$$

The set of quantities $\{l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i\}$ has no explicit dynamics and is taken as constant. On the other hand, $\{R_\mu, S_i\}$ may dynamically evolve and will be refered to as *dynamical variables*. Note that there are in this model a lot of different symbols that may be easy to confuse. We will at least try to keep the following conventions:

- Quantities related to resources have subscripts in greek alphabet (*e.g.* the resource μ has abundance R_μ). Quantities related to species have subscripts in latin alphabet (*e.g.* the species i has abundance S_i). Finally, quantities related to both have both indices.

- Vectors (*i.e.* quantities with one index) are written with the latin alphabet (*e.g.* the resource μ has death rate m_μ).

- Matrices (*i.e.* quantities with two indices, usually relating resources and species) are written with the greek alphabet (*e.g.* $\gamma_{i\mu}$ is the rate at which species i consumes resource μ).

Our model takes numerous phenomena into account and it may be helpful to take the time to explain the different terms of each differential equation. The temporal evolution of the biomass R_μ of a resource μ is essentially driven by the following processes:

- Constant external inflow coming from the experimental setup: this corresponds to the constant $+l_\mu$ term.

- Natural diffusion/deterioration at rate m_μ: this corresponds to the $-m_\mu R_\mu$ term.

[2]In a chemostat, new nutrients are continuously added, while at the same time microorganisms and resources are removed in order to keep the culture volume constant [17].

- Consumption by the species j at a rate $\gamma_{j\mu}$: $-\gamma_{j\mu}R_\mu S_j$. Summing up the contributions of every species, we get the Lotka-Volterra style [9] term $-\sum_j \gamma_{j\mu}R_\nu S_j$,
- Intrasystemic inflow coming from the syntrophy of species j at a rate $\alpha_{\mu j}$: $+\sum_j \alpha_{\mu j}S_j$.

On the other hand, biomass of species S_i changes because of the following processes:

- Consumption of resource R_ν at a rate $\gamma_{i\nu}$. Only a fraction $\sigma_{i\nu}$ of this is allocated to biomass growth: $+\sum_\nu \sigma_{i\nu}\gamma_{i\nu}R_\nu S_i$.
- Cell death/diffusion at rate d_i: this is the $-d_i S_i$ term.
- Syntrophic interaction : release of resource ν at rate $\alpha_{\nu i}$. In total $-\sum_\nu \alpha_{\nu i}S_i$.

The aim of the project is to study equilibria points of this model and their stability. In particular, we are interested in how syntrophy changes the robustness of the equilibria.

1.2.1 Equilibria of the model

We say that abundances $\{R^*_\mu, S^*_j\}$ are an *equilibrium*[3] of our model if they are fixed points of it, that means if the following equations are fulfilled:

$$\begin{cases} 0 = l_\mu - m_\mu R^*_\mu - \sum_j \gamma_{j\mu}R^*_\mu S^*_j + \sum_j \alpha_{\mu j}S^*_j & \text{(1.2a)} \\ 0 = \sum_\nu \sigma_{i\nu}\gamma_{i\nu}R^*_\nu S^*_i - d_i S^*_i - \sum_\nu \alpha_{\nu i}S^*_i & \text{(1.2b)} \end{cases}$$

1.2.2 Attack strategy and important notions

Before jumping right into the matter, it is important to explain how we will study this system of differential equations. Mainly two different but complimenteray approaches will be used: analytical and numerical. Note that the $\sim$ 5'000 lines of code we wrote from scratch and that we use to get the results of Section 3 are available at the address `https://gitlab.ethz.ch/palberto/consumersresources.git`.

Metaparameters and matrix properties

Studying the equilibria of our CRM will lead us to establish and study several relations involving the different *parameters* of the problem. Namely, these are: $l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i, R^*_\mu$ and S^*_i $\forall i = 1, \dots, N_S; \mu = 1, \dots, N_R$. We define the *parameters space* $\mathcal{P}$ as the space that contains all the parameters:

$$\mathcal{P} \equiv \left\{p : p = (l_\mu, m_\mu, d_i, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, R^*_\mu, S^*_i)\right\} \tag{1.3}$$

Without taking into account the constraints on these parameters, there are $3N_R + 2N_S + 3N_R N_S$ free parameters, so $\mathcal{P} \simeq \mathbb{R}_+^{3N_R+2N_S+3N_R N_S}$. Our goal is to study microbial communities with a large number of consumers and resources, typically $N_R, N_S \approx 25, 50, 100, \dots$

[3]For the sake of brevity, we will sometimes drop the μ and j subscripts and simply write $\{R^*, S^*\}$.

i.e. $\mathcal{P} \simeq \mathbb{R}^{\sim 2000}$. It is clear that a precise study on each one of the 2000 elements is way too tenuous of a job. Another, simpler, approach is needed.

We decide to look at the problem from a statistical point of view, *i.e.* we write a matrix $q_{i\mu}$ as [18]:

$$q_{i\mu} = \mathfrak{Q} Q_{i\mu} \tag{1.4}$$

where $\mathfrak{Q}$ is a random variable of mean Q_0 and standard deviation σ_Q. $Q_{i\mu}$ is a binary matrix that, if interpreted as an adjacency matrix, tells about the network structure of the quantity $q_{i\mu}$.
We apply this way of thinking to the parameters of our problem, namely we write:

$$\begin{cases} l_\mu = \mathfrak{L} & \text{(1.5a)} \\ m_\mu = \mathfrak{M} & \text{(1.5b)} \\ \gamma_{i\mu} = \mathfrak{G} G_{i\mu} & \text{(1.5c)} \\ \alpha_{\mu i} = \mathfrak{A} A_{\mu i} & \text{(1.5d)} \\ \sigma_{i\mu} = \mathfrak{S} & \text{(1.5e)} \\ d_i = \mathfrak{D} & \text{(1.5f)} \\ R^*_\mu = \mathfrak{R} & \text{(1.5g)} \\ S^*_i = \mathfrak{S} & \text{(1.5h)} \end{cases}$$

Note that we do not add any explicit topological structure on $l_\mu, m_\mu, d_i, R^*_\mu, S^*_i$ and $\sigma_{i\mu}$ because we require these to always be larger than zero. In particular, we require positive-valued equilibria [19].

In order to make computations analytically tractable, we require the standard deviation on the parameters involved in the problem to be small, *i.e.* not larger than typically 10%. In that regime, every random variable $\mathcal{Q}$ is well approximated by its average value Q_0. We call Q_0 a *metaparameter*. While studying things analytically we will hence often come back to the following approximation:

$$\begin{cases} l_\mu \approx l_0 & \text{(1.6a)} \\ m_\mu \approx m_0 & \text{(1.6b)} \\ \gamma_{i\mu} \approx \gamma_0 G_{i\mu} & \text{(1.6c)} \\ \alpha_{\mu i} \approx \alpha_0 A_{\mu i} & \text{(1.6d)} \\ \sigma_{i\mu} \approx \sigma_0 & \text{(1.6e)} \\ d_i \approx d_0 & \text{(1.6f)} \\ R^*_\mu \approx R_0 & \text{(1.6g)} \\ S^*_i \approx S_0 & \text{(1.6h)} \end{cases}$$

This simplification is mathematically equivalent to collapsing the parameter space $\mathcal{P}$ to a lower dimensional space. Formally that lower dimensional space is the Cartesian product of $\mathcal{M}$ and $\mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$, where $\mathcal{M}$ is the *metaparameters space*:

$$\mathcal{M} \equiv \{m : m = (l_0, m_0, d_0, \gamma_0, \alpha_0, \sigma_0, R_0, S_0)\} \tag{1.7}$$

and $\mathcal{B}_{N\times M}$ is the set of binary matrices of dimensions $N \times M$. To summarize, we simply designed a *collapsing procedure* $\mathcal{C} : \mathcal{P} \to \mathcal{M} \times \mathcal{B}_{N_S\times N_R} \times \mathcal{B}_{N_R\times N_S}$ in order to simplify our problem.

Mathematically, when we do analytical computations, we mostly work in the collapsed space $\mathcal{C}(\mathcal{P})$ because it reduces the number of parameters from $3N_R + 2N_S + 3N_RN_S$ (continous) to 8 (continous) + $3N_RN_S$ (binary). And to make the problem even simpler, instead of looking at each entry of the binary matrices G and A individually, we will consider only some globally defined quantities of these matrices. For a matrix M_{ij} the metrics interesting to us are most of all:

- its **nestedness**[4]: this measures how "nested" the system is, *i.e.* if there are clusters grouped together[5]. It is known [21, 18] that nestedness can play a profound role in the dynamics of ecological communities. Although it is somewhat controversed [22], we will keep the definition of the nestedness $\eta(M)$ of a binary matrix M as it was used in [21]:

$$\eta(M) \equiv \frac{\sum_{i<j} n_{ij}}{\sum_{i<j} \min(n_i, n_j)} \tag{1.8}$$

  where the number of links n_i is simply the degree of the i-th row of M

$$n_i \equiv \sum_k M_{ik}, \tag{1.9}$$

  and n_{ij} is the overlap matrix defined as

$$n_{ij} \equiv \sum_k M_{ik}M_{jk}. \tag{1.10}$$

- its **connectance**: this measure, simply defined as the ratio of non-zero links in a matrix, is central in the study of plants-and-animals systems [18]. It is formally defined for a matrix q_{ij} of size $N \times M$ as:

$$\kappa(q) \equiv \frac{\sum_{ij} Q_{ij}}{NM} \tag{1.11}$$

  where Q is the (binary) network adjacency matrix of q.

Losing complexity – how to gain it back

As explained above, the idea is to simplify the study of a system with a large number of parameters to a system with a manageable number of so-called "metaparameters". Of course, collapsing a very high dimensional space to a low-dimensional space makes us lose information. Losing some information – and hence complexity – is desired when

[4]For the matrix consumption G, we will call it especially the "ecological overlap".

[5]In typical Lotka-Volterra models, where only species-species interactions are considered, *e.g.* [20], measuring the nestedness of the γ consumption matrix would be in the same spirit as counting how many niches there are in the community.

doing analytical computations but it is not when we want to produce precise and detailed numerical results.

So, how do we bridge the gap between what we work with analytically, *i.e.* a set of metaparameters and binary matrices, to precise measurements of quantities defined in our model Eq.(1.1)? The answer is simple: we define a function

$$\mathcal{A} : \mathcal{M} \times \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S} \to \mathcal{P} \tag{1.12}$$

which brings us from the collapsed space to the parameter space[6]. Numerically, from a set of metaparameters $m \in \mathcal{M}$ and binary matrices $B = (G, A) \in \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$, we produce a (or several) set(s) of parameters $p = \mathcal{A}(m, B) \in \mathcal{P}$ and study properties of it. Section 2.1.4 details how $\mathcal{A}$ is constructed.

[6]Note that since the collapsed space is lower dimensional than the parameters space, $\mathcal{A}$ is not the inverse of $\mathcal{C}$.

Chapter 2

Methods

2.1 Feasibility

We want to be able to build feasible models numerically, *i.e.* we would like to generate a set of constant numbers $\{l_\nu, m_\nu, R^*_\nu, S^*_j, \gamma_{j\nu}, \alpha_{\nu j}, \sigma_{j\nu}\}$ such that the equilibria equations Eqs.(**??**) are fulfilled.

2.1.1 Basic concepts

Since its very inception [23], the study of ecological interactions has been and still is tightly close to the one of random matrices [24, 25, 26]. Usually, the procedure is assuming we are at a feasible equilibrium point, where some matrix of the model (*e.g.* the species-interaction matrix or the jacobian) is approximated as random, and then study the dynamical stability of said feasible point.

This framework is not satisfying for the study we would like to conduct, because the question "does a given set of random parameters lead to a feasible system?" is not trivial at all. Indeed for our model to make sense, we impose two conditions on any system deemed as feasible : the model parameters must be "biological" and biomass must be conserved.

Asking for the model parameters to be biological simply means we want them to have the intended biological interpretation. This means *e.g.* that any syntrophic interaction has to be non-negative $\alpha_{\mu i} \geqslant 0$ otherwise it cannot be interpreted as a syntrophic interaction anymore. More generally this is equivalent to requiring that all the model parameters are non-negative:

$$p \geqslant 0 \ \forall p \in \mathcal{P}. \tag{2.1}$$

In our study, this equation will be slightly restricted since we are looking for positive-valued equilibria, so we require $R^*_\mu, S^*_i > 0$ specifically for these two parameters. Also, we require also a non-zero efficiency[7]. Finally every resource feeding rate should be non-zero in order to avoid resource depletion and every resource and consumer must eventually die out in the absence of interaction. In the end this means we require:

$$R^*_\mu, S^*_i, \sigma_{i\mu}, l_\mu, d_i, m_\mu, \sigma_{i\mu} > 0 \text{ and } \gamma_{i\mu}, \alpha_{\mu i} \geqslant 0. \tag{2.2}$$

[7]It wouldn't make sense to say that species i eats resource μ with efficiency 0, since this is equivalent to species i not eating resource μ, and this is already encoded in the network structure.

Remember that not all the parameters of our models are free : there are $3N_R + 2N_S + 4N_RN_S$ parameters constrained by $N_R + N_S$ equations. So if we set $2N_R + N_S + 4N_RN_S$ parameters, the remaining $N_R + N_S$ are not free but set by the equilibrium equations **Insert ref equation**. Traditionally, we would solve for R^* and S^* and choose the rest of the parameters, but for reasons explained in **insert ref**, we will solve for the consumers death rates d_i and the resources diffusion rate m_μ. This means that if we *choose* non-negative $\gamma, \alpha, \sigma, \tau, l, R^*$ and S^*, Eq.(2.2) is equivalent to :

$$\begin{cases} d_i = \sum_\nu \left(\sigma_{i\nu}\gamma_{i\nu}R_\nu - \alpha_{\nu i}\right) > 0 \ \forall i = 1, \dots, N_S & \text{(2.3a)} \\ m_\mu = \dfrac{l_\mu - \sum_j \left(\gamma_{j\mu}R_\mu - \alpha_{\mu j}\right) S_j}{R_\mu} > 0 \ \forall \mu = 1, \dots, N_R & \text{(2.3b)} \end{cases}$$

In addition to Eqs.(2.3), we want any feasible system to conserve biomass *at equilibrium*[8]. This means no species should be able to produce more biomass than it physically can. More specifically, a consumer i attains, from consuming resources, a total biomass of $\sum_\nu \gamma_{i\nu}R^*_\nu S^*_i$. From this available biomass, only a part $\sum_\nu \sigma_{i\nu}\gamma_{i\nu}R^*_\nu S^*_i$ is devoted to growth. From the remaining $\sum_\nu (1-\sigma_{i\nu})\gamma_{i\nu}R^*_\nu S^*_i$, a part $\sum_\nu \alpha_{\nu i}S^*_i$ is given back to the resources as a syntrophic interaction. We simply impose that the syntrophic interaction is smaller than or equal to the available remaining biomass :

$$\sum_\nu (1-\sigma_{i\nu})\gamma_{i\nu}R^*_\nu \geqslant \sum_\nu \alpha_{\nu i} \ \forall i = 1, \dots, N_S. \tag{2.4}$$

From now on, we will say that a parameter set p is *feasible* if it satisfies Eqs.(2.3) and (2.4). This is completely deterministic, in the sense that for a given parameters set $p \in \mathcal{P}$ one can without a doubt say whether it is feasible or not. We can hence define the *parameters set feasibility function* $\mathfrak{F} : P \to \{0, 1\}$, which takes a parameter set as an input and tells you whether this parameter set is feasible or not:

$$\mathfrak{F}(p) = \begin{cases} 1 \text{ if } p \text{ is feasible,} \\ 0 \text{ else.} \end{cases} \tag{2.5}$$

However as explained above we will usually not work with a parameter set $p \in \mathcal{P}$ directly – because there are too many variables to keep track of – but with a metaparameter set $m \in \mathcal{M}$ and a binary consumption matrix $G \in \mathcal{B}_{N_R \times N_S}$ instead. We can similarly define a *metaparameters set feasibility function* $\mathcal{F} : \mathcal{M} \to [0, 1] \times \mathcal{B}_{N_R \times N_S}$ which is the probability that a given set of metaparameters $m \in \mathcal{M}$ coupled with binary matrices $B = (G, A)$ gives rise – through the algorithmic procedure $\mathcal{A}$ – to a feasible parameter set :

$$\boxed{\mathcal{F}(m, B) = \text{Probability}\left\{\mathfrak{F}\left(\mathcal{A}(m, B)\right) = 1\right\}} \tag{2.6}$$

We will in general work with $\mathcal{F}$ rather than $\mathfrak{F}$ because it is easier to handle metaparameters. In practice $\mathcal{F}(m, B)$ is estimated numerically by generating N parameters sets from (m, B)

[8]This weak condition should hold only at equilibrium : we allow transition periods where biomass may not be conserved.

and calculating the number of feasible ones :

$$\mathcal{F}(m,B) = \lim_{N\to\infty} \sum_{i=1}^{N} \frac{\mathfrak{F}(\mathcal{A}(m,B))}{N} \approx \sum_{i=1}^{N} \frac{\mathfrak{F}(\mathcal{A}(m,B))}{N} \text{ for } N \gg 1. \tag{2.7}$$

2.1.2 The feasibility volume

The algorithmic procedure detailed in Section **??** explains how feasible systems can be built. However, it implies that we first found a combination of metaparameters that will most of the time lead to the realisation of feasible systems when they are taken as an input of the algorithm.

Overall we have six metaparameters that we can play with : γ_0, α_0, l_0, σ_0, S_0 and R_0. However, following the analysis of [26], we notice that our system (**??**) is arbitrary on some level. Indeed we have a "scale freedom", that means we decide in which set of units we work. There are two physical quantities at stake here : biomass and time, and we may choose, however we want it, a specific set of units describing both of them.

We will measure biomass in units of the average resource abundance at equilibrium[9], that means :

$$\langle R_\mu \rangle = R_0 = 1. \tag{2.8}$$

Similarly, we will measure time such that the average external resource uptake rate is one, that is :

$$\langle l_\mu \rangle = l_0 = 1. \tag{2.9}$$

After this manipulation, our number of metaparameters is reduced from six to four : only γ_0, S_0, α_0 and σ_0 remain.

For the sake of simplicity, we will keep the same σ_0 throughout our whole study. We take a value close to the efficiency of real microbial systems [**insert ref**], that is $\sigma_0 = 0.25I$.

Overall, we need to choose the last three remaining metaparameters: α_0, γ_0 and S_0. As soon if we choose γ_0 and S_0, we will get a range of α_0 which will give rise to feasible systems. We will then choose γ_0 and S_0 such that they lead to feasible systems for every consumption matrix considered here **when there is no syntrophy**, *i.e.* $\alpha_0 = 0$. We will then study the impact of varying α_0 at those values of γ_0 and S_0.

Formally, we can define for a consumption adjacency matrix G coupled with a syntrophy adjacency matrix A the x-feasible volume $\mathcal{V}_x^{G,A}$ of the metaparameters space $\mathcal{M}$ that will lead to at least a ratio x of feasible systems *i.e.* :

$$\mathcal{V}_x^{G,A} \equiv \{m \in \mathcal{M} : \mathcal{F}(m,(G,A)) \geqslant x\}. \tag{2.10}$$

It is clear that $\mathcal{V}_0^{G,A} = \mathcal{M}\ \forall G$ and $\text{Vol}\left(\mathcal{V}_x^{G,A}\right) \leqslant \text{Vol}\left(\mathcal{V}_y^{G,A}\right)\ \forall x > y, G$. We can similarly define for a set $S = \{(G_1,A_1),(G_2,A_2),\ldots,(G_N,A_N)\}$ of N couples of matrices their *common feasibility* volume $\mathcal{V}_x^S$, which is the region of the metaparameters space where feasibility is at least x for every couple in the set:

$$\mathcal{V}_x^S \equiv \bigcap\nolimits_{(G,A)\in S} \mathcal{V}_x^{G,A}. \tag{2.11}$$

[9]Note that this is not a completely innocent choice. Indeed we will see later that the matrix $\alpha_{\nu i} - \gamma_{i\nu} R_\nu^*$ is a crucial quantity here. Setting $\langle R^* \rangle = 1$ allows us to simply study the impact of γ against α instead of the more complicated γR^* versus α.

We also define for a matrix set S, its critical feasibility $f^*(S)$, which is the largest feasibility we can get while still having a non-zero common volume :

$$f^*(S) \equiv \max_{x\in[0,1]} \left\{x : \text{Vol}\left(\mathcal{V}_x^S\right) > 0\right\}. \tag{2.12}$$

For actual computations, we will choose a matrix set S_M, stick to it during the whole thesis, and work in its critical feasibility volume $\mathcal{V}^*$, defined as :

$$\boxed{\mathcal{V}^* \equiv \mathcal{V}^{S_M}_{f^*(S_M)}.} \tag{2.13}$$

2.1.3 Estimating the fully feasible volume $\mathcal{V}_1^{G,A}$

Now that we defined the x-feasible volume of a given couple consumption-syntrophy network (G, A) in Eq.(2.10), the goal is to get an estimate of the fully feasible volume $\mathcal{V}_1^{G,A}$ in terms of the metaparameters, such that we can directly and efficiently run simulations in the feasible region.

Biomass conservation

As stated before, we require that biomass is conserved in our model. This was equivalent to fulfilling Eq.(2.4), which we rewrite here :

$$\sum_\nu (1-\sigma_{i\nu})\,\gamma_{i\nu} R^*_\nu \geqslant \sum_\nu \alpha_{\nu i}\ \forall i = 1,\dots,N_S. \tag{2.14}$$

The idea is to neglect the variance of every quantity involved *i.e.* we use the approximation

$$\gamma_{i\mu} \approx \gamma_0 G_{i\mu},\ \alpha_{\mu i} \approx \alpha_0 A_{\mu i},\ \sigma_{i\nu} = \sigma_0 \text{ and } R^*_\nu \approx R_0. \tag{2.15}$$

This means the RHS of Eq.(2.4) is roughly given by

$$\sum_\nu \alpha_{\nu i} \approx \deg(A,i)\alpha_0 \tag{2.16}$$

where $\deg(A,i)$ is the degree of the i-th column of the α matrix : $\deg(A,i) = \sum_\nu A_{\nu i}$. Similarly,

$$\sum_\nu (1-\sigma_{i\nu})\,\gamma_{i\nu} R^*_\nu \approx (1-\sigma_0) R_0 \sum_\nu \gamma_{i\nu} \approx \deg(G,i)(1-\sigma_0)R_0\gamma_0, \tag{2.17}$$

Then energy conservation Eq.(2.4) is equivalent to

$$\deg(A,i)\alpha_0 \lessapprox \deg(G,i)(1-\sigma_0)R_0\gamma_0\ \forall i = 1,...,N_S \tag{2.18}$$

Since $\deg(G,i) > 0$ (it is the number of resources species i eats), we have:

$$\frac{\deg(A,i)}{\deg(G,i)}\alpha_0 \lessapprox (1-\sigma_0)R_0\gamma_0\ \forall i = 1,...,N_S \tag{2.19}$$

This is fulfilled if :

$$\boxed{\max_i \left\{\frac{\deg(A,i)}{\deg(G,i)}\right\}\alpha_0 \lessapprox (1-\sigma_0)R_0\gamma_0}. \tag{2.20}$$

Systems where the ratio $\frac{\#\text{resources released to}}{\#\text{resources consumed}}$ is small for each species allow for a larger individual syntrophy interaction (which is very intuitive).

Positivity of the parameters

As said before, the consumers death rates d_i have to be positive. This implied Eq.(2.3a), which may be recast as :

$$\sum_\mu \sigma_{i\mu}\gamma_{i\mu}R^*_\mu > \sum_\mu \alpha_{\mu i} \tag{2.21}$$

Using a reasoning similar to above, we get :

$$\boxed{\max_i \left\{\frac{\deg(A,i)}{\deg(G,i)}\right\}\alpha_0 \lessapprox \sigma_0 R_0 \gamma_0}. \tag{2.22}$$

Also, the resources diffusion rates m_ν need to be positive:

$$l_\nu + \sum_j \alpha_{\nu j} S^*_j > \sum_j \gamma_{j\nu} R^*_\nu S^*_j \ \forall \nu = 1, \ldots, N_R \tag{2.23}$$

Which is equivalent to

$$l_0 + k^\alpha_\nu \alpha_0 S_0 \gtrapprox \deg(G,\nu)\gamma_0 R_0 S_0 \ \forall \nu \tag{2.24}$$

Since $k^\gamma_\nu > 0$ (every resource is at least consumed by one species), the N_R equations above can be rewritten as:

$$\boxed{\min_\nu \left\{\frac{l_0}{\deg(G,\nu)S_0} + \frac{\deg(A,\nu)}{\deg(G,\nu)}\alpha_0\right\} \gtrapprox \gamma_0 R_0} \tag{2.25}$$

This says that systems where the ratio $\frac{\#\text{number of species that release to me}}{\#\text{number of species that consume me}}$ is large for every resource are more feasible. The strategy should be then to have γ's that have large $\deg(G,\nu)$ (*i.e.* resources are consumed by many species) and large $\deg(G,i)$ (*i.e.* species consume a lot of species), and the other way around for α (not sure about this for the last one).

Combining both conditions

The two upper bounds Eqs.(2.4)-(2.22) on α_0 can be combined in a single inequality :

$$\max_i \frac{\deg(A,i)}{\deg(G,i)}\alpha_0 \lessapprox \min(1-\sigma_0, \sigma_0)\gamma_0 R_0 \tag{2.26}$$

Note that when $\alpha_0 > 0$, we will trivially require that the syntrophy matrix is not empty, *i.e.* there exists at least an i for which $\deg(A,i) \geqslant 1$. Note also that the largest value $\deg(G,i)$ can get (for any i) is N_R. Hence,

$$\max_i \left\{\frac{\deg(A,i)}{\deg(G,i)}\right\} \geqslant \frac{1}{N_R}, \tag{2.27}$$

and we can find a largest allowed theoretical non-zero α_0 :

$$\alpha_0 \lessapprox \min(1-\sigma_0, \sigma_0)\gamma_0 R_0 N_R. \tag{2.28}$$

Finally, Eq.(2.26) and (2.25) can be combined into a single one, which gives us the volume of the metaparameters space that make the system feasible:

$$\boxed{\max_i \left\{\frac{\deg(A,i)}{\deg(G,i)}\right\}\alpha_0 \lessapprox \min(1-\sigma_0,\sigma_0)\gamma_0 R_0 \lessapprox \min\left(1-\sigma_0,\sigma_0\right) \min_\nu \left\{\frac{l_0}{\deg(G,\nu)S_0} + \frac{\deg(A,\nu)}{\deg(G,\nu)}\alpha_0\right\}} \tag{2.29}$$

2.1.4 Algorithmic procedure

We hereby detail the procedure used to numerically build feasible systems. It goes like this:

1. We first draw randomly each R_ν^* and S_i^* from a uniform distribution of mean equal to the corresponding metaparameter, *i.e.* :

$$R_\nu^* = \mathcal{R}\ \forall \nu = 1, \dots, N_R \text{ and } S_i^* = \mathcal{S}\ \forall i = 1, \dots, N_S, \tag{2.30}$$

   where $\mathcal{R}$ and $\mathcal{S}$ are random variables coming from a distribution of mean equal to the corresponding metaparameter and relative standard deviation[10] ϵ. In our simulations, we chose uniform distributions :

$$\mathcal{R} \sim \text{Unif}(R_0, \epsilon) \text{ and } \mathcal{S} \sim \text{Unif}(S_0, \epsilon). \tag{2.31}$$

2. The efficiency matrix $\sigma_{i\nu}$ is then drawn similarly, from a distribution with average σ_0. In order to simplify the problem[11], we will take a zero-variance à la Butler and O'Dwyer in [19], *i.e.* all species consume resources at the same global efficiency :

$$\sigma_{i\nu} = \sigma_0. \tag{2.32}$$

3. We build the consumption matrix $\gamma_{i\nu}$. Its adjacency matrix G is loaded through a user-provided file. While G gives the structure of γ, *i.e.* if a given $\gamma_{i\nu}$ is zero or not, the actual values of $\gamma_{i\nu}$ need then to be determined. They are drawn from a uniform distribution of mean γ_0 and relative standard deviation ϵ:

$$\gamma_{i\nu} = \text{Unif}(\gamma_0, \epsilon) G_{i\nu}. \tag{2.33}$$

4. We draw the resources external feeding rates, similarly to the other parameters :

$$l_\mu = \text{Unif}(l_0, \epsilon)\ \forall \mu = 1, \dots, N_R. \tag{2.34}$$

5. The last free parameter is the syntrophy matrix $\alpha_{\nu i}$, the d_i and l_μ are determined through the equations of evolution at equilibrium. This is the tricky part of the algorithm because α has to follow three constraints, namely energy conservation/dissipation Eq.(**??**) and positiveness of d_i and l_μ [**insert reference to equation**]. The general strategy is to choose the metaparameters in a way that these constraints should *almost always* be satisfied, *i.e.* we pick metaparameters that follow the feasibility constraint Eq.(**??**). The adjacency matrix A of α needs then to be specified. At the moment, it can be chosen in three different ways : fully connected, or in a way that no resource eaten by a given species can be released by that same species (*i.e.* $G_{i\mu} > 0 \iff A_{\mu i} = 0$) or by a user provided matrix. After the adjacency matrix is loaded, we can build α from a uniform distribution of mean α_0 and relative standard deviation ϵ:

$$\alpha_{\nu i} = \text{Unif}(\alpha_0, \epsilon) A_{\nu i}. \tag{2.35}$$

[10] By relative standard deviation, we mean the standard deviation measured in units of the average value.
[11] Indeed, a non uniform σ_0 introduces instability in the system

6. With all of these parameters drawn, we can solve Eq.(**??**) for the species death rate d_i.

7. We solve Eq.(**??**) for m_ν. All the parameters of the model are now fully determined.

8. We check if the constraints Eq.**(insert reference)** on the parameters are fulfilled. If they are not, we go back to step 1. Otherwise, we can exist the algorithm, a feasible system has been built.

2.2 Dynamical stability

As stated in the introduction, we are interested in the equilibria points of the set of coupled differential equations (1.1). In particular we want to know how *stable* a given equilibrium is. However the notion of *stability* itself is not trivial: what does it mean exactly that a system is stable under a given perturbation? How is a perturbation even defined? These questions have many different possible answers. Throughout this thesis different notions of stability will be tackled: the first is *dynamical stability*.

The main idea behind dynamical stability is simple. We want to answer the following question:

Given an equilibrium point $\{R_\mu^*, S_i^*\}$, does the system go back to a positive-valued equilibrium when the consumers and resources abundances are changed? If yes, how much can they be changed before the system evolves in such a way that it does not reach a positive-valued equilibrium?

2.2.1 Definitions

Local dynamical stability

The first notion to introduce is *local dynamical stability*: a system is said to be *locally dynamically stable* if it goes back to *its initial equilibrium point* $\{R_\mu^*, S_i^*\}$ after R_μ^* and S_i^* have been perturbed by an infinitesimal amount $\{\Delta R_\mu, \Delta S_i\}$. Because the perturbations are supposed to be very small, we can write:

$$\begin{pmatrix} \frac{dR_\mu(t)}{dt} \\ \frac{dS_i(t)}{dt} \end{pmatrix} = \begin{pmatrix} \frac{d\Delta R_\mu(t)}{dt} \\ \frac{d\Delta S_i(t)}{dt} \end{pmatrix} = J \begin{pmatrix} \Delta R_\mu(t) \\ \Delta S_i(t) \end{pmatrix} \tag{2.36}$$

where J is the *jacobian* of our system *i.e.* the jacobian matrix of its temporal evolution (1.1):

$$J \equiv \begin{pmatrix} \frac{\partial \dot{R}_\mu}{\partial R_\nu} & \frac{\partial \dot{R}_\mu}{\partial S_j} \\ \frac{\partial \dot{S}_i}{\partial R_\nu} & \frac{\partial \dot{S}_i}{\partial S_j} \end{pmatrix} = \begin{pmatrix} \left(-m_\mu - \sum_j \gamma_{j\mu} S_j\right) \delta_{\mu\nu} & -\gamma_{j\mu} R_\mu + \alpha_{\mu j} \\ \sigma_{i\nu}\gamma_{i\nu} S_i & \left(\sum_\nu \sigma_{i\nu}\gamma_{i\nu} R_\nu - d_i - \sum_\nu \alpha_{\nu i}\right) \delta_{ij} \end{pmatrix}, \tag{2.37}$$

where δ is the Kronecker delta symbol. Using the fact that we are only interested in equilibria where every resource is positive and Eq.(1.2b), this can be rewritten as:

$$J = \begin{pmatrix} \frac{l_\mu + \sum_j \alpha_{\mu j} S_j^*}{R_\mu^*} \delta_{\mu\nu} & -\gamma_{j\mu} R_\mu + \alpha_{\mu j} \\ \sigma_{i\nu}\gamma_{i\nu} S_i & \left(\sum_\nu \sigma_{i\nu}\gamma_{i\nu} R_\nu - d_i - \sum_\nu \alpha_{\nu i}\right) \delta_{ij} \end{pmatrix}, \tag{2.38}$$

We can then define for a given equilibrium point $\{R_\mu^*, S_i^*\}$ the *jacobian at equilibrium* J^* as the jacobian of said equilibrium. We will hence say that a given equilibrium is *locally dynamically stable* if its jacobian J^* is negative definite, *i.e.* if the largest real part of the eigenvalues of J^* is negative.

Note that since we are interested only in positive valued equilibria (*i.e.* $S_i^* > 0 \; \forall i$), then Eq.(1.2b) is equivalent to:

$$\sum_\nu \sigma_{i\nu}\gamma_{i\nu}R_\nu^* - d_i - \sum_\nu \alpha_{\nu i} = 0, \tag{2.39}$$

which means that the lower right block of the jacobian in Eq.(2.37) will be zero. Hence at equilibrium the jacobian J^* will have the following block form:

$$J^* = \begin{pmatrix} -\Delta & \Gamma \\ \mathrm{B} & 0 \end{pmatrix}, \tag{2.40}$$

where

- $\Delta_{\mu\nu} = \mathrm{diag}(m_\mu + \sum_j \gamma_{j\mu}S_j^*) = \mathrm{diag}\left(\frac{l_\mu + \sum_j \alpha_{\mu j}S_j^*}{R_\mu^*}\right)$ is a positive $N_R \times N_R$ diagonal matrix,
- $\Gamma_{\mu j} = -\gamma_{j\mu}R_\mu^* + \alpha_{\mu j}$ is a $N_R \times N_S$ matrix which does not have entries with a definite sign.
- $\mathrm{B}_{i\nu} = \sigma_{i\nu}\gamma_{i\nu}S_i^*$ is a $N_S \times N_R$ matrix with positive entries.

The locally dynamically stable volume $\mathcal{D}_{L,x}^{G,A}$

Similarly to what was conducted in Methods **??**, one can define the *parameters set local dynamical stability function* $\mathfrak{D}_L : \mathcal{P} \to \{0, 1\}$, which tells you whether a given set of parameters $p \in \mathcal{P}$ is locally dynamically stable or not:

$$\mathfrak{D}_L(p) \equiv \begin{cases} 1 \text{ if } p \text{ is locally dynamically stable} \\ 0 \text{ else.} \end{cases} \tag{2.41}$$

Note that trivially p has to be feasible in order to be locally dynamically stable:

$$\mathfrak{D}_L(p) = 1 \implies \mathfrak{F}(p) = 1. \tag{2.42}$$

Again, we also define the *metaparameters set local dynamical stability function* $\mathcal{D}_L : \mathcal{M} \times \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S} \to [0, 1]$ which tells you, given a set of metaparamters $m \in \mathcal{M}$ and a consumption-syntrophy network $B = (G, A)$ the chance that you draw a locally dynamically stable set of parameters:

$$\boxed{\mathcal{D}_L(m, B) \equiv \text{Probability}\left\{\mathfrak{D}_L(\mathcal{A}(m, B)) = 1\right\}}. \tag{2.43}$$

We also define the x locally dynamically stable (lds) volume $\mathcal{D}_{L,x}^{G,A}$ by the region of the metaparameters space that gives rise to a percentage of x dynamically stable systems:

$$\mathcal{D}_{L,x}^{G,A} \equiv \{m \in \mathcal{M} : \mathcal{D}_L(m, (G, A)) \geqslant x\} \tag{2.44}$$

Clearly, $\mathcal{D}_{L,0}^{G,A} = \mathcal{M}$, $\mathrm{Vol}\left(\mathcal{D}_{L,x}^{G,A}\right) \leqslant \mathrm{Vol}\left(\mathcal{D}_{L,y}^{G,A}\right) \; \forall x \geqslant y$, and more importantly, Eq.(2.42) is equivalent to $\mathcal{D}_{L,x}^{G,A} \subset \mathcal{V}_x^{G,A}$. We can also define for a set of N couples of matrices $S = \{(G_1, A_1) \ldots, (G_N, A_N)\}$ their common x (lds) volume $\mathcal{D}_{L,x}^S$:

$$\mathcal{D}_{L,x}^S \equiv \bigcap\nolimits_{(G,A)\in S} \mathcal{D}_{L,x}^{G,A}. \tag{2.45}$$

For such a set S we define also its critical local dynamical stability $d_L^*(S)$ which is the largest local dynamical stability we can get while still having a non-zero common volume:

$$d_L^*(S) = \max_{x\in[0,1]} \{x : \text{Vol}(D_{L,x}) > 0\}\,. \tag{2.46}$$

Finally the critical common local dynamical stability volume $\mathcal{D}_L^S$ is the common lds volume at critical lds:

$$\mathcal{D}_L^* \equiv \mathcal{D}_{L,d_L^*(S)}^S. \tag{2.47}$$

Global dynamical stability

If we established that a system is locally dynamically stable, we know that it will come back to the same equilibrium after an infinitesimal perturbation of the resources and consumers abundances. The next natural question is: *how much* can these be perturbed before the system goes to a point where either at least a species has gone extinct or reaches another positive valued equilibrium $\{\tilde{R}_\mu^*, \tilde{S}_i^*\}$ or simply does not reach a new dynamical equilibrium? One way of studying this [18] is to simply take an equilibrium point $\{R_\mu^*, S_i^*\}$ and perturb the abundance of the species and resources at that point by a fixed number $\Delta_D \in [0,1]$ which allows us to quantify the perturbation:

$$\begin{cases} R_\mu^* \to R_\mu(t_0) \equiv R_\mu^*\,(1+\Delta_D\nu_\mu)\,, & (2.48)\\ S_i^* \to S_\mu(t_0) \equiv S_i^*\,(1+\Delta_D\nu_i)\,, & (2.49) \end{cases}$$

where the $\nu_{\mu,i}$ are random numbers drawn from a uniform distribution between -1 and +1 and t_0 is the time where the previously at equilibrium system is perturbed. This is basically the same procedure as local dynamical stability except we allow the perturbation Δ_D to be non-zero. The question we will often ask is precisely how big can Δ_D be. The system with the initial values $\{R(t_0), S(t_0)\}$ can then be time evolved from $t = t_0$ until it reaches an equilibrium $\{\tilde{R}^*, \tilde{S}^*\}$ which may be different from the equilibrium $\{R^*, S^*\}$ initially considered.

A certain number of quantities, that all depend on the perturbation Δ_D, can then be measured to quantify the dynamical stability of the system:

- The resilience t_R: this is the time scale over which the system reaches its new equilibrium.
- The number of extinctions E: this is the number of species or resources which died during the time it took the system to reach its new equilibrium.
- The angle α between two equilibria: this quantifies how close the old and new equilibria are. α is defined through its standard scalar product formula:

$$\cos(\alpha) \equiv \frac{\sum_\mu R_\mu^*\tilde{R}_\mu^* + \sum_j S_j^*\tilde{S}_j^*}{\sqrt{\sum_\mu \left(R_\mu^*\right)^2 + \sum_i \left(S_i^*\right)^2}\sqrt{\sum_\mu \left(\tilde{R}_\mu^*\right)^2 + \sum_i \left(\tilde{S}_i^*\right)^2}}. \tag{2.50}$$

These quantities have either been already introduced in previous papers or are natural extensions of standard quantities [27, 18]. They allow us to quantify the robustness of a given equilibrium.

2.2.2 The quest for a full solution

The question of global dynamical stability is mathematically tedious, so we start by focusing on local dynamical stability.

Here we aim to find the spectrum of the jacobian at equilibrium, which will tell us whether the system is locally dynamically stable or not (see above).

How to determine local dynamical stability

As explained, we need to determine the sign of the largest real part of all the eigenvalues of J^*. More precisely, mathematically we are interested in the real part of λ_1, which itself is defined by the following property:

$$\forall \lambda \in \sigma(J^*),\ \mathrm{Re}\,(\lambda) \leqslant \mathrm{Re}\,(\lambda_1)\,, \tag{2.51}$$

where $\sigma(J^*)$ is the set of eigenvalues of J^*, called the *spectrum* of J^*. The sign of λ_1, will govern the local stability of the system at equilibrium. Namely **Need to add resource?**:

- $\mathrm{Re}\,(\lambda_1) < 0$: any perturbation on the abundances will be exponentially supressed. The system is stable.
- $\mathrm{Re}\,(\lambda_1) > 0$: any perturbation on the abundances will be exponentially amplified. The system is unstable.
- $\mathrm{Re}\,(\lambda_1) = 0$: a second order perturbation analysis is required to assess the system local stability. We will call such systems *marginally stable* [28].

The master equation for local dynamical stability

In order to get $\mathrm{Re}\,(\lambda_1)$, we have to get the full spectrum of J^*, as sadly easier standard techniques like the Perron-Frobenius theorem [29] cannot be applied. The eigenvalues of J^* are obtained through the eigenvalue problem:

$$\det\left(J^* - \lambda\right) = 0. \tag{2.52}$$

More explicitly, using Eq.(2.40), we state the *master equation for local dynamical stability*:

$$\boxed{\det\begin{pmatrix} -\Delta - \lambda & \Gamma \\ \mathrm{B} & 0 - \lambda \end{pmatrix} = 0} \tag{2.53}$$

Before proceeding any further, we first eliminate systems where local dynamical stability cannot be decided with the first order perturbation analysis that we are conducting, *i.e.* marginally stable equilibria.

Marginally stable equilbria

We want to avoid the case $\mathrm{Re}\,(\lambda_1) = 0$, because it is way harder to handle mathematically. To make things easier, we will in our analytical computations[12] get rid of all systems where 0 is part of the spectrum. Even though this is a harsh condition, we know that for such systems, local dynamical stability can be decided by the computations we conduct. $\lambda = 0$ is part of the spectrum if and only if it solves the master equation (2.53):

$$\det \begin{pmatrix} -\Delta & \Gamma \\ \mathrm{B} & 0 \end{pmatrix} = 0 \tag{2.54}$$

Using the fact that Δ is invertible, we can make use of the equality[13]:

$$\det \begin{pmatrix} -\Delta & \Gamma \\ \mathrm{B} & 0 \end{pmatrix} = \det(-\Delta)\det(\Gamma\Delta^{-1}\mathrm{B}) = (-1)^{N_R}\det(\Gamma\mathrm{B}). \tag{2.55}$$

Eq.(2.54) then becomes:

$$\det(\Gamma\mathrm{B}) = \det(\mathrm{B}\Gamma) = 0 \tag{2.56}$$

which means that $\Gamma\mathrm{B}$ and $\mathrm{B}\Gamma$ are not full rank, *i.e.* Γ or B is not full rank.

Non marginal equilibria

For now we will concentrate on equilibria that are clearly either stable or unstable[14], *i.e.* :

$$\lambda_1 \neq 0. \tag{2.57}$$

For the sake of simplicity, we will first look for the non zero solutions of the spectrum, *i.e.* for now we assume:

$$\lambda \neq 0. \tag{2.58}$$

This immediately implies

$$\det(\lambda) \neq 0, \tag{2.59}$$

where λ actually stands for the $N_S \times N_S$ identity matrix multiplied by a scalar λ. One can use this condition to simplify Eq.(2.53) using the properties of block matrices [30]:

$$\det \begin{pmatrix} -\Delta - \lambda & \Gamma \\ \mathrm{B} & 0 - \lambda \end{pmatrix} = \det(-\lambda)\det\left(-\Delta - \lambda + \frac{1}{\lambda}\Gamma\mathrm{B}\right). \tag{2.60}$$

Hence Eq.(2.53) becomes:

$$\boxed{\det\left(\lambda^2 + \Delta\lambda - \Gamma\mathrm{B}\right) = 0.} \tag{2.61}$$

The complexity here is already reduced because we go from the determinant of a $N_R + N_S$ square matrix to a N_R square matrix. We see from the previous expression that the

[12] For numerical computations we will get rid of marginally stable systems individually, meaning we may have some systems where $0 \in \sigma(J^*)$.

[13] This uses a formula which is trivially analogous to one found in [30].

[14] The case of marginally stable systems, where the maximum eigenvalue is zero, will be covered later.

dynamics is essentially dictated by the ΓB N_R-dimensional square matrix, which is given by:

$$(\Gamma \mathrm{B})_{\mu\nu} = \sum_i \Gamma_{\mu i} \mathrm{B}_{i\nu} = \sum_i \left(\alpha_{\mu i} - \gamma_{i\mu} R^*_\mu\right) \sigma_{i\nu} \gamma_{i\nu} S^*_i. \tag{2.62}$$

There are many strategies here to find regimes of stability. One is the so-called "Reductio ad absurdum", which is explored later in Methods 2.2.4.

2.2.3 Bounds on the eigenvalues

Gerschgorin's circle theorem

Gerschgorin's circle theorem [31] allows us to get a better idea of the location of the eigenvalues in the complex plane. It states that every eigenvalue of a $N \times N$ square matrix A is located in one of the N discs D_i defined by:

$$D_i \equiv \left\{ z \in \mathbb{C} : |z - A_{ii}| \leqslant \sum_{j \neq i} |A_{ij}| \right\}. \tag{2.63}$$

This can be reformulated as:

$$\sigma(A) \subset \bigcup_{i=1}^{N} D_i. \tag{2.64}$$

The geometrical interpretation is that the eigenvalues of a matrix deviate from the diagonal elements by a value bounded by the sum of the off-diagonal elements. It is then easy to see that if all the discs D_i are located to the left of the imaginary axis (*i.e.* the discs contain only numbers with a negative real part), then the eigenvalues of A are all negative. Geometrically this corresponds to the following lemma:

Lemma 1. *If for a matrix A the following equations are verified:*

$$\mathit{Re}\,(A_{ii}) + \sum_{j \neq i} |A_{ij}| < 0, \forall\ i, \tag{2.65}$$

then $\mathit{Re}\,(\lambda) < 0\ \forall \lambda\ \in \sigma(A)$.

Gerschgorin's circle theorem allows us to get a precious bound on the modulus of each eigenvalue and hence on the interesting one λ_1. Indeed we know that all eigenvalues of J^* will be located in one of the discs (as defined in Eq.(2.63)) of J^*. There are precisely $N_R + N_S$ discs of J^*, these are the "resources" discs:

$$D^R_\mu \equiv \left\{ z \in \mathbb{C} : |z + \Delta_\mu| \leqslant \sum_j |\Gamma_{\mu j}| \right\}\ \forall \mu = 1, \ldots, N_R, \tag{2.66}$$

and the "consumers" discs:

$$D^C_i \equiv \left\{ z \in \mathbb{C} : |z| \leqslant \sum_\nu |B_{i\nu}| \right\}\ \forall i = 1, \ldots, N_S. \tag{2.67}$$

The circle's theorem Eq.(2.64) tells us that all eigenvalues will be in the union of these circles, *i.e.* there exists $\forall \lambda \in \sigma\left(J^*\right)$ at least one μ^* or one i^* such that:

$$|\lambda| \leqslant \sum_{\nu} |B_{i^*\nu}| \tag{2.68}$$

or

$$|\lambda + \Delta_{\mu^*}| \leqslant \sum_{j} |\Gamma_{\mu^* j}| \tag{2.69}$$

Note furthermore that, because $\Delta_\mu > 0$ and $\mathrm{Im}\left(\Delta_\mu\right) = 0$, Eq.(2.69) implies

$$|\lambda| \leqslant \sum_{j} |\Gamma_{\mu^* j}| \,. \tag{2.70}$$

The only way both Eq.(2.68) and (2.70) are satisfied for all eigenvalues, and especially the one with the highest real part λ_1 is if they are bound by the maximum of both RHS of these equations. More precisely:

$$\boxed{|\lambda| \leqslant R_C \ \forall \lambda \in \sigma(J^*),} \tag{2.71}$$

where we defined the critical radius as:

$$R_C \equiv \max\left\{ \max_i \left\{ \sum_{\nu} |B_{i\nu}| \right\}, \max_{\mu} \left\{ \sum_{j} |\Gamma_{\mu j}| \right\} \right\}. \tag{2.72}$$

Intuitively, this means every eigenvalue must lie in a circle around the origin. The radius of this circle is given by whichever is larger between the largest column-sums of the B and Γ matrices.

Critical radius in terms of metaparameters

We would like to estimate R_C in terms of metaparameters, so that we can choose metaparameters that will give rise to systems in an LRI regime, which we know will be dynamically stable.

Using techniques very similar to previous computations:

$$\sum_{j} |\Gamma_{\mu j}| = \sum_{j} \left| \alpha_{\mu j} - \gamma_{j\mu} R^*_{\mu} \right| \approx \deg(\Gamma, \mu) \left| \alpha_0 - \gamma_0 R_0 \right| . \tag{2.73}$$

The difficult part is estimating $\deg(\Gamma, \mu) \approx \deg(A - G^T, \mu)$. If we assume that $\deg(A, \mu) \approx \deg(G^T, \mu) \ll N_S$ then we may use the very loose approximation

$$\deg(A - G^T, \mu) \approx \deg(A, \mu) + \deg(G, \mu). \tag{2.74}$$

In that regime we then have:

$$\max_{\mu} \left\{ \sum_{j} |\Gamma_{\mu j}| \right\} \approx \max_{\mu} \left\{ \deg(A, \mu) + \deg(G, \mu) \right\} \left| \alpha_0 - \gamma_0 R_0 \right| . \tag{2.75}$$

Similarly we find

$$\sum_{\nu} |B_{i\nu}| \approx \deg(G, i)\sigma_0\gamma_0 S_0, \tag{2.76}$$

such that R_C can be estimated roughly as:

$$\boxed{R_C \approx \max\left\{\max_i \left(\deg(G, i)\right)\sigma_0\gamma_0 S_0, \max_\mu \left(\deg(A, \mu) + \deg(G, \mu)\right)|\alpha_0 - \gamma_0 R_0|\right\}} \tag{2.77}$$

The main factor that will determine R_C (and hence the largest magnitude of any eigenvalue) is the structure of the food consumption matrix.

2.2.4 Low intra resources interaction (LRI) regime

Reductio ad absurdum

Now that we have a bound on how big the eigenvalues can be, we need to find strategies to find regimes where we *know* $\text{Re}\,(\lambda_1) < 0$, *i.e.* local dynamical stability is guaranteed. We inspire ourselves from the general idea of the mathematical proofs of [19]. We first rewrite Eq.(2.61) as[15]:

$$\det\left(-\Delta^{-1}\right)\det\left(-\Delta^{-1}\lambda^2 - \lambda + \Delta^{-1}\Gamma\mathrm{B}\right) = 0 \iff \boxed{\det\left(S(\lambda) - \lambda\right) = 0} \tag{2.78}$$

with

$$\boxed{S(\lambda) = \Delta^{-1}\Gamma\mathrm{B} - \Delta^{-1}\lambda^2}, \tag{2.79}$$

or, component-wise:

$$S_{\mu\nu} = \frac{1}{\Delta_\mu}\left[\left(\sum_i \Gamma_{\mu i}\mathrm{B}_{i\nu}\right) - \lambda^2\delta_{\mu\nu}\right] \tag{2.80}$$

The idea is to assume we are in an unstable regime, there exists at least $\lambda \in \sigma(J^*)$ which satisfies Eq.(2.61) and such that $\text{Re}\,(\lambda) > 0$. By Eq.(2.78), λ is also an eigenvalue of $S(\lambda)$. If we find conditions under which the real part of the spectrum of $S(\lambda)$ is entirely negative, we will know that $\text{Re}\,(\lambda) \leqslant 0$. As this is a contradiction to the hypothesis that the regime is unstable, we must conclude that the regime is stable[16].

Hence, the general strategy is to find regimes where we know that the spectrum of S, written as $\sigma(S)$, will be entirely negative for a positive λ.

Strong LRI regime

Theorem 1. *If a model is not marginally stable at equilibrium and it verifies:*

$$(\Gamma B)_{\mu\mu} < -\sum_{\nu\neq\mu}\left|(\Gamma B)_{\mu\nu}\right| - R_C^2 \;\forall\mu, \tag{2.81}$$

then it is dynamically stable.

[15] We can do this because since $m_\mu > 0$, we know Δ will always be invertible.

[16] Indeed, Eq.(2.61) assumes already that either $\text{Re}\,(\lambda_1) > 0$ or $\text{Re}\,(\lambda_1) < 0$.

Proof. We assume

$$(\Gamma \mathrm{B})_{\mu\mu} < -\sum_{\nu\neq\mu} \left|(\Gamma \mathrm{B})_{\mu\nu}\right| - R_C^2 \ \forall \mu. \tag{2.82}$$

This implies:

$$(\Gamma \mathrm{B})_{\mu\mu} + R_C^2 < -\sum_{\nu\neq\mu} \left|(\Gamma \mathrm{B})_{\mu\nu}\right| \ \forall \mu. \tag{2.83}$$

Using Eq.(2.71) and $\mathrm{Im}\,(\lambda)^2 \leqslant |\lambda|^2$, we get:

$$(\Gamma \mathrm{B})_{\mu\mu} + \mathrm{Im}\,(\lambda)^2 < -\sum_{\nu\neq\mu} \left|(\Gamma \mathrm{B})_{\mu\nu}\right| \ \forall \mu. \tag{2.84}$$

It is not difficult to prove that for any complex number:

$$\mathrm{Im}\,(c)^2 \geqslant -\mathrm{Re}\left(c^2\right) \ \forall c \in \mathbb{C}. \tag{2.85}$$

Using this result and dividing Eq.(2.90) by[17] Δ_μ, we get:

$$\frac{1}{\Delta_\mu}\left[\left(\sum_i \Gamma_{\mu i}\mathrm{B}_{i\mu}\right) - \mathrm{Re}\left(\lambda^2\right)\right] < -\sum_{\nu\neq\mu}\left|\frac{\sum_i \Gamma_{\mu i}\mathrm{B}_{i\nu}}{\Delta_\mu}\right| \ \forall \mu. \tag{2.86}$$

Looking at Eq.(2.80), we see that this is equivalent to:

$$\mathrm{Re}\,(S_{\mu\mu}) + \sum_{\nu\neq\mu} |S_{\mu\nu}| < 0 \ \forall \mu. \tag{2.87}$$

Using Lemma 1, we know that all the eigenvalues of $S(\lambda)$ will have a negative real part. As explained before that means that if $\mathrm{Re}\,(\lambda_1) > 0$ in Eq.(2.79) (unstable regime), then $\mathrm{Re}\,(\lambda_1) < 0$, which leads to a contradiction. This then implies that the equilibrium is dynamically stable. □

Weak LRI regime

A weaker version of Theorem 1 can also be stated.

Theorem 2. *If a model is not marginally stable at equilibrium and it verifies:*

$$(\Gamma B)_{\mu\mu} < -\sum_{\nu\neq\mu} \left|(\Gamma B)_{\mu\nu}\right| \ \forall \mu, \tag{2.88}$$

then its real eigenvalues are negative.

Proof. We assume

$$(\Gamma \mathrm{B})_{\mu\mu} < -\sum_{\nu\neq\mu} \left|(\Gamma \mathrm{B})_{\mu\nu}\right| \ \forall \mu. \tag{2.89}$$

[17] This works because $\Delta_\mu > 0$

Let $\lambda \in \mathbb{R}$, then the following will also trivially hold:

$$(\Gamma \mathrm{B})_{\mu\mu} - \lambda^2 < -\sum_{\nu \neq \mu} \left| (\Gamma \mathrm{B})_{\mu\nu} \right| \ \forall \mu. \tag{2.90}$$

Dividing Eq.(2.90) by[18] Δ_μ, we get:

$$\frac{1}{\Delta_\mu} \left[\left(\sum_i \Gamma_{\mu i} \mathrm{B}_{i\mu} \right) - \lambda^2 \right] < -\sum_{\nu \neq \mu} \left| \frac{\sum_i \Gamma_{\mu i} \mathrm{B}_{i\nu}}{\Delta_\mu} \right| \ \forall \mu. \tag{2.91}$$

Looking at Eq.(2.80), we see that this is equivalent to:

$$S_{\mu\mu} + \sum_{\nu \neq \mu} |S_{\mu\nu}| < 0 \ \forall \mu. \tag{2.92}$$

Using Lemma 1, we know that all the real eigenvalues of $S(\lambda)$ will have a negative real part. We can conclude with the statement of the theorem. □

Feasibility of the low intra resources interaction regime

So we found that if a system has parameters that respect Eq.(**??**) then it is dynamically stable. A naturally arising question is then to ask in what measure this is compatible with the feasability equations Eqs.(2.20), (**??**) and (**??**).
Finding an approximation of the resource interaction matrix $(\Gamma \mathrm{B})_{\mu\nu}$ using the metaparameters allows to find a necessary condition on the metaparameters. Indeed, using the metaparameters approximations Eq.(1.6), we get:

$$(\Gamma \mathrm{B})_{\mu\nu} \approx \sigma_0 \gamma_0 S_0 \left(\alpha_0 \sum_i A_{\mu i} G_{i\nu} - \gamma_0 R_0 \sum_i G_{i\mu} G_{i\nu} \right) \equiv \sigma_0 \gamma_0 S_0 \left(\alpha_0 O_{\mu\nu} - \gamma_0 R_0 C_{\mu\nu} \right) \tag{2.93}$$

where we defined the syntrophy overlap matrix $O_{\mu\nu}$ and the consumption overlap matrix $C_{\mu\nu}$ as:

$$O_{\mu\nu} \equiv (AG)_{\mu\nu} \text{ and } C_{\mu\nu} \equiv \left(G^T G \right)_{\mu\nu}. \tag{2.94}$$

The fight syntrophy vs. consumption between these two binary matrices essentially builds the dynamics of our model and an intuition about their meaning can be very helpful.

The syntrophy overlap matrix $O_{\mu\nu}$ is defined as:

$$O_{\mu\nu} \equiv \sum_k A_{\mu k} G_{k\nu}. \tag{2.95}$$

Although A and G are binary, O does not have to and usually won't be. A given consumer k contributes to $O_{\mu\nu}$ if and only if both $A_{\mu k}$ and $G_{k\nu}$ are non zero, that is if consumer k releases resource μ *and* consumes resource ν. Hence $O_{\mu\nu}$ essentially tells how many species effectively link resource μ to resource ν through the indirect interaction of the species consumption.

[18] This works because $\Delta_\mu > 0$

Similarly, the consumption overlap matrix is defined as:

$$C_{\mu\nu} = \sum_k G_{k\mu} G_{k\nu}. \tag{2.96}$$

Like O, C usually will not be binary. The intuition behind $C_{\mu\nu}$ is straight forward: it counts how many species eat both resource μ and ν. Note that $C_{\mu\nu} = C_{\nu\mu}$ (interesting: hard part comes from the antisymmetric part of S).

We then find a lowerbound for the RHS of Eq.(**??**):

$$-\sum_{\nu\neq\mu} |\Gamma\mathrm{B}|_{\mu\nu} \geqslant -\sum_{\nu\neq\mu} \max_{\nu\neq\mu} |\Gamma\mathrm{B}|_{\mu\nu} \geqslant -\deg(\mu, O-C) \max_{\nu\neq\mu} |\Gamma\mathrm{B}|_{\mu\nu}\,. \tag{2.97}$$

Combining this with the approximation of ΓB above we get an approximative LRI regime condition on the metaparameters:

$$\boxed{\alpha_0 O_{\mu\mu} - \gamma_0 R_0 C_{\mu\mu} \lessapprox -\deg(\mu, O-C) \max_{\nu\neq\mu} |\alpha_0 O_{\mu\nu} - \gamma_0 R_0 C_{\mu\nu}| - \frac{R_C^2}{\sigma_0 \gamma_0 S_0}\ \forall\mu.} \tag{2.98}$$

Since R_C essentially scales with the largest degree of G (see Eq.2.77) we only expect systems with a low connectance food consumption adjacency matrix to be able to achieve an LRI state.

This allows to give a necessary condition on the magnitude of α_0. Indeed, since the RHS of the previous equation is negative, we need:

$$\alpha_0 O_{\mu\mu} - \gamma_0 R_0 C_{\mu\mu} < 0\ \forall\mu \implies \alpha_0 \max_\mu \left\{\frac{O_{\mu\mu}}{C_{\mu\mu}}\right\} < \gamma_0 R_0. \tag{2.99}$$

It is clear that systems with the maximal ratio of $O_{\mu\mu}$ and $C_{\mu\mu}$ is small will be more easily into a LRI regime. The most favoured systems will be those where $S_{\mu\mu} = 0\ \forall\mu$, *i.e.* systems where no species consumes what it itself produces. In that way we may say that coprophagy tends to destabilize microbial communities. Combining Eq.(**??**) and (2.99) gives us a necessary condition on α_0 for feasible systems (need more details?):

$$\boxed{\alpha_0 \left[\max_\mu \{O_{\mu\mu}\} - \min_\mu \left\{\frac{k_\mu^\alpha}{k_\mu^\gamma}\right\}\right] \leqslant \frac{l_0}{\min_\mu (k_\mu^\gamma)\, S_0}}. \tag{2.100}$$

Although that equation gives us a necessary condition, it is not sufficient. Eq.(2.98), on the other hand, is and provides an intuitive way of finding a syntrophy adjacency matrix $A_{\mu i}$ that would put a system with a given consumption adjacency $G_{\mu i}$ in an LRI regime. Section **??** explains in details how this can be achieved numerically.

Monte Carlo algorithm for the optimal syntrophy matrix

We want to find a general algorithm which, for a given food consumption adjacency matrix G gives back an optimal syntrophy adjacency matrix A. Strategically, we would like an A such that Eq.(2.98) is as close to being satisfied as possible. If it were satisfied, it would put the system in an LRI regime, which we have proven is dynamically stable.

One way of trying to satisfy Eq.(2.98) is to increase the magnitude of its LHS and minimize the magnitude of the RHS. The LHS is minimized if $(AG)_{\mu\mu}$ is set to its lowest possible value for every μ, that is zero. On the other hand, the RHS is minimized if $\alpha_0(AG)_{\mu\nu} \approx \gamma_0 R_0 (G^T G)_{\mu\nu}\ \forall \nu \neq \mu$.

Intuitively, we then search for systems where AG is zero on the diagonal, *i.e.* where no coprophagy is observed, and $AG \approx \frac{\gamma_0 R_0}{\alpha_0} G^T G$ outside the diagonal. It can be formalized by writing a proper Metropolis-Hastings Markov Chain Monte Carlo (MCMC) method. We designed the following algorithmic procedure to build a syntrophy adjacency matrix A:

1. Create a random A. Its connectance is chosen as the one of the consumption matrix G.

2. Do the following for a given number of steps:

   - Choose a random row or, every other iteration, a column.
   - In that row/column, try to swap a zero and a one while preserving the "releasers": if a species releases some resource, it has to keep releasing something (the resource can change though). The "releasees" are preserved as well: if a resource is being released by some species, it has to keep being released (but it does not have to be by the same species). **why do we impose those conditions?**
   - The swap is accepted, *i.e.* A is modified, if the energy difference ΔE is negative or if a random number drawn uniformly between zero and one is smaller than $e^{-\Delta E/T}$ where T is the current temperature. More on ΔE and T below.

3. Return A.

A couple comments on this algorithm can be made:

- The algorithm preserves the connectance of A but not its nestedness. The question of what value to choose is open, but we choose $\kappa(A) = \kappa(G)$ as a first approach, *i.e.* syntrophy and consumption networks have the same connectance.

- The temperature T changes dynamically during the simulation. It is obtained in a way close to the spirit of simulated annealing techniques [32]: the temperature T is multiplied by a factor $\lambda = 0.99$ at a fixed frequency (for instance every 1000 steps). We add the requirement that if new moves are rejected during too many consecutive steps, we multiply the temperature by $1/\lambda$.

- The energy difference ΔE between the new proposed A' and the old A is computed by assigning an energy E to both A' and A and subtracting them:

$$\Delta E \equiv E(A', G) - E(A, G). \tag{2.101}$$

  The choice of the energy function E is crucial. In essence, this MCMC algorithm will find the specific A which minimizes $E(A)$. Since we want to work with systems in the LRI regime, we use the simplest and most natural function that is compatible with

the intuitively expected characteristics of A explained above (*i.e.* AG is zero on the diagonal and equal to $\frac{\gamma_0 R_0}{\alpha_0} G^T G$ outside of it):

$$E(A,G) \equiv \sum_{\mu} \left(|\alpha_0 (AG)_{\mu\mu}| + \sum_{\nu \neq \mu} \left| (\alpha_0 AG - \gamma_0 R_0 G^T G)_{\mu\nu} \right| \right). \tag{2.102}$$

The energy function and hence the optimal syntrophy adjacency matrix A depend on the ratio $\frac{\alpha_0}{\gamma_0 R_0}$. This prompts then the question of which α_0 can be deemed sensible. As a first step, we will take the value of Eq.(**??**): $\alpha_0 = \min(1-\sigma_0, \sigma_0)\gamma_0 R_0 N_R$. This means that the outcome of the algorithm is an optimized A **for the largest feasible syntrophy**. Since the expression we have for the largest feasible syntrophy is independent of the G matrix, this choice of α_0 provides us a sensible way of comparing different consumption networks.

2.2.5 Flux analysis - a way to get a sense of scales

A natural scale free order parameter that at first sight controls the behaviour of the system is the ratio of the syntrophy and consumption fluxes.

The rate of consumption (or *consumption flux*) of species i is given by $\sum_\nu \sigma_{i\nu}\gamma_{i\nu} R_\nu S_i$. Hence the total consumption flux C_{tot} is given by:

$$C_{\text{tot}} = \sum_{i,\nu} \sigma_{i\nu}\gamma_{i\nu} R_\nu S_i. \tag{2.103}$$

We can similarly define the total syntrophy flux of the system S_{tot}:

$$S_{\text{tot}} = \sum_{i,\nu} \alpha_{\nu i} S_i. \tag{2.104}$$

A natural order parameter O is then

$$O \equiv \frac{S_{\text{tot}}}{C_{\text{tot}}} = \frac{\sum_{i,\nu} \alpha_{\nu i} S_i}{\sum_{i,\nu} \sigma_{i\nu}\gamma_{i\nu} R_\nu S_i} \approx \frac{N_S N_R \alpha_0 S_0}{\sigma_0 R_0 S_0 N_S N_R \gamma_0} = \frac{\alpha_0}{\sigma_0 R_0 \gamma_0}. \tag{2.105}$$

2.2.6 Application: fully connected consumption and syntrophy networks

As an application, we consider the very special case of the fully connected consumption and syntrophy networks. In this "mean-field" theory, every consumer consumes from and releases each resource, *i.e.*

$$G_{i\mu} = A_{\mu i} = 1 \ \forall \mu = 1, \ldots, N_R \ \ \forall i = 1, \ldots, N_S. \tag{2.106}$$

Our goal is to find the spectrum of systems with such matrices as their consumption and syntrophy networks. We follow the path explained by Barbier and Arnoldi in [26] and perform a *standard deviation expansion*. They showed that the variance of the interaction matrix drives the strength of the interactions in their model [26], so it is worth a try to follow their path.

Standard deviation expansion The idea behind the standard deviation expansion (abbreviated SDE) is the following. Let $q_{i\mu}$ be an arbitrary matrix of size $N_S \times N_R$. Then the nice trick done in [26] is to write the elements of the q matrix in terms of new variables $\tilde{q}_{i\mu}$:

$$q_{i\mu} = \langle q \rangle + \sigma_q \tilde{q}_{i\mu}. \tag{2.107}$$

In that expression, $\langle q \rangle$ is the average of q, element-wise

$$\langle q \rangle \equiv \frac{\sum_{\mu,i} q_{i\mu}}{N_S N_R}, \tag{2.108}$$

and σ_q is the standard deviation of q, again element-wise:

$$\sigma_q \equiv \sqrt{\langle q^2 \rangle - \langle q \rangle^2} \text{ with } \langle q^2 \rangle \equiv \frac{\sum_{\mu,i} q_{i\mu}^2}{N_S N_R}. \tag{2.109}$$

The main advantage of this procedure is that we get a clear idea about the scales involved. A matrix element $q_{i\mu}$ is roughly the mean $\langle q \rangle$ plus a deviation σ_q multiplied by a factor of magnitude sim 1. Indeed the $\tilde{q}_{i\mu}$ are not large since they follow the two equalities [26]:

$$\langle \tilde{q} \rangle = 0 \text{ and } \langle \tilde{q}^2 \rangle = 1. \tag{2.110}$$

We apply this framework to our problem by noticing that if the $q_{i\mu}$ are all random samples coming from the same distribution law $\mathfrak{Q}$, we can write the following approximation in the case $N_R, N_S \gg 1$:

$$\langle q \rangle \approx \langle \mathfrak{Q} \rangle \equiv q_0. \tag{2.111}$$

We can then rewrite the free parameters of our model[19]:

$$\begin{cases} l_\nu \approx l_0 + \sigma_l \tilde{l}_\nu & \text{(2.112a)} \\ R^*_\nu \approx R_0 + \sigma_R \tilde{r}_\nu & \text{(2.112b)} \\ S^*_i \approx S_0 + \sigma_S \tilde{s}_i & \text{(2.112c)} \\ \gamma_{i\nu} \approx \gamma_0 + \sigma_\gamma \tilde{g}_{i\nu} & \text{(2.112d)} \\ \alpha_{\nu i} \approx \alpha_0 + \sigma_\alpha \tilde{\alpha}_{\nu i} & \text{(2.112e)} \\ \sigma_{i\nu} \approx \sigma_0 + \sigma_\sigma \tilde{\sigma}_{i\nu} & \text{(2.112f)} \end{cases}$$

The general idea is then to assume that the standard deviations are small and proceed to a first order Taylor expansion.

Rewriting the jacobian at equilibrium The different blocks of the jacobian at equilibrium (**??**) can be written with the new variables :

$$\begin{cases} \dfrac{l_\mu + \sum_j \alpha_{\mu j} S^*_j}{R^*_\mu} = \dfrac{l_0 + N_S \alpha_0 S_0 + \sigma_l \tilde{l}_\mu + \sigma_\alpha S_0 \sum_j \tilde{\alpha}_{\mu j} + \sigma_\alpha \sigma_S \sum_j \tilde{\alpha}_{\mu j} \tilde{s}_j}{R_0 + \sigma_R \tilde{r}_\mu} & \text{(2.113)} \\ -\gamma_{j\mu} R^*_\mu + \alpha_{\mu j} = -\gamma_0 R_0 + \alpha_0 + \sigma_\gamma R_0 \tilde{\gamma}_{j\mu} + \sigma_R \gamma_0 \tilde{r}_\mu + \sigma_\alpha \tilde{\alpha}_{\mu j} + \sigma_\gamma \sigma_R \tilde{\gamma}_{j\mu} \tilde{r}_\mu & \text{(2.114)} \\ \sigma_{i\nu} \gamma_{i\nu} S^*_i = \sigma_0 \gamma_0 S_0 + \sigma_\sigma \gamma_0 S_0 \tilde{\sigma}_{i\nu} + \sigma_\gamma \sigma_0 S_0 \tilde{\gamma}_{i\nu} + \sigma_S \sigma_0 \gamma_0 \tilde{s}_i + \sigma_\sigma \sigma_\gamma S_0 \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \\ \qquad + \sigma_\sigma \sigma_S \gamma_0 \tilde{\sigma}_{i\nu} \tilde{s}_i + \sigma_\gamma \sigma_S \sigma_0 \tilde{\gamma}_{i\nu} \tilde{s}_i + \sigma_\sigma \sigma_\gamma \sigma_S \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \tilde{s}_i & \text{(2.115)} \end{cases}$$

[19]This works with γ and α because G and A have a trivial topology. Otherwise we would have to take it into account and the computations would not be as easy.

It's easier to work with relative standard deviations, *i.e.* we rewrite for all parameters :

$$\sigma_P \equiv \epsilon_P \langle P \rangle, \ \forall P \in \{l_\nu, R^*_\nu, S^*_i, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \tag{2.116}$$

The previous relations then become :

$$\begin{cases} \dfrac{l_\mu + \sum_j \alpha_{\mu j} S^*_j}{R^*_\mu} = \dfrac{l_0 \left(1 + \epsilon_l \tilde{l}_\mu\right) + \alpha_0 S_0 \left(N_S + \epsilon_\alpha \sum_j \tilde{\alpha}_{\mu j} + \epsilon_\alpha \epsilon_S \sum_j \tilde{\alpha}_{\mu j} \tilde{s}_j\right)}{R_0 \left(1 + \epsilon_R \tilde{r}_\mu\right)} & (2.117) \\ -\gamma_{j\mu} R^*_\mu + \alpha_{\mu j} = -\gamma_0 R_0 \left(1 + \epsilon_\gamma \tilde{\gamma}_{j\mu} + \epsilon_R \tilde{r}_\mu + \epsilon_\gamma \epsilon_R \tilde{\gamma}_{j\mu} \tilde{r}_\mu\right) + \alpha_0 \left(1 + \epsilon_\alpha \tilde{\alpha}_{\mu j}\right) & (2.118) \\ \sigma_{i\nu} \gamma_{i\nu} S^*_i = \sigma_0 \gamma_0 S_0 \left(1 + \epsilon_\sigma \tilde{\sigma}_{i\nu} + \epsilon_\gamma \tilde{\gamma}_{i\nu} + \epsilon_S \tilde{s}_i + \epsilon_\sigma \epsilon_\gamma \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} + \epsilon_\sigma \epsilon_S \tilde{\sigma}_{i\nu} \tilde{s}_i \right. \\ \qquad \left. + \epsilon_\gamma \epsilon_S \tilde{\gamma}_{i\nu} \tilde{s}_i + \epsilon_\sigma \epsilon_\gamma \epsilon_S \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \tilde{s}_i\right) & (2.119) \end{cases}$$

Standard deviation expansion at first order

We can then assume the relative standard deviations are small :

$$\epsilon_P \ll 1, \ \forall P \in \{l_\nu, R^*_\nu, S^*_i, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \tag{2.120}$$

The previous equations can be rewritten as :

$$\begin{cases} \dfrac{l_\mu + \sum_j \alpha_{\mu j} S^*_j}{R^*_\mu} = \dfrac{l_0 + N_S \alpha_0 S_0}{R_0} - \epsilon_R \dfrac{l_0 + N_S \alpha_0 S_0}{R_0} \tilde{r}_\mu + \epsilon_l \dfrac{l_0}{R_0} \tilde{l}_\mu + \epsilon_\alpha \dfrac{\alpha_0 S_0}{R_0} \sum_j \tilde{\alpha}_{\mu j} & (2.121) \\ -\gamma_{j\mu} R^*_\mu + \alpha_{\mu j} = -\gamma_0 R_0 + \alpha_0 - \epsilon_\gamma \gamma_0 R_0 \tilde{\gamma}_{j\mu} - \epsilon_R \gamma_0 R_0 \tilde{r}_\mu + \epsilon_\alpha \alpha_0 \tilde{\alpha}_{\mu j} & (2.122) \\ \sigma_{i\nu} \gamma_{i\nu} S^*_i = \sigma_0 \gamma_0 S_0 \left(1 + \epsilon_\sigma \tilde{\sigma}_{i\nu} + \epsilon_\gamma \tilde{\gamma}_{i\nu} + \epsilon_S \tilde{s}_i\right) & (2.123) \end{cases}$$

where we neglect all terms of order $\mathcal{O}(\epsilon^2)$. We now assume furthermore that the relative standard deviations of every parameter in the model more or less have the same value which again is assumed small, *i.e.* we set :

$$\epsilon_P \approx \epsilon \ll 1, \ \forall P \in \{l_\nu, R^*_\nu, S^*_i, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \tag{2.124}$$

This allows us to rewrite the jacobian at equilibrium as :

$$J^* = J_0 + \epsilon \tilde{J}, \tag{2.125}$$

with

$$J_0 = \begin{pmatrix} -\Delta_0 & \Gamma_0 \\ B_0 & 0 \end{pmatrix} \tag{2.126}$$

where

$$(\Delta_0)_{\mu\nu} \equiv \Delta_0 \delta_{\mu\nu} = \frac{l_0 + N_S \alpha_0 S_0}{R_0} \delta_{\mu\nu} \tag{2.127}$$

$$(\Gamma_0)_{\mu i} \equiv \Gamma_0 = -\gamma_0 R_0 + \alpha_0 \tag{2.128}$$

$$(\mathsf{B}_0)_{i\nu} \equiv \mathsf{B}_0 = \sigma_0 \gamma_0 S_0 \tag{2.129}$$

and

$$\tilde{J} = \begin{pmatrix} \tilde{\Delta} & \tilde{\Gamma} \\ \tilde{\mathsf{B}} & 0 \end{pmatrix} \tag{2.130}$$

with

$$\tilde{\Delta}_{\mu\nu} \equiv \left(-\Delta_0 \tilde{r}_\mu + \frac{l_0}{R_0} \tilde{l}_\mu + \frac{\alpha_0 S_0}{R_0} \sum_j \tilde{\alpha}_{\mu j} \right) \delta_{\mu\nu} \tag{2.131}$$

$$\tilde{\Gamma}_{\mu i} \equiv \alpha_0 \tilde{\alpha}_{\mu i} - \gamma_0 R_0 \left(\tilde{\gamma}_{i\mu} + \tilde{r}_\mu \right) \tag{2.132}$$

$$\tilde{B}_{i\mu} \equiv B_0 \left(\tilde{\sigma}_{i\mu} + \tilde{\gamma}_{i\mu} + \tilde{s}_i \right) \tag{2.133}$$

Using Jacobi's formula [33], the equation $\det\left(J^* - \lambda\right) = 0$ can be rewritten as :

$$\boxed{\det(J_0 - \lambda + \epsilon J^*) = \det(J_0 - \lambda) + \epsilon \, \text{Tr}\left(\text{adj}(J_0 - \lambda)\tilde{J}\right) = 0.} \tag{2.134}$$

where adj $(\dots)$ is the adjugate operator (*i.e.* which yields the transpose of the cofactor matrix). This equation is a complicated polynomial of degree $N_R + N_S$. As of now it does not seem to have an easily computable solution for $\epsilon > 0$. An explicit solution can however be computed when $\epsilon = 0$.

Zero variance case When $\epsilon = 0$, Eq.(2.134) becomes:

$$\det(J_0 - \lambda) = \det \begin{pmatrix} -\Delta_0 - \lambda & \Gamma_0 \\ B_0 & -\lambda \end{pmatrix} = 0 \tag{2.135}$$

If we assume that $\lambda \neq 0$, using a reasoning similar to Section **??** we can write the previous equation as:

$$\det\left(\lambda^2 + \Delta_0 \lambda - \Gamma_0 \mathrm{B}_0\right) = 0 \tag{2.136}$$

Component-wise, we have :

$$\left(\lambda^2 + \Delta_0 \lambda - \Gamma_0 \mathrm{B}_0\right)_{\mu\nu} = \left(\lambda^2 + \Delta_0 \lambda\right) \delta_{\mu\nu} - \Gamma_0 \mathrm{B}_0. \tag{2.137}$$

Using Eq.(**??**), the non-zero solutions of Eq.(2.136) are immediately found:

$$\left(\lambda + \Delta_0\right)^{N_R - 1} \left(\lambda^2 + \Delta_0 \lambda - N_R \Gamma_0 \mathrm{B}_0\right) = 0. \tag{2.138}$$

That equation gives us $N_R - 1 + 2 = N_R + 1$ non-zero eigenvalues, which means that there are $N_S - 1$ zero eigenvalues. The two eigenvalues different from $-\Delta_0$ or 0 are the roots of the second degree polynomial:

$$\lambda^2 + \Delta_0 \lambda - N_R \Gamma_0 \mathrm{B}_0 = 0. \tag{2.139}$$

In the end, the spectrum is given by:

- if $\Gamma_0 < -\frac{\Delta_0^2}{4 N_R \mathrm{B}_0}$:

$$\sigma(J_0) = \left\{ 0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2} \left(1 \pm i \sqrt{-\left(1 + \frac{4 N_R \Gamma_0 \mathrm{B}_0}{\Delta_0^2} \right)} \right) \right\} \tag{2.140}$$

- if $\Gamma_0 = -\frac{\Delta_0^2}{4N_R \mathrm{B}_0}$:

$$\sigma(J_0) = \left\{0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2}, -\frac{\Delta_0}{2}\right\} \tag{2.141}$$

- if $\Gamma_0 > -\frac{\Delta_0^2}{4N_R \mathrm{B}_0}$:

$$\sigma(J_0) = \left\{0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2}\left(1 \pm \sqrt{1 + \frac{4N_R \Gamma_0 \mathrm{B}_0}{\Delta_0^2}}\right)\right\} \tag{2.142}$$

It then becomes clear that the system is dynamically unstable if and only if $\frac{4N_R\Gamma_0\mathrm{B}_0}{\Delta_0^2} > 0$. Because $N_R, \mathrm{B}_0 > 0$, we get the condition:

$$\boxed{\text{The non-variance system is dynamically unstable } \iff \Gamma_0 > 0.} \tag{2.143}$$

If $\Gamma_0 \leqslant 0$, the fully connected system will be marginally stable. Note that the RHS of the feasibility condition Eq.(2.29) is equivalent in the fully connected case to :

$$\alpha_0 \lessapprox \min(1-\sigma_0, \sigma_0)\gamma_0 R_0 \iff \Gamma_0 \lessapprox \left[\min(1-\sigma_0, \sigma_0) - 1\right]\gamma_0 R_0 < 0, \tag{2.144}$$

so we expect the fully connected case to be feasible and the non-zero eigenvalues to be negative. We see that the non-zero part of the spectrum is essentially given by Δ_0. However the sign of the parameter $4N_R\Gamma_0\mathrm{B}_0 + \Delta_0^2$ plays an essential role in the stability. Looking at its sign allows us to find possibly more locally dynamically stable zones of the metaparameters space $\mathcal{M}$. Indeed we expect to find more stable systems if

$$4N_R\Gamma_0\mathrm{B}_0 + \Delta_0^2 \ll 0 \iff 4N_R\sigma_0\gamma_0(\alpha_0 - \gamma_0 R_0) + \frac{l_0^2}{S_0} + 2N_S\alpha_0 l_0 + N_S^2\alpha_0^2 S_0 \ll 0. \tag{2.145}$$

2.3 Structural stability

When studying dynamical stability, we investigate what happens when the equilibria abundances $\{R_\mu^*, S_i^*\}$ of a given equilibrium point are perturbed. The question of *structural stability* looks also at the behaviour of a given system when perturbed away from equilibrium. However, structural stability focuses on the perturbations of the parameters of the model *i.e.* $\{l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i\}$. Namely we will try to answer the following question :

Given an equilibrium point, does the system go back to a positive-valued equilibrium when some of the model parameters are changed? If yes, how much can they be changed before the system evolves in such a way that it does not reach a positive-valued equilibrium?

2.3.1 Definitions

Studying how a system responds to the perturbation of all parameters $\{l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i\}$ is a quite difficult problem. So we will try to simplify it by perturbing *only one* parameter. We make the somewhat arbitrary choice of perturbing the external feeding rate l_μ, since it is essentially the only parameter one can control experimentally [**is this true?**]. More precisely, consider $\Delta_S \in [0,1]$, we say that a given system $p \in \mathcal{P}$ is *structurally stable* under the perturbation Δ_S, if under the transformation

$$l_\mu \to \hat{l}_\mu \equiv l_\mu \left(1 + \Delta_S \nu_\mu\right) \tag{2.146}$$

the transformed set of parameters $\left\{\hat{l}_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i\right\}$ gives rise under time evolution to a positive valued-equilibrium $\left\{\hat{R}_\mu^*, \hat{S}_i^*\right\}$. In the equation above, ν_μ is a random variable drawn from a uniform distribution of support $[-1,1]$.

Similarly to what was done for feasibility and dynamical stability, we can define the *parameters set structural stability function* $\mathfrak{S} : [0,1] \times \mathcal{P} \to \{0,1\}$ in the following way $\forall \Delta_S \in [0,1], p \in \mathcal{P}$:

$$\mathfrak{S}(\Delta_S, p) = \begin{cases} 1 \text{ if } p \text{ is structurally stable under the perturbation } \Delta_S, \\ 0 \text{ otherwise.} \end{cases} \tag{2.147}$$

For a fixed p, we expect $\mathfrak{S}(\Delta_S, p)$ to behave as a step function of Δ_S : it makes sense that we may only perturb the parameters so much before they become structurally unstable.

The corresponding metaparameters set function, the *metaparameters set structural stability function* $\mathcal{S}$ can also be defined as the function which, given a set of metaparameters and a consumption-syntrophy couple of binary matrices, tells you how probable it is that you draw a system structurally stable under a perturbation Δ_S. Mathematically, $\mathcal{S} : [0,1] \times \mathcal{M} \times \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S} \to [0,1]$ is defined $\forall \Delta_S \in [0,1], m \in \mathcal{M}, B = (G, A) \in \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$:

$$\boxed{\mathcal{S}(\Delta_S, m, B) = \text{Probability}\left\{\mathfrak{S}(\Delta_S, \mathcal{A}(m, B)) = 1\right\}} \tag{2.148}$$

Because we expect a step-like drop of $\mathfrak{S}$ as Δ_S increases, we expect also a somewhat sharp drop from $\mathcal{S} \approx 1$ to $\mathcal{S} \approx 0$. To quantify this, one can define the *critical structural perturbation* $\Delta_S^*(m, G, A)$ of a consumption-syntrophy network couple implicitly as :

$$\mathcal{S}(\Delta_S^*(m, G, A), m, G, A) = 0.5 \tag{2.149}$$

Methods 2.3.2 below explains how $\Delta_S^*(m, G, A)$ can be estimated numerically.

For a given consumption-syntrophy network couple, we may finally define the (Δ_S, x)-structurally stable region $\mathcal{S}_{\Delta_S,x}^{G,A}$ of the metaparameter space $\mathcal{M}$:

$$\mathcal{S}_{\Delta_S,x}^{G,A} \equiv \{m \in \mathcal{M} : \mathcal{S}(\Delta_S, m, G, A) = x\} \tag{2.150}$$

which is the set of all points in the metaparameters space that have a probability x of being stable under a perturbation Δ_S.

2.3.2 Numerical estimate of the critical structural perturbation

As explained above, the critical structural perturbation of set of metaparameters m and a consumption-syntrophy couple of matrices is the point where we shift from structural stability to instability. In that sense, $\Delta_S^*(m, G, A)$ is a measure of how good (m, G, A) respond to structural perturbation and can be interpreted geometrically as the radius of a sphere of "tolerance" around m **Shouldn't we change all metaparameters in order to achieve this?**. It turns out that Δ_S^* can be estimated numerically quite easily.

Indeed, to decide whether a system is structurally stable or not, one can simply perturb the parameters of the system, let it time-evolve until it reaches a new equilibrium and count how many of the original consumers are still present at the new equilibrium. By repeating this procedure many times one gets a good estimate of the *probability of observing an extinction* $P_E(\Delta_S, m, G, A)$ after a structural perturbation Δ_S and it is clear that

$$\mathcal{S}(\Delta_S, m, G, A) = 1 - P_E(\Delta_S, m, G, A). \tag{2.151}$$

So $\Delta_S^*(m, G, A)$ can be found by computing $P_E(\Delta_S, m, G, A)$ over the range $[0, 1]$ and finding at which point it is equal to zero. In practice, a general solver involving methods from the C++ `GSL` library was implemented in order to get a good estimate on $\Delta_S^*(m, G, A)$ (see Results **??**).

Chapter 3

Results

3.1 Feasibility

As explained in Methods **??**, before addressing the question of the stability of a system, be it dynamical or structural, it is important to study whether that system is *feasible*. In short we must answer the question: "does it make sense to talk about this system? Does it even exist?".

We will say that it makes sense to talk about a system if it is *feasible* (see Methods **??**). In order to be feasible, a system should respect two conditions: it must conserve biomass and its parameters must have a direct biological interpretation.

3.1.1 The feasibility volume $\mathcal{V}_x^{G,A}$

Formally we can define $\forall x \in [0,1]$ the x-feasibility volume $\mathcal{V}_x^{G,A} \subset \mathcal{M}$ of the consumption network coupled with the syntrophy network $(G,A) \in \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$ (see Methods **??**). Every metaparameter set $m \in \mathcal{M}$ contained in the x-feasibility volume $\mathcal{V}_x^{G,A}$ will give rise to a percentage x of feasible systems. A first order approximation of the fully feasible volume $\mathcal{V}_1^{G,A}$ is given by Eq.(**??**). In the absence of syntrophy $\alpha_0 = 0$, it becomes:

$$\gamma_0 R_0 \lessapprox \frac{l_0}{\max_\nu\{\deg(G,\nu)\} S_0}. \tag{3.1}$$

This relation is interesting in many ways. First of all it tells us that at fixed consumption rate γ_0 and resource equilibrium abundance R_0, feasibility increases when:

- the external resource input rate increases. This result was somewhat expected: if you give more food to a goldfish you expect it to thrive more.

- the consumer equilibrium abundance S_0 increases. What this means is that if you want to maintain the same consumption interaction but get a higher abundance of resources at equilibrium, you must at the same time decrease the consumers equilibrium abundance.

- the largest column-degree of the consumption matrix decreases. The degree of a given column ν of the consumption matrix tells you how many species eat from resource ν. This encourages communities of specialists, where each consumer eats from its own and no other resource.

Overall we see that feasibility increases when the *consumption flow* $\simeq \gamma_0 S_0 \deg(G,\nu)$ is low $\forall\nu$. Eq.(3.1) can be confronted to simulations. First of all note that for all the matrices in the set we chose, there existed a fully feasible zone. Fortunately, there was an overlap between all of these, such that the critical feasibility $f^*(S) = 1$ and the critical volume is the fully feasible volume which we denote $\mathcal{V}^1$.

Figure 3.1 shows the proportion of feasible systems without syntrophy $\mathcal{F}(\gamma_0, S_0, \alpha_0 = 0, G)$ and $R_0 = l_0 = 1$, $\sigma_0 = 0.25$ (see Methods **??**), for two matrices G_1 and G_2 of our set. G_1 has connectance $\kappa_1 = 0.17$ and nestedness $\eta_1 = 0.2$, G_2 is more connected and more nested: $\kappa_2 = 0.37$ and $\eta_2 = 0.4$.

We observe a very sharp transition from a fully feasible to a fully unfeasible regime. Theoretically, this sharp transition happens when both sides of the inequality (3.1) are equal, *i.e.* at $\gamma_0 R_0 = l_0/\max_\nu\{\deg(G,\nu)\}S_0$. Numerically we fit the points which are at "the boundary" of the common feasible volume, *i.e.* points where $0.4 \leqslant \mathcal{F}(\gamma_0, S_0, G) \leqslant 0.6$.

For G_1, the theoretical expectation is $S_0 = 0.125/\gamma_0$ and a fit on the numerical results gives $S_0 = (0.124 \pm 7\times10^{-8})/\gamma_0 - (6.8\times10^{-4} \pm 3\times10^{-7})$ so the theoretical relation is already very good. For G_2, we expect $S_0 = 0.077/\gamma_0$. A fit gives $S_0 = (0.075 \pm 2\times10^{-8})/\gamma_0 + (3.6\times10^{-4} \pm 1\times10^{-7})$. Again, the theoretical value is very close to the measured value.

The numerical estimate does not always match that well the theoretical value. Fig.3.2 shows the relative error $\Delta_G = 1-$ (theoretical value)/(numerical estimate). We see that in general the theoretical expectation tends to overestimate the fully feasible region. This is probably due to the noise (*i.e.* the deviations away from the metaparameters) in the actual systems and the structure of the G matrix. Indeed Fig.3.2 shows that the lower the nestedness and connectance of G, the worse the theoretical estimate. In the future a better approximation can surely be found taking into account the variance of the metaparameters and the nestedness of G.

We can similarly measure the common fully feasible volume $\mathcal{V}^*$, which according to Eq.(3.1) is inversely proportional to the largest maximal row degree of the matrix set. For the set we considered, this yields in theory: $S_0 = 0.053\gamma_0$. A fit on the points at the edge yields the critical boundary $S_0 = (0.043 \pm 10^{-8})/\gamma_0 - (4.6\times10^{-3} \pm 3\times10^{-8})$. The theoretical prediction is not as good as before with an error of $\sim 20\ \%$. The discrepancy is probably due to the fact that numerically we determine the common feasibility volume by counting the points for which $\mathcal{F}(\gamma_0, S_0, G) = 1\ \forall G \in S_G$. while the theoretical value matches better a fit of the points $0.4 \leqslant \mathcal{F}\ \leqslant 0.6$ **make this more understandable**.

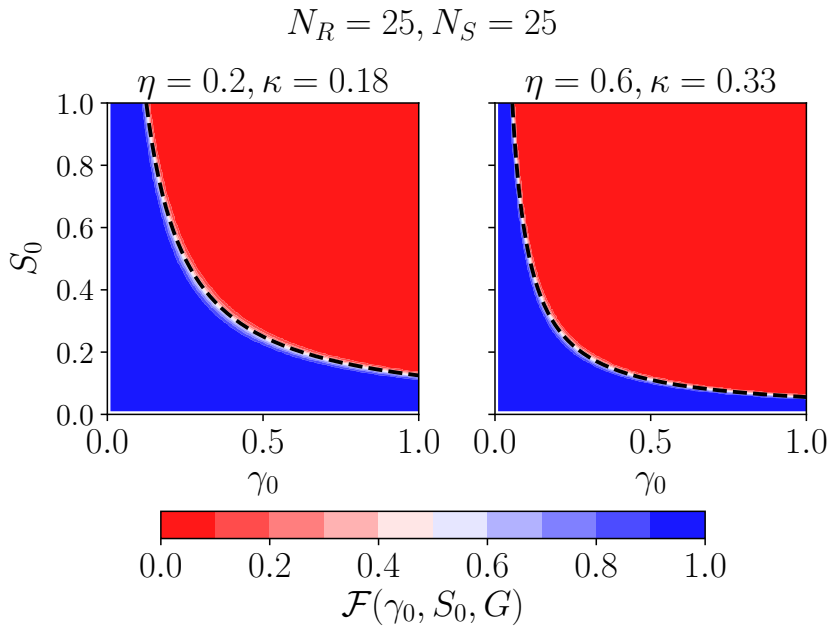

Figure 3.1: Plot of the feasability region. The color curve indicicates the feasibility function $\mathcal{F}(\gamma_0, S_0, \alpha_0 = 0, G)$ for G_1 (left) and G_2 (right). We observe a steep descent which marks a very clear transition from a totally feasible regime to a totally unfeasible regime, which allows us to precisely get the boundary of $\mathcal{V}_1^G$. The dashed lines indicate the theoretical predictions, which for both G_1 and G_2 are accurate to the order of 0.1%.

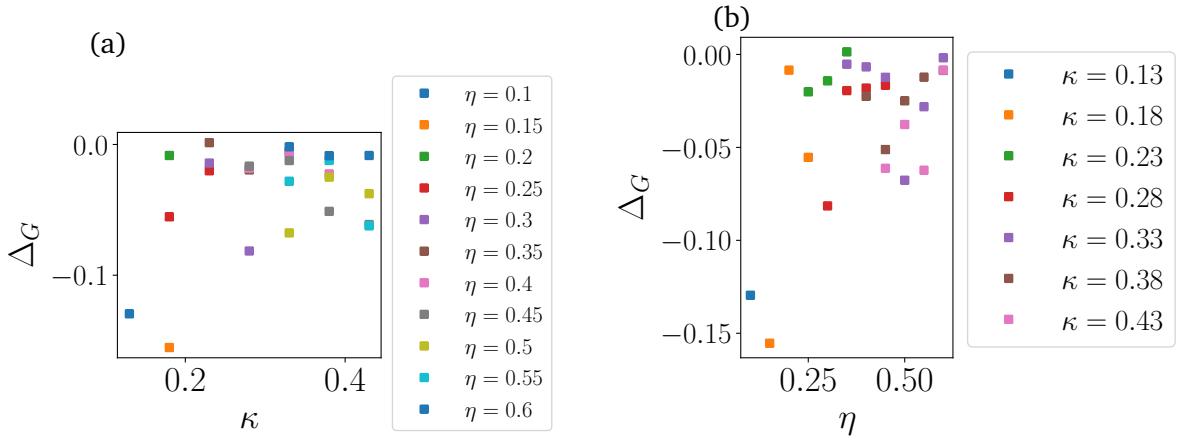

Figure 3.2: Relative error in the determination of the boundary of $\mathcal{V}_1^{G,0}$ (a) varying connectance at fixed nestedness and (b) varying nestedness at fixed connectance. The theoretical prediction tends to overestimate the measured value. The larger the nestedness and connectance, the better the estimate.

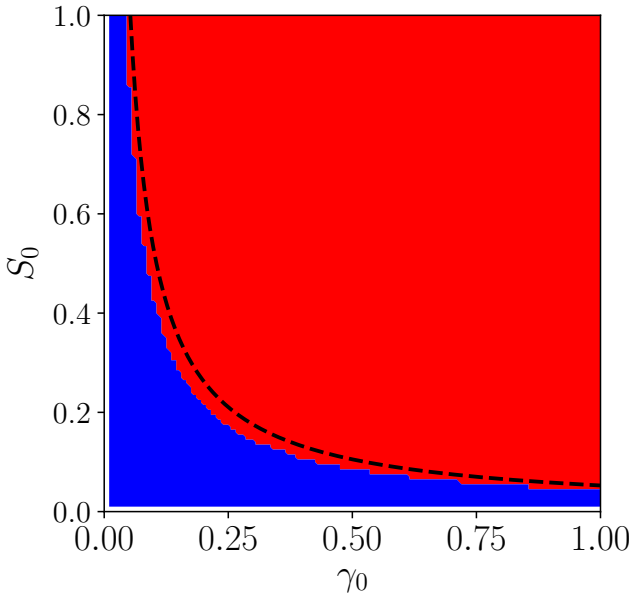

Figure 3.3: Plot of the common feasability region. The blue area indicates the common feasibility volume, computed numerically, while the dashed line shows the analytical prediction. Although the match is not as good as before, the relative error is only of the order of 20%. The red part is the area where not all matrices are fully feasible. From now on, it will not be considered anymore.

3.1.2 Evolution of the feasible volume with syntrophy

Above we computed feasibles volumes when there is no syntrophy *i.e.* $\mathcal{V}_1^G(\alpha_0 = 0)$. Since $\alpha_0 = 0$, we did not need to specify what the structure of A was. The next naturally arising question is then: what happens to the feasible volume of a given matrix G when we add a syntrophic interaction? More precisely, how does $\mathcal{V}_1^G$'s shape change?

The problem gets a bit more complex here. When we computed $\mathcal{V}_1^G(\alpha_0 = 0)$, we only had to take into account the structure of G, since there was no syntrophy. Now, we have an additional complexity because we have to think about the structure of the syntrophy network A as well. The choice of A, which is at the core of the problem, is far from trivial. We investigated three different cases:

- "fully connected": A is filled with ones only, $A_{\mu i} = 1$. This corresponds to a so-called "mean-field" approximation. Every consumer releases every resource at (up to some noise) the same intensity.
- "no intraspecific syntrophy": the structure of A is such that consumers are not allowed to release what they consume, *i.e.* there is no coprophagy.
- "optimal LRI": A is the outcome of the Monte Carlo algorithm [20] described in Methods 2.2.4, whose purpose is to make systems "more dynamically stable" and hence more feasible.

The maximal common feasible syntrophy can be estimated with the help of Eq.(**??**):

$$\alpha_0 \lessapprox \frac{\min(1-\sigma_0, \sigma_0)\gamma_0 R_0}{\max_{(G,A)\in S}\left\{\max_i\left\{\frac{\deg(A,i)}{\deg(G,i)}\right\}\right\}} \approx 0.01\gamma_0 \leqslant 0.01. \tag{3.2}$$

So we will look at ten different α_0 values from 0 to 0.015. We first consider the effect of syntrophy on each consumption network G then on the common fully feasible volume.

The influence of matrix structure

Because Eq.(**??**) depends on the structure of G and of A, we expect $\mathcal{F}_1^{G,A}$ to depend heavily on the topology of both consumption and syntrophy matrices. Figure 3.4 shows that indeed $\mathcal{F}_1^G$ not only changes with syntrophy but also with the network structure of the consumption matrix. First of all, we observe a general trend among all matrices, the fully feasible region moves horizontally to the right towards a higher γ_0. This can be explained with Eq.(**??**): when $\alpha_0 > 0$, it provides a lower bound to γ_0. Note that S_0 remains unbounded, so at a fixed γ_0, every S_0 from 0 to the upper boundary critical curve $\sim \gamma_0^{-1}$ discussed before will be a fully feasible point. So in general, as syntrophy increases, systems with a high consumption rate and a low consumers abundance at equilibrium will remain feasible.

Not only does $\mathcal{F}_1^{G,A}$ change its location, it also shrinks in size: as syntrophy is increased, the set of possible consumption rate and average consumers abundance is more restricted. Figure 3.5 shows that $\text{Vol}\left(\mathcal{F}_1^G(\alpha_0)\right)$ decays exponentially as α_0 increases. We can define

[20] Note that we took a constant value α_0 (given in Methods 2.2.4) and $\gamma_0 = 0.2$. A more thorough analysis should build the optimal LRI matrix *corresponding to each* (γ_0, α_0). That would take too much time which is why we decided to keep γ_0 and α_0 constant.

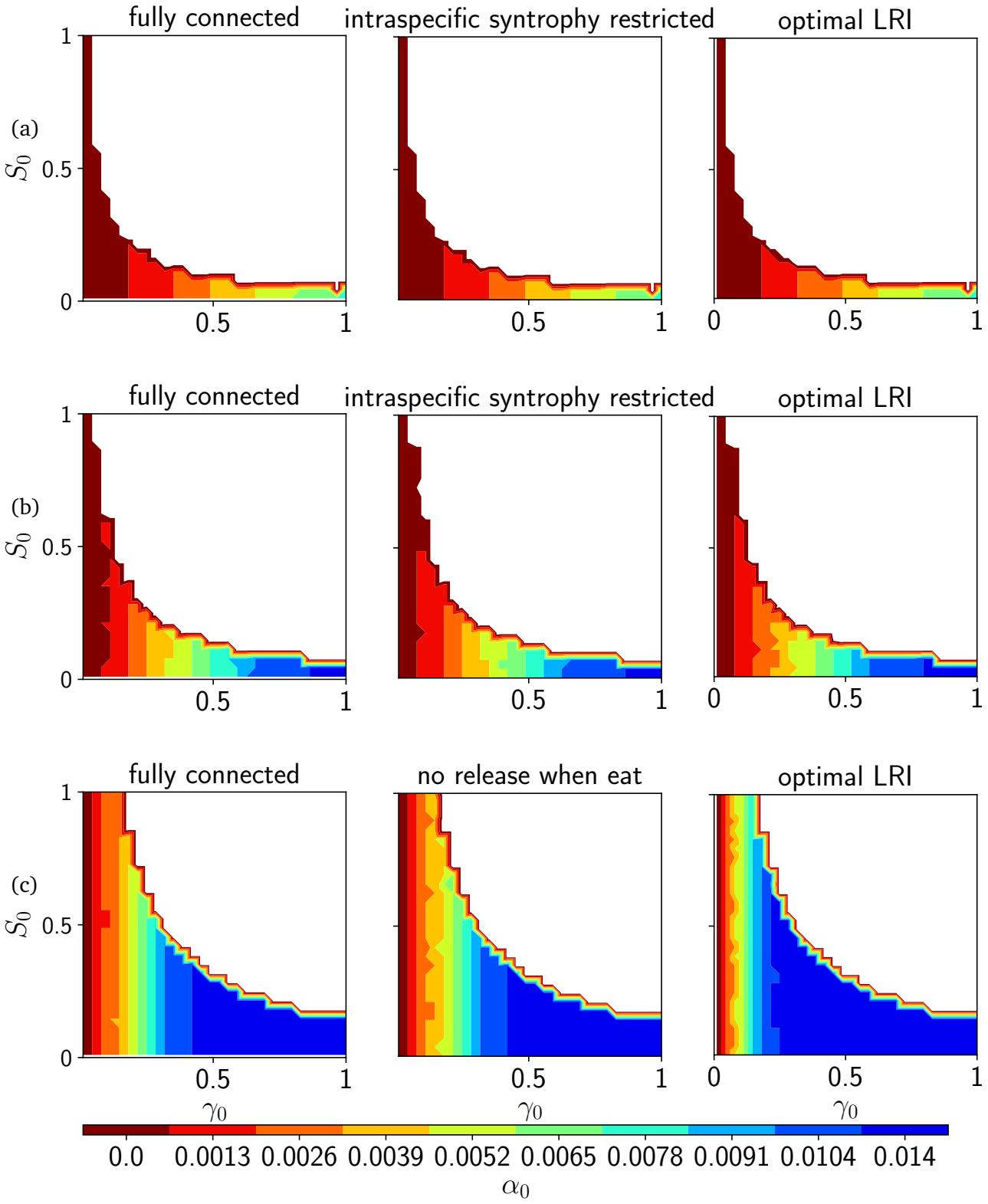

Figure 3.4: Fully feasible region in the $(\gamma_0, S_0) \in [0,1] \times [0,1]$ unit square as a function of syntrophy for different consumption matrices G: (a) $\eta_G = 0.6$, $\kappa_G = 0.32$, (b) $\eta_G = 0.35$, $\kappa_G = 0.22$ and (c) $\eta_G = 0.15$, $\kappa_G = 0.18$. The white zone corresponds to points that are never fully feasible. The colour of a given point tells until which syntrophy that point is fully feasible, *e.g.* a light blue point is fully feasible for $0 \leqslant \alpha_0 \leqslant 9.1 \times 10^{-3}$. The size of the feasibility regions depend heavily on the topology of the matrix, which makes the problem far from trivial.

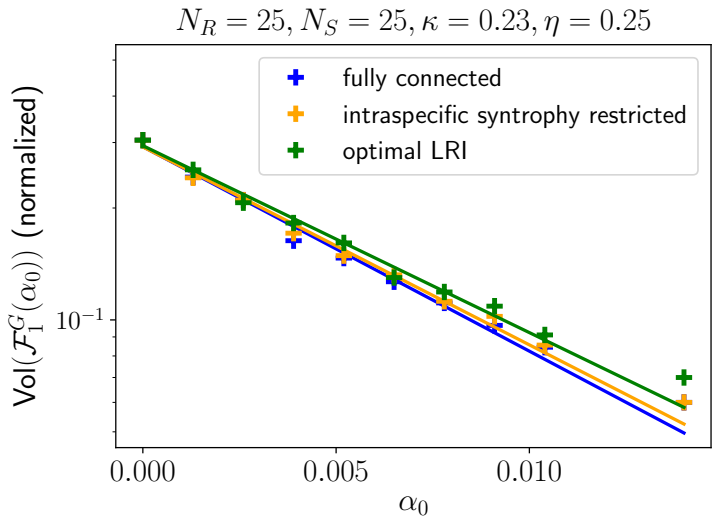

Figure 3.5: Decay of the volume of the fully feasible region $\mathcal{F}_1^G(\alpha_0)$ for a matrix consumption G with ecological overlap $\eta_G = 0.25$ and connectance $\kappa_G = 0.23$ on a logarithmic scale. The solid lines represent the exponential fit explained in the main text. The three difeerent colors represent the three different structures considered for the syntrophy matrix. The decay of Vol($\mathcal{F}_1^G$) seems well approximated by an exponential decay and the structure of the A-matrix seems to not play a large role in that prospect.

the *feasibility decay rate* by performing a fit in order to find the constants $c_1, c_2, d_F \in \mathbb{R}^+$ that satisfy best:

$$\text{Vol}\left(\mathcal{F}_1^{G,A}(\alpha_0)\right) \approx c_1 \exp\left(-d_F x\right) - c_2. \tag{3.3}$$

The value of $d_F(G, A)$ tells us how fast the feasible volume shrinks for a given consumption-syntrophy couple (G, A). In that sense $d_F(G, A)$ provides a measure of how good a consumption-syntrophy network (G, A) can sustain an increase in syntrophy[21]. If d_F is low then the system can bear an increase of syntrophy and stay feasible. If d_F is high, the system will quickly not be feasible at all anymore.

Figure 3.6 shows how d_F changes for different matrices and different structures of the syntrophy matrix. A few comments can be made. First of all it seems like the structure of A does not provide any real difference, except for G with $\eta_G = 0.15$ and $\kappa_G = 0.18$. For that specific matrix, the optimal LRI structure for A reduced by a factor of three the decay rate, compared to the fully connected case. That result is only true for this specific matrix and does not hold for the others, where $d_F(G, A)$ seems to almost not depend on A. Furthermore a very strong trend can be observed, for all matrices and all structures of A: for a given connectance, d_F is increased if the ecological overlap is increased and

[21]One could also desire to define a *critical feasible syntrophy* as the smallest syntrophy that gives a zero fully feasible volume. This would be a very interesting quantity to study and could be easily found as the root of the RHS of Eq.(3.3). We tried doing this, because the errors on d_F and on the two other fitting coefficients are already quite large (see the caption of Fig.3.6), the errors on the critical feasible syntrophy we obtained were way too large, making our results essentially meaningless.

for a given ecological overlap d_F decreases if the connectance is increased. This means clearly that systems where there is a small ecological overlap but a lot of links in the food consumption network will be favoured. Microbial communities where consumers eat from a lot of different resources (*i.e.* each their own) can sustain a larger syntrophic interaction than others.

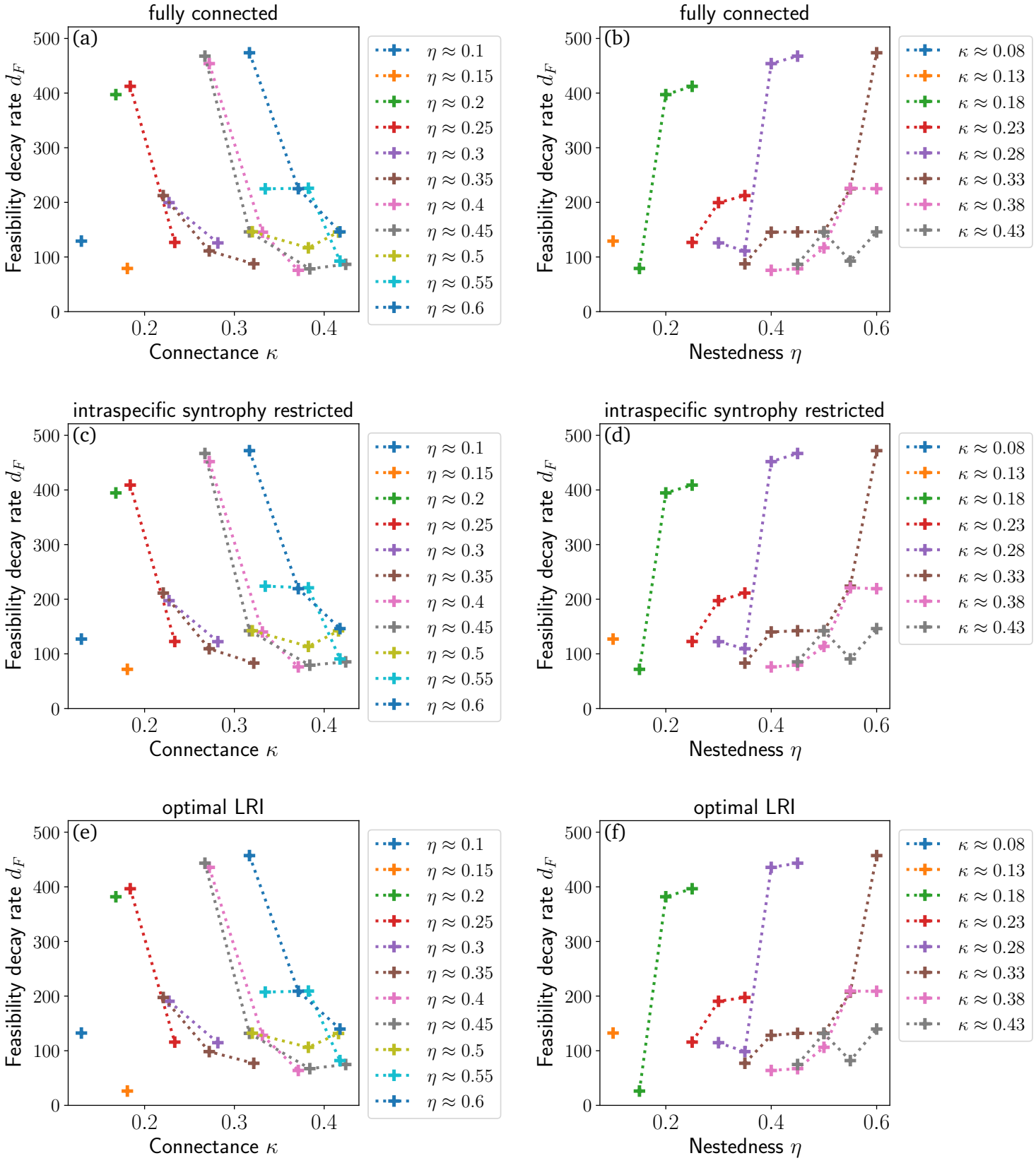

Figure 3.6: Feasibility decay rate for $N_R = 25$ and $N_S = 25$. Plots on the first column (a)-(c)-(e) show how d_F changes with connectance for a given ecological overlap, while plot on the second column (b)-(d)-(f) show how it evolves when ecological overlap is changed and connectance is kept fixed. Different structures of the A matrix are considered: (a)-(b) fully connected, (c)-(d) without intraspecific syntrophy and (e)-(f) outcome of the MC algorithm.

Common fully feasible volume

After studying the structure of each matrix individually, we can focus on the common fully feasible volume $\mathcal{F}_1^{S_M}$. Figure 3.7 show the evolution of the common feasible volume as syntrophy increases. Once again, the structure of the syntrophy matrix does not seem to play a significant role in the shape of $\mathcal{F}_1^{S_M}$. Note that at $\alpha_0 = 9.1 \times 10^{-3}$, Vol $\left(\mathcal{F}_1^{S_M}\right) = 0$ but at $\alpha_0 = 7.8 \times 10^{-3}$ the common feasible region was still non-zero. So the *critical common feasibility* α_0^F, defined as the smallest syntrophy which gives rise to a zero fully feasible volume respects $7.8 \times 10^{-3} < \alpha_0^F \leqslant 9.1 \times 10^{-3}$. It was estimated above that $\alpha_0^F \approx 0.01$, which is the right order of magnitude. Figure 3.8 shows that unsurpisingly Vol $\left(\mathcal{F}_1^{S_M}(\alpha_0)\right)$ also decays exponentially with a rate $d_F(S_M) = 480 \pm 50$ per unit of syntrophy which is a bit larger than the largest $d_F(G, A)$ observed in Fig.3.6. Without any surprise as well we observe the same shift of $\mathcal{F}_1^{S_M}$ towards points with a high γ_0 and consequently a small S_0 **Add interpretation on this?**.

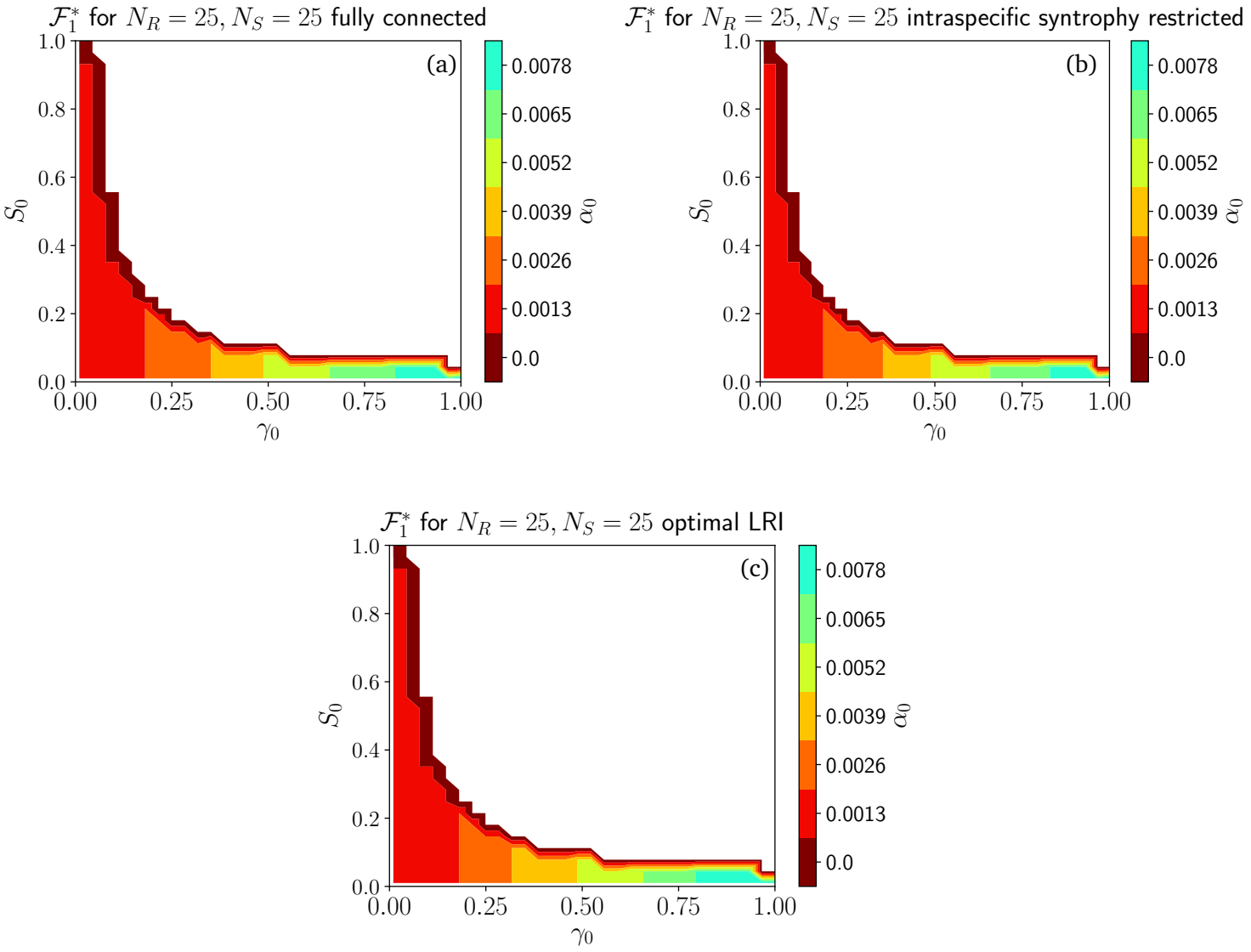

Figure 3.7: Surface plot of the fully feasible volume $\mathcal{V}^1(\alpha_0)$. The color bar on the side indicates the value of α_0 to which the surface corresponds. The white part of the plot corresponds to points that *never* are fully feasible. Note that even though it is not very clear on the figure $\mathcal{V}^1(\alpha_0^+) \subset \mathcal{V}^1(\alpha_0^-)\ \forall \alpha_0^+ > \alpha_0^-$, *i.e.* the common fully feasible region of higher syntrophy is included in the one of lower syntrophy. The different subplots correspond to different structures for the syntrophy matrix: (a) A is fully connected, (b) A has a structure such that intraspecific syntrophy is restricted and (c) A is obtained through the LRI MC algorithm described in Methods 2.2.4.

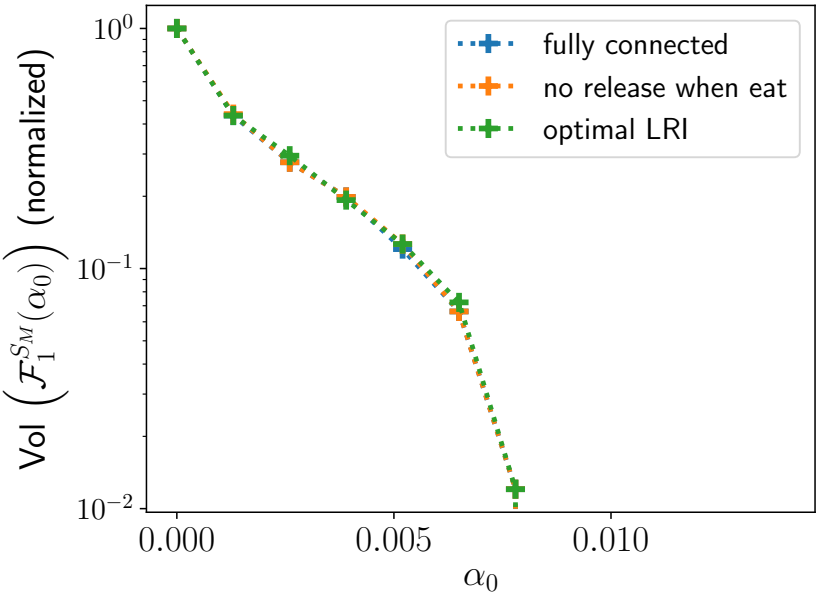

Figure 3.8: Volume of the common feasibility region $\mathcal{V}^1(\alpha_0)$ as a function of syntrophic interaction strength α_0. The curves indicate the different structures used for the syntrophy network A.

The influence of the matrix dimensions

For now we only focused on systems with the same number of consumers and resources: $N_R = N_S = 25$. But such a system lies at what has been called in the literature as May's stability bound [28], which is precisely defined as an ecological community where the number of resources is equal to the number of species. According to the competitive exclusion principle[22], an ecological system which has as many resources as consumers is the border case where coexistence, *i.e.* feasibility of our system, starts to exist, so in a way the study conducted before can be seen as a borderline case and it can be very fruitful to investigate the behaviour of systems where the number of resources has been increased to $N_R = 50$.

Figure 3.9 is the $N_R = 50$ equivalent to Figure 3.4. We may observe that increasing N_R has a non-trivial effect, which will be different for each consumption-syntrophy network. For instance, for G with $\eta_G = 0.6$ and $\kappa_G = 0.32$, adding more resources increases the maximal syntrophy bearable by the system (Fig.3.9a). On the contrary, for G with $\eta_G = 0.15$ and $\kappa_G = 0.12$, it decreases it from 1.4×10^{-2} to 7.8×10^{-3}. A global trend can however be discovered by looking at the common feasibility region (Fig.3.10). We see that, compared with the $N_R = 25$ case (Fig.3.7), an overall lesser syntrophy can be achieved. Finally Figure 3.11 shows that indeed it is really hard to predict the effect that increasing the number of resources will have on a specific consumption-syntrophy network. Indeed d_F, which we use to measure how big of a syntrophy a consumption-syntrophy network (G, A) can bear, does not have a clear pattern, at least not under the matrix metrics we chose to measure. It is a sign that this question needs a further and deeper investigation.

[22]The heavily debated and often misunderstood [34] *competitive exclusion principle*, also known as Gause's principle, states that "Complete competitors cannot coexist" [34], or more generally that "the number of consumer species in steady coexistence cannot exceed that of resources" [35].

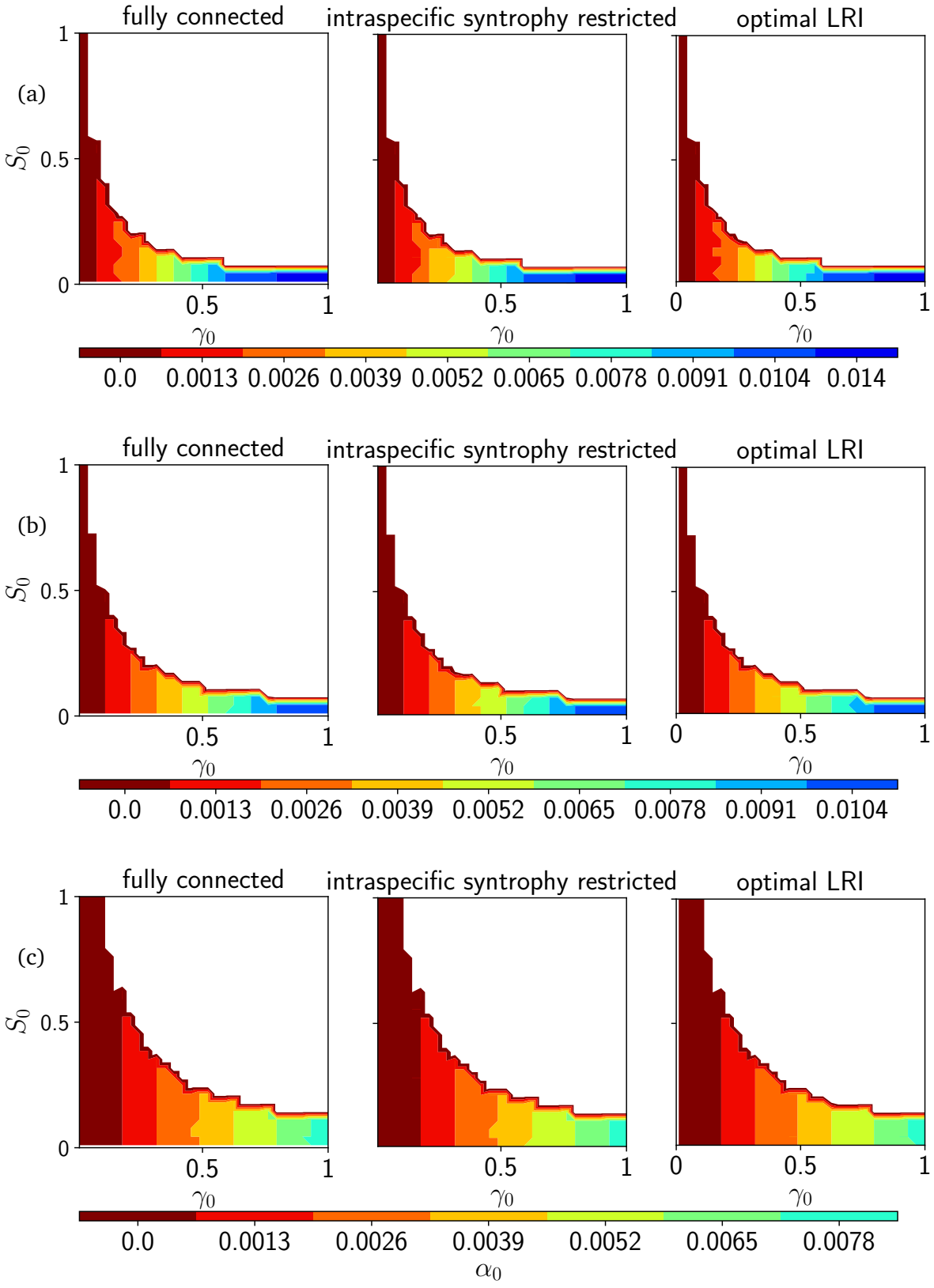

Figure 3.9: Surface colour plot of the fully feasible region $\mathcal{F}_1^{G,A}(\alpha_0)$ as a function of the syntrophy α_0 for the case $N_R = 50$, with different structures of A: fully connected (left column), no intraspecific syntrophy (middle) and LRI matrix (right). The rows correspond to different choices of the consumption matrix G: (a) $\eta_G = 0.6$ and $\kappa_G = 0.33$, (b) $\eta_G = 0.35$ and $\kappa_G = 0.23$, (c) $\eta_G = 0.15$ and $\kappa_G = 0.12$. These are matrices with similar properties than Fig.3.4, except that the number of resources is here doubled. This affects $\mathcal{F}_1^{G,A}(\alpha_0)$ quite drastically.

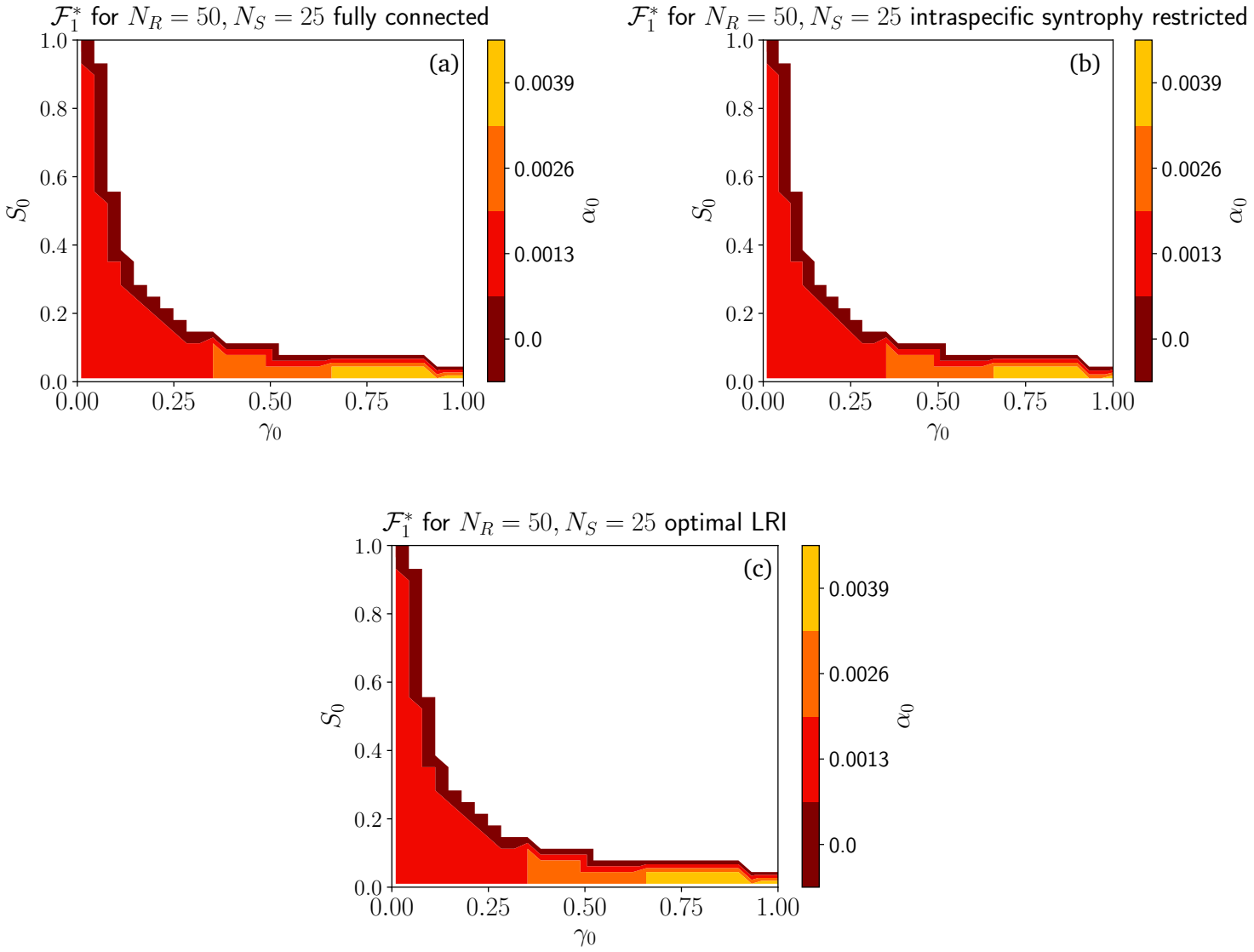

Figure 3.10: Common feasibility region $\mathcal{F}_1^{S_M}(\alpha_0)$ for $N_R = 50$ and $N_S = 25$, to compare with 3.7. We considered different structures of the syntrophy matrix: (a) fully connected, (b) intraspecific syntrophy restricted and (c) LRI matrix. As the number of resources increases, the feasibility volume for a given α_0 decreases.

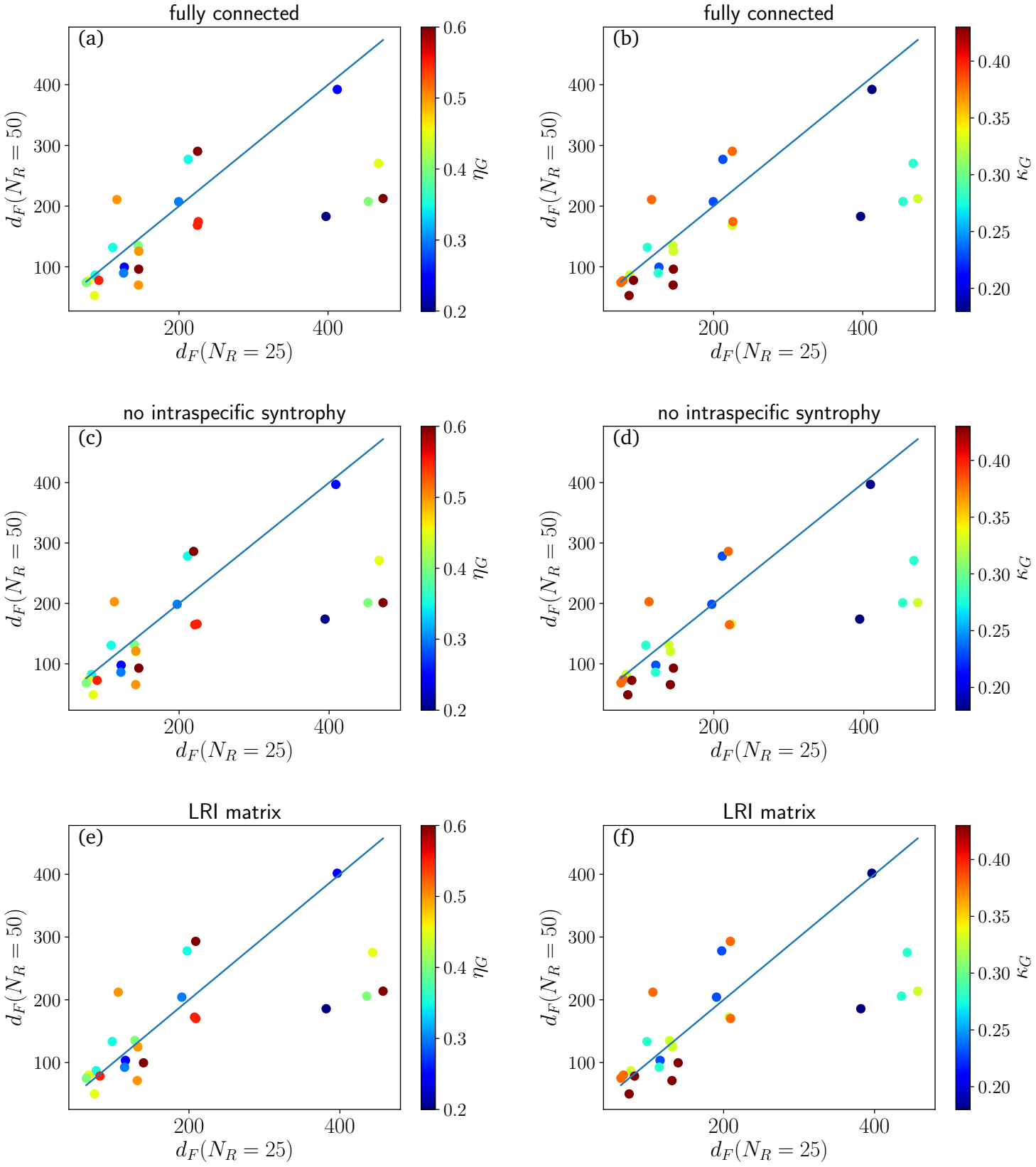

Figure 3.11: Comparison of the feasibility decay rates obtained for matrices with same ecological and connectance, at $N_R = 25$ and $N_R = 50$. The color of the points indicates their ecological overlap (left) or connectance (right). The straight line represents equal $d_F(N_R = 50)$ and $d_F(N_R = 25)$. Different structures for the syntrophy matrix A are considered: (a) fully connected, (b) no intraspecific syntrophy and (c) LRI matrix. No clear tendency on what effect increasing resources will have can be drawn.

3.2 Dynamical stability

3.2.1 LRI regime – Outcome of the Monte Carlo algorithm

Methods 2.2.4 explains how we designed an algorithm whose goal is to find the shape of the syntrophy matrix A that should bring us closer to a dynamically stable regime for a given consumption matrix G. It works by minimizing an energy $E(A, G)$ given by Eq.(2.102) and should, for a given G, yield an A matrix with the following properties:

- Each diagonal element of AG is minimized. Note that $(AG)_{\mu\mu} = \sum_{i=1}^{N_S} A_{\mu i} G_{i\mu}$ corresponds to the number of species that both consume and release resource μ.
- Outside of the diagonal, we should have $\frac{\alpha_0}{\gamma_0 R_0} AG \approx G^T G$. A direct ecological interpretation is more difficult to draw. For a couple of different resources (μ, ν), $(AG)_{\mu\nu} = \sum_{i=1}^{N_S} A_{\mu i} G_{i\nu}$ is the number of species that both consume resource ν and release resource μ and $(G^T G)_{\mu\nu} = \sum_{i=1}^{N_S} G_{i\mu} G_{i\nu}$ is the number of consumers that eat both ν and μ.

So intuitively the LRI MC algorithm should give us a syntrophy matrix that both limits intraspecific syntrophy and such that for every couple of different resources (μ, ν), the number of consumers that eat both μ and ν is proportional to the number of consumers that eat μ and release ν. The proportionality constant, which is the same for all (μ, ν) is equal to the ratio of the syntrophy and consumption interactions. Since the connectance is and the dimensions of A are fixed, the number of links of A is already decided and the algorithm simply decides the optimal cells to put them in. Figure 3.12 shows that typically the algorithm will put links in a cell (μ, i) if (i, μ) is zero, meaning that not only intraspecific syntrophy tries to be avoided but also species that consume a lot of resources will tend to release few of them and vice-versa. Figure 3.13 shows that indeed we obtain for a given G matrix a syntrophy matrix A such that the two requirements above are best satisfied. Note that the algorithm works better for matrices with a low connectance. It is worth noticing that this procedure produces highly nested syntrophy matrices (Fig.3.14) where only a few species produce most of the syntrophic flow. The obtained matrices have an even larger nestedness if we increase the number of resources.

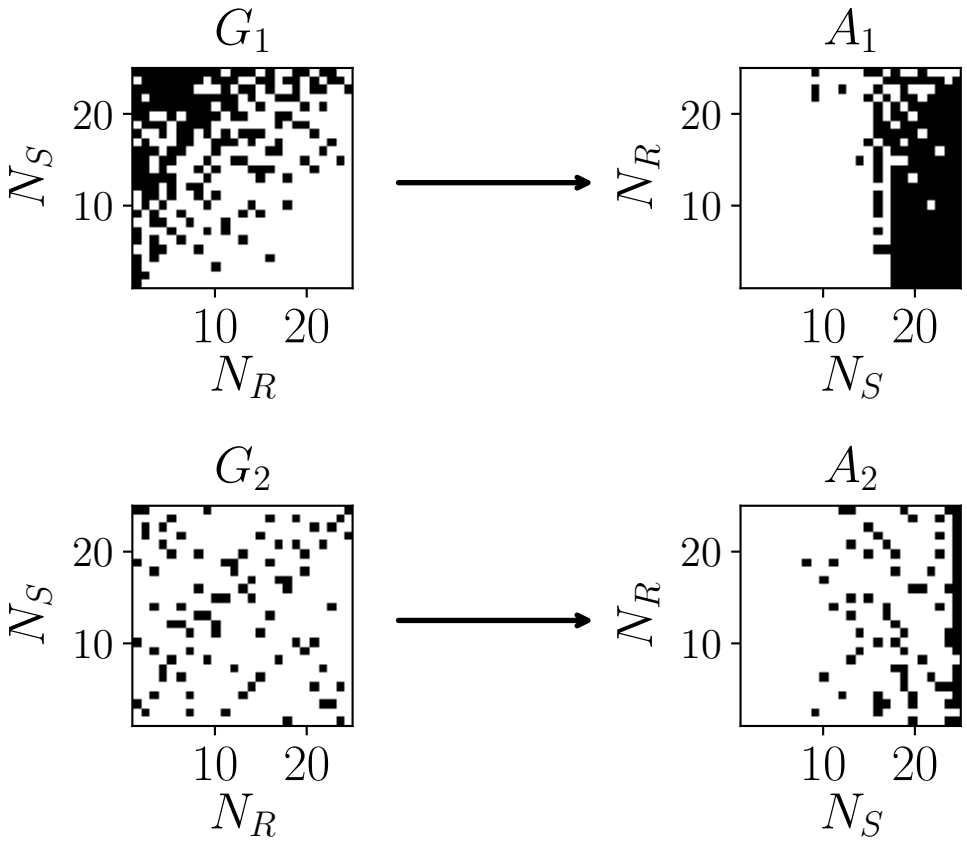

Figure 3.12: Typical shape of the consumption G_i and syntrophy A_i matrices. The white cells symbolize a zero matrix element and the black cells, a one. A_i here is the outcome of the LRI MC algorithm described in Methods 2.2.4. The first row has a consumption matrix with $\eta_1 = 0.6$ and $\kappa_1 = 0.32$, the LRI MC solver gives rise to a syntrophy matrix with same connectance and ecological overlap ≈ 0.85. The second row has G_2 with $\eta_2 = 0.1$ and $\kappa_2 = 0.13$ and the corresponding syntrophy matrix A_2 has ecological overlap ~ 0.42. We observe that under this optimisation, species that consume few resources end up releasing many and the other way around.

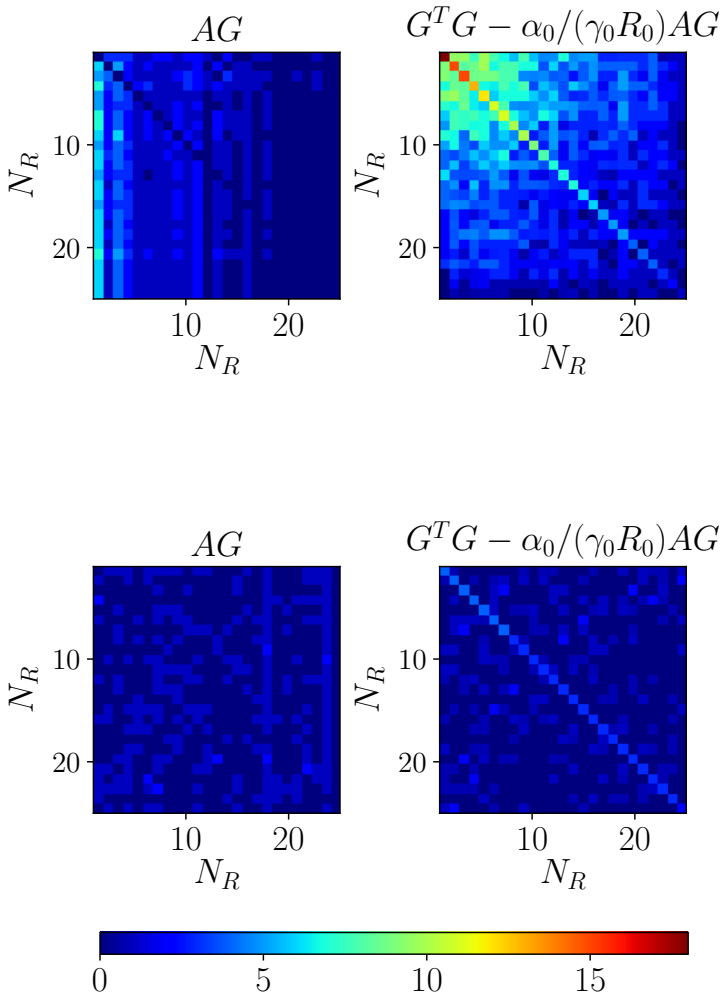

Figure 3.13: Plotting of AG and $G^TG - \alpha_0/(\gamma_0 R_0)AG$. The A and G matrices of the first and second rows correspond to the respective A and G of Figure 3.12. As expected, we obtain an A such that intraspecific coprophagy is limited (the diagonal of AG is roughly zero) and, outside the diagonal, $G^TG - \alpha_0/(\gamma_0 R_0)AG \approx 0$. Both relations are better satisfied for consumption (and hence syntrophy) matrices with a low connectance.

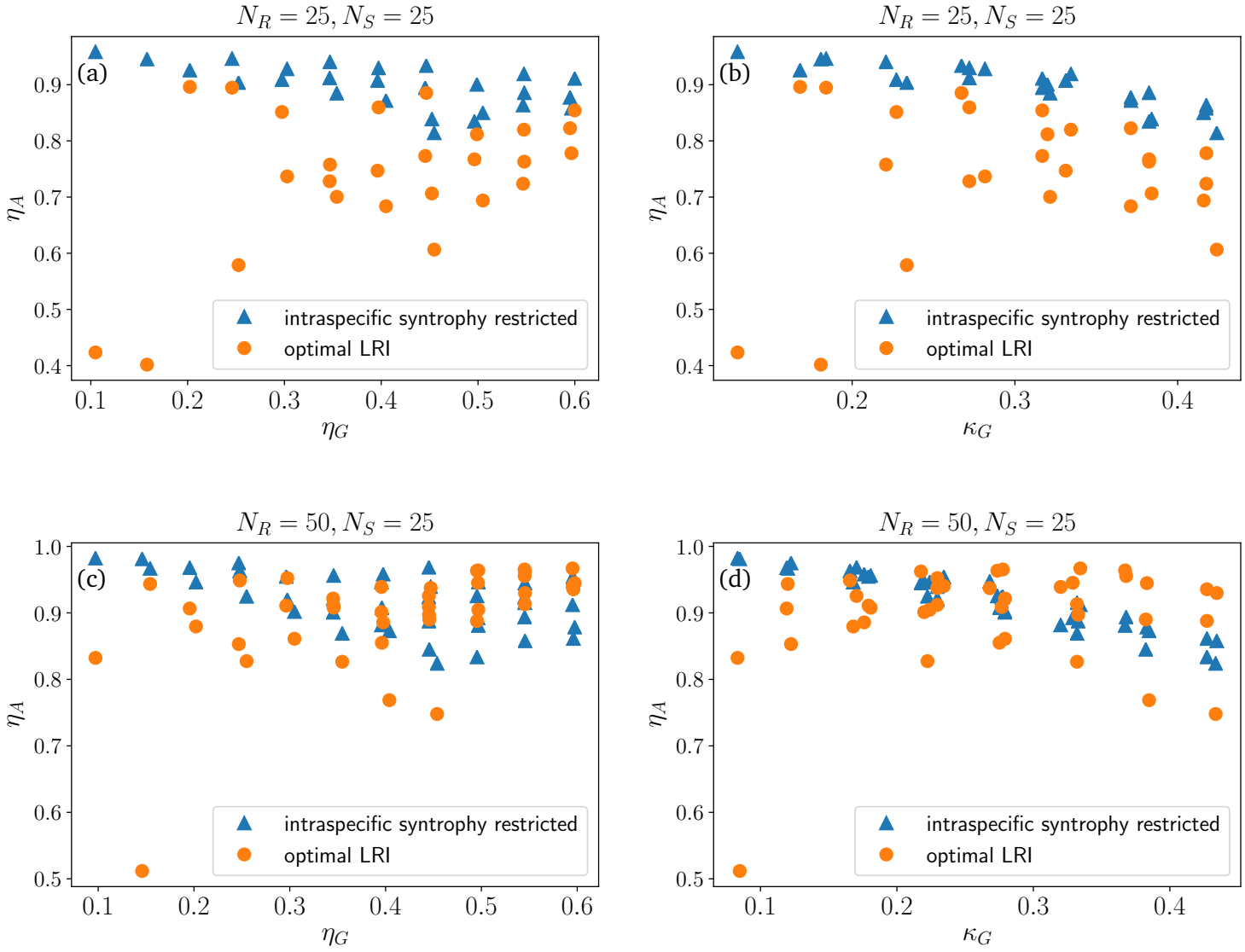

Figure 3.14: Properties of the syntrophy matrix against the consumption matrix. (a)-(c) Ecological overlap of A as a function of the ecological overlap of G for $N_S = 25$ and $N_R = 25$ (a) or $N_R = 50$ (c). (b)-(d) Ecological overlap of A as a function of the connectance of G for $N_S = 25$ and $N_R = 25$ (b) or $N_R = 50$ (d). The nestedness of the "intraspecific syntrophy restricted" is also plotted as a matter of comparison. As η_G or κ_G increase, the two results will without surprise give matrices with similar properties. **Explain this?**

3.2.2 Fully dynamically stable region

The same way we studied the fully feasible volume $\mathcal{F}_1^{G,A}(\alpha_0)$, we investigate now the behaviour of its special subset, the locally fully dynamically stable region $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$. A formal definition of $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ is provided in Methods 2.2.1. Intuitively, $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ corresponds to the set of all metaparameters $m \in \mathcal{M}$ such that $\mathcal{A}(m, G, A)$ is a feasible, locally dynamically stable system with probability 1. Since we require $\mathcal{A}(m, G, A)$ to be feasible, it is clear that $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ is indeed a subset of $\mathcal{F}_1^{G,A}(\alpha_0)$.

Figure 3.15 shows that $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ is geometrically more complex than $\mathcal{F}_1^{G,A}(\alpha_0)$ (which is after all not suprising, the question of dynamical feasibility is harder than the one of feasibility!). It may sometimes have holes, even without syntrophy, and sometimes not, even for matrices that are topologically very close. Compare for instance Fig.3.15a with Fig.3.15c, these two networks have the same ecological overlap, but even though their connectance is very similar, their fully locally dynamically stable regions have a very different shape: one of them can sustain only a tiny bit of syntrophy before becoming dynamically unstable (Fig.3.15a) while the second can endure basically any feasible syntrophic interaction (Fig.3.15c). A general trend is hence harder to find but it seems that points with a larger γ_0 are in general more dynamically stable (more on this below). We required that a system had to be feasible in order to be dynamically stable, which is the mathematical equivalent of $\mathcal{D}_{L,1}^{G,A}(\alpha_0) \subset \mathcal{F}_1^{G,A}(\alpha_0)$, *i.e.* local dynamical stability implies feasibility. We may also ask the reverse question, does feasibility imply local dynamical stability? The answer to this is, again, unsurpisingly, "it depends on the matrix", as shows Figure 3.15. For instance, for G with $\kappa_G = 0.13$ and $\eta_G = 0.1$, we have Vol $\left(\mathcal{D}_{L,1}^{G,A}(\alpha_0)\right) <$ Vol $\left(\mathcal{F}_{L,1}^{G,A}(\alpha_0)\right) \forall \alpha_0$, which means that for this consumption matrix feasibility does not imply stability. The fully connected case gives a larger dynamically stable volume than the regime without intraspecific syntrophy which is itself better than the LRI regime. This hints that the LRI regime, despite what it was designed for, apparently does not give better results than other structures of A. On the contrary, for G with $\kappa_G = 0.32$ and $\eta_G = 0.6$, both volumes are equal at every syntrophy that is feasible, for the three structures of A considered, which shows that for this specific matrix, feasibility implies local dynamical stability. A good way to measure how systems react to syntrophy is to compute the *critical locally dynamically stable syntrophy* $\alpha_0^D(G, A)$ (see Methods **??**). An easy way this can be done is by getting some points of the volume of $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ curve and finding its intercept to zero. Figure 3.17 shows the typical shrinkage of the fully locally dynamically stable volume. To find $\alpha_0^D(G, A)$ for each G, we fit with a linear function the last four points of the curve corresponding to Fig.3.17 and find its intercept to zero $\alpha_0^D(G, A)$. Figure 3.18 shows a very interesting and clear behaviour[23]: for a given connectance of the consumption matrix, systems that can sustain the largest syntrophy have a small ecological overlap. And for a given ecological overlap, systems with a larger connectance will stay stable longer under the action of syntrophy. In the end, optimal systems have a small ecological overlap and a large connectance: many resources are eaten by the consumers, but they do not share them.

[23]Those results have to be taken with a grain of salt because we did not check whether $\alpha_0^D(G, A)$ was feasible, *i.e.* the actual critical locally dynamically stable syntrophy is the minimum between the value measured in Fig.3.18 and the largest feasible α_0 for that couple (G, A).

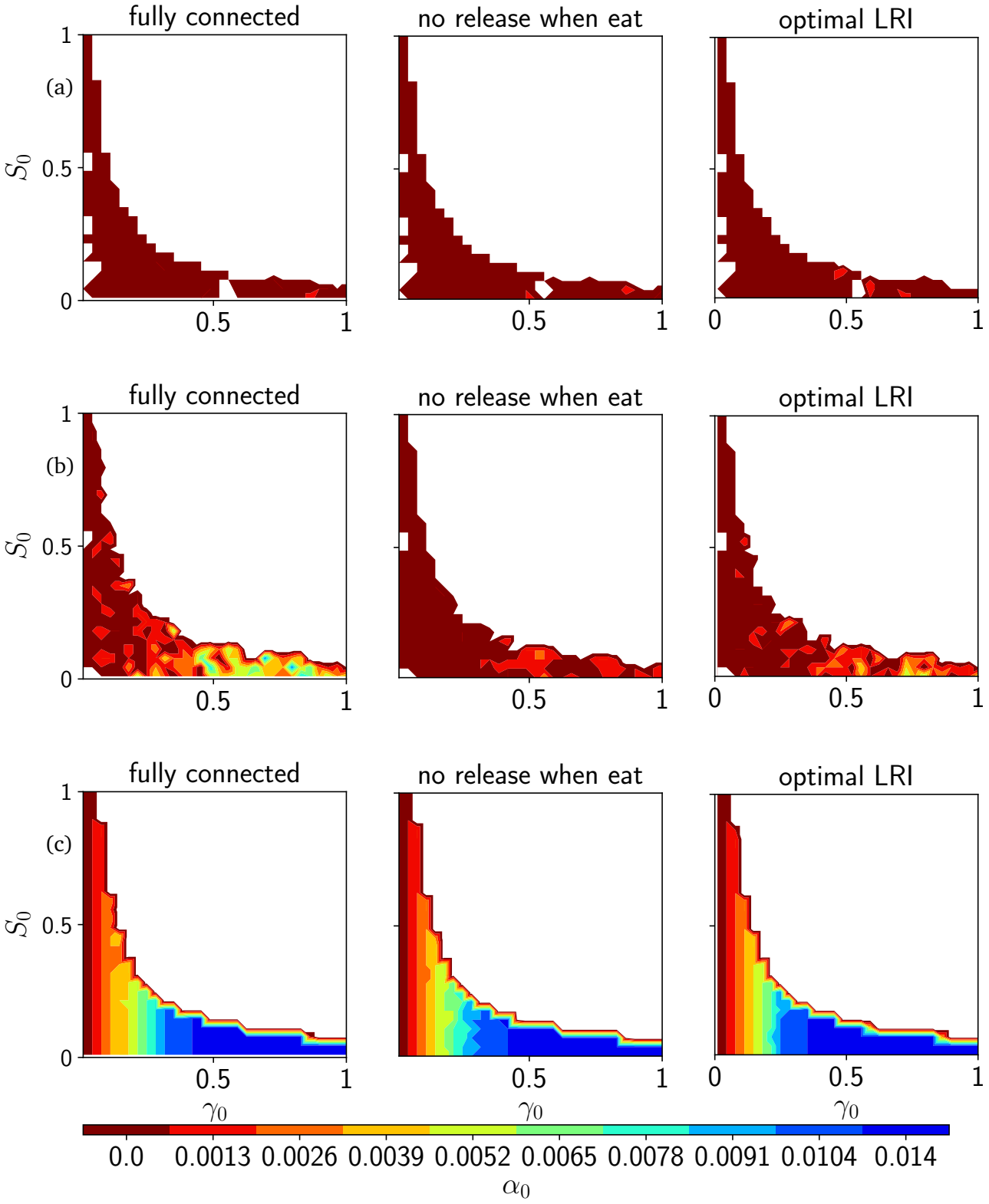

Figure 3.15: Locally fully dynamically stable region $\mathcal{D}_{L,1}^{G,A}$ as a function of syntrophy for different matrices G. The white zone corresponds to points that are never fully locally dynamically stable. The colour of a given point tells until which syntrophy that point is fully locally dynamically stable, *e.g.* a green point is fully locally dynamically stable for $0 \leqslant \alpha_0 \leqslant 6.5 \times 10^{-3}$. Row (a) corresponds to G with $\eta_G = 0.35$ and $\kappa_G = 0.23$, (b) has $\eta_G = 0.35$ and $\kappa_G = 0.33$ and (c) $\eta_G = 0.35$ and $\kappa_G = 0.27$. Even at fixed ecological overlap, different connectances of G give rise to completely different systems in terms of local dynamical stability.

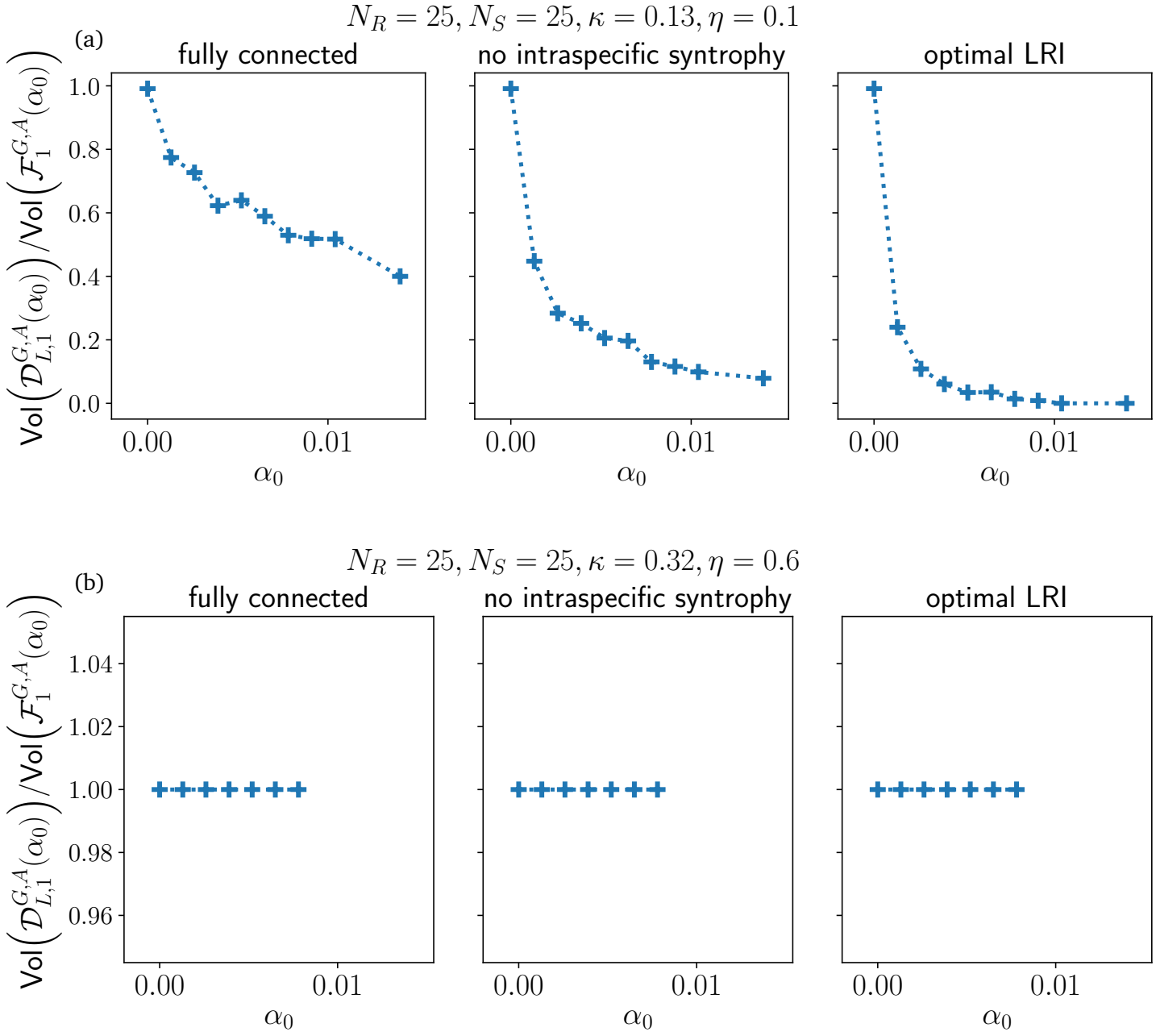

Figure 3.16: Ratio of the size of the fully dynamically stable volume and the fully feasible volume for two consumption matrices G (a) with $\eta_G = 0.1$ and $\kappa_G = 0.13$, (b) with $\eta_G = 0.6$ and $\kappa_G = 0.32$. We observe different behaviours for different matrices: for (a) feasibility does not imply local dynamical stability even without syntrophy (it is barely feasible but the ratio is a bit below one for $\alpha_0 = 0$). On the other hand, for (b) feasibility implies local dynamical stability, indeed both regions have the same volume and since $\mathcal{D}_{L,1}^{G,A}(\alpha_0) \subset \mathcal{F}_1^{G,A}(\alpha_0)$, we conclude that both are equal.

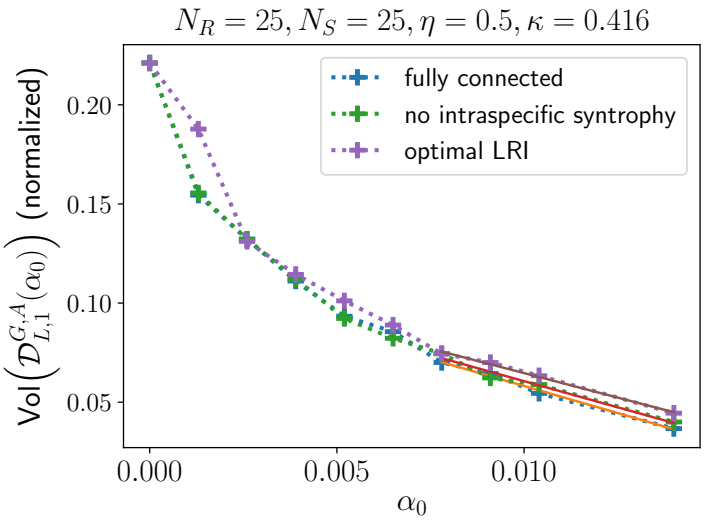

Figure 3.17: Evolution of the volume of $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ with α_0 for a consumption matrix G with $\eta_G = 0.5$ and $\kappa_G = 0.42$. The size of the fully locally dynamically feasible region shrinks as syntrophy increases. A linear fit is performed on the last four points to determine $\alpha_0^D(G, A)$ (see Fig.3.18), the point where this curves reaches zero.

Finally, one can take a look at the *common fully locally dynamically stable region*, which is the intersection of the $\mathcal{D}_{L,1}^{G,A}(\alpha_0) \forall (G, A) \in S_M$ (Figure 3.19). Because of the fractured and heterogenous nature of each $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$, we observe a very fractured and small common fully locally dynamically stable region, which is the same for all structures of A considered. It has a non-zero volume for $\alpha_0 = 0$, but for the next point investigated $\alpha_0 = 1.3 \times 10^{-3}$, no point is fully locally dynamically stable for every matrix considered, which means that the critical common syntrophy is smaller than this.

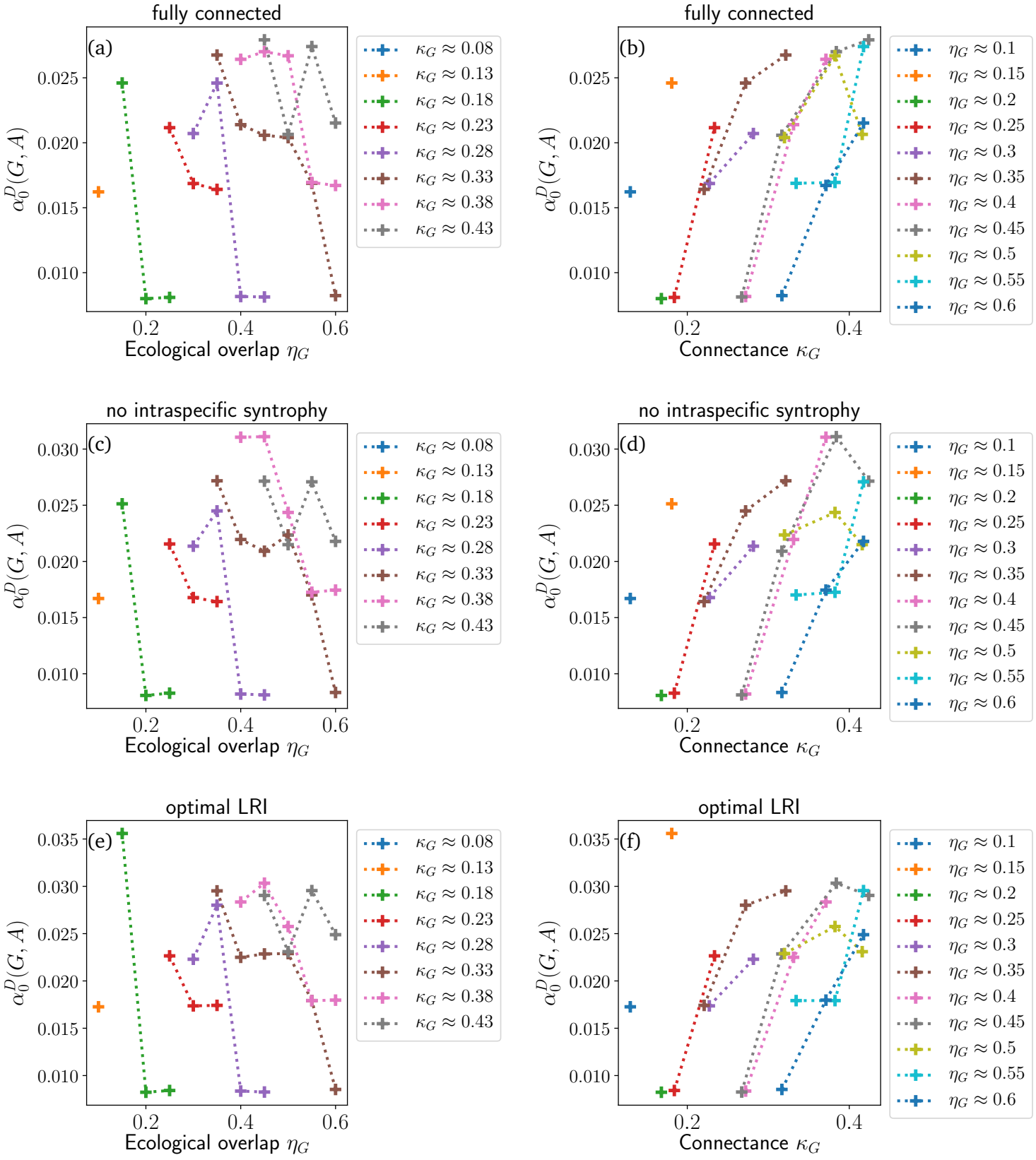

Figure 3.18: Critical syntrophy α_0^D, defined as the smallest syntrophy for which we can still find metaparameters that will give rise to fully dynamically stable systems. How α_0^D is estimated is explained in the main text. Errors on α_0^D are not plotted but are at most around 10%. (a)(c)(e) Evolution of α_0^D with ecological overlap η at different connectance. (b)(d)(f) Evolution of α_0^D with connectance κ for different ecological overlap. We observe a strong trend: for a given connectance, α_0^D decreases as ecological overlap increases. Also, for a given ecological overlap, α_0^D increases as connectance is increased.

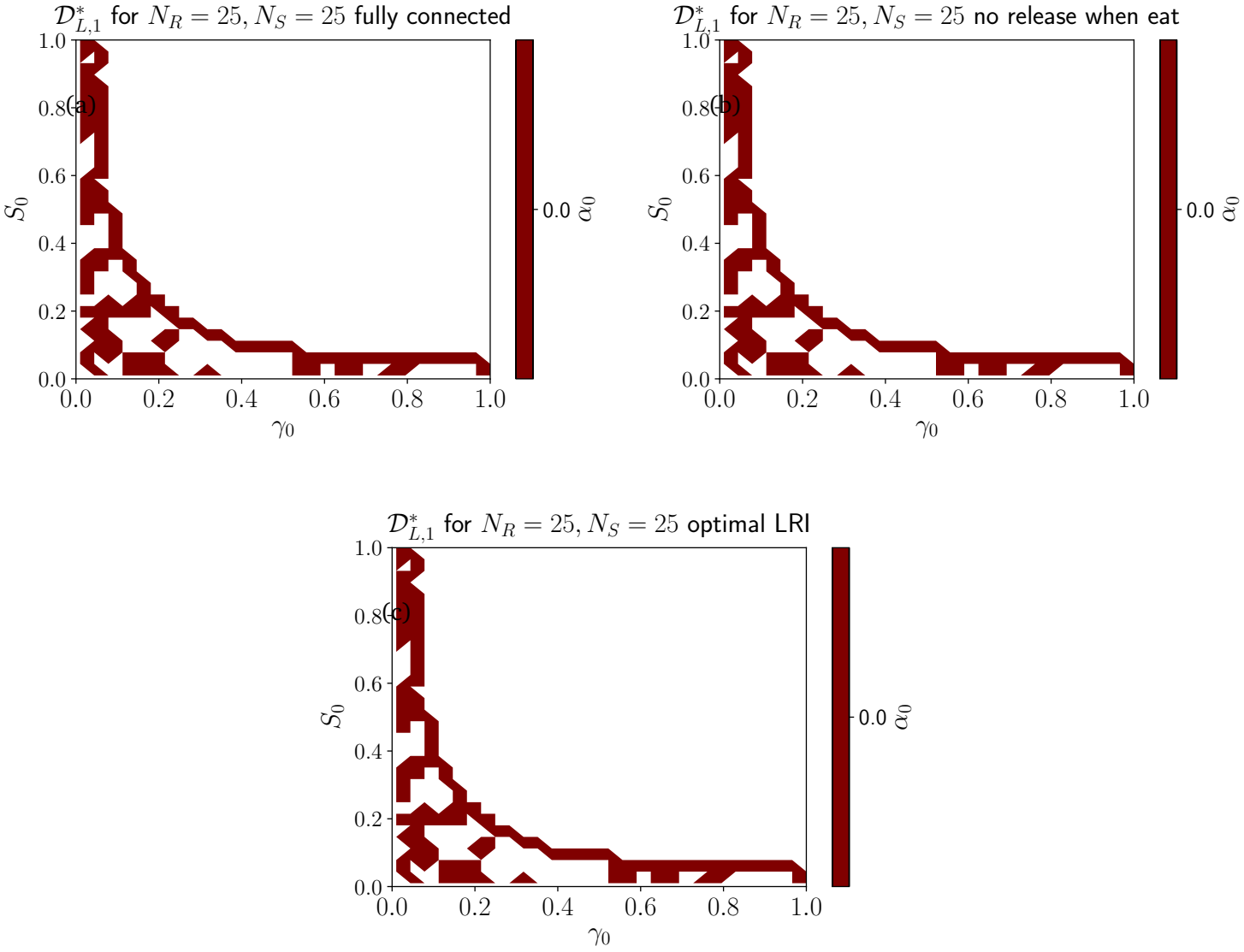

Figure 3.19: Common full local dynamical stability volume for different A structures: (a) fully connected, (b) no intraspecific syntrophy and (c) LRI algorithm. The points coloured in dark red give rise to locally dynamically stable systems with probability 1 for *all the matrices considered*. Very few spots verify this property when there is no syntrophic interaction, and no point gives rise to a fully dynamically stable system for $\alpha_0 = 1.3 \times 10^{-3}$. This is independent of the structure of A that we chose. The white points never give rise to fully dynamically stable systems.

3.2.3 Largest eigenvalue of the jacobian

Equation (**??**) from Methods **??** gives a relationship that the metaparameters should approximately follow in order to give rise to locally dynamically stable systems. Although strictly speaking it is only valid for the case where both G and A are fully connected, we expect it to work as well when G and A are *not too far away* from the fully connected case. It tells us that in order to get more local dynamically stable systems you should:

- decrease N_S, l_0 **WEIRD RESULT: would expect that increasing l0 would make systems more dynamically stable (observed in simulations I think)** or α_0,
- if $\alpha_0 - \gamma_0 R_0 < 0$, increase N_R, σ_0 and γ_0,
- be careful in how you handle S_0: increasing S_0 reduces the l_0^2/S_0 term but increases the $N_S^2\alpha_0^2 S_0$ term. It is very easy to show (Appendix 4.1.2) that if $S_0 > l_0/(N_S\alpha_0)$ it should be decreased, and otherwise it should be increased until it reaches $l_0/(N_S\alpha_0)$.

Combining these with the feasibility conditions Eq.(**??**) we expect that – for all other metaparameters fixed – systems get more and more locally dynamically stable as γ_0 is increased and S_0 is the largest possible. In short, points at the upper border of $\mathcal{D}_{L,1}^{G,A}$ should have a lower and lower $\mathrm{Re}\,(\lambda_1)$ as γ_0 increases. Figure 3.20 shows that indeed this trend is followed. This tells us that if we keep the consumption flux $N_S\gamma_0 S_0$ constant, increasing γ_0 (and hence decreasing S_0) will give rise to more stable systems. Notice that contrarily to the prediction made above, increasing α_0 does not decrease stability but increases the maximal $|\mathrm{Re}\,(\lambda_1)|$ observed as shows Fig.3.21a. This is coupled with the shrinkage of the fully locally dynamically stable volume seen on Fig.3.21b. This means that overall increasing syntrophy makes the system *more stable* but at *fewer points*. This hints that systems in a high syntrophic regime, where consumers produce a lot of resources, should be very fine-tuned and occur for very specific consumption strength and average abundance of consumers.

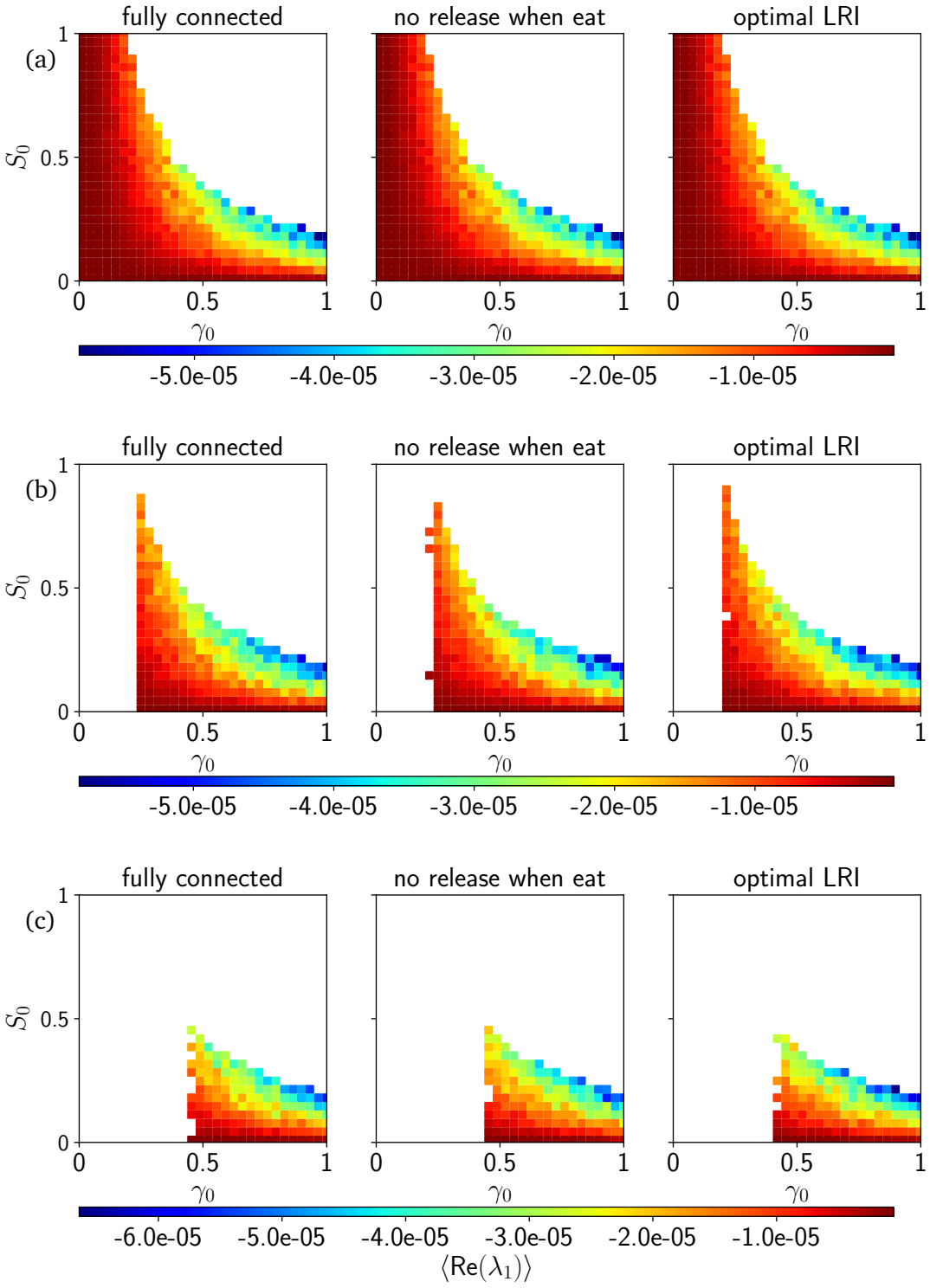

Figure 3.20: Largest real eigenvalue $\mathrm{Re}(\lambda_1)$ averaged over 300 realisations for each (γ_0, S_0) points for the consumption matrix G with consumers overlap $\eta_G = 0.1$ and connectance $\kappa_G = 0.13$. The white points correspond to not fully dynamically stable systems. Each row corresponds to a different syntrophy value (a) $\alpha_0 = 0$ (no syntrophic interaction), (b) $\alpha_0 = 3.9 \times 10^{-3}$ and (c) $\alpha_0 = 7.8 \times 10^{-3}$. The first column corresponds to the regime where A is fully connected, the second where A forbids intraspecific syntrophy and the third is the outcome of the MC algorithm. As syntrophy increases, the size of the fully dynamically stable region decreases. Furthermore, the boundary points close to the $\gamma_0 \sim S_0^{-1}$ curve are the most stable in every situation.

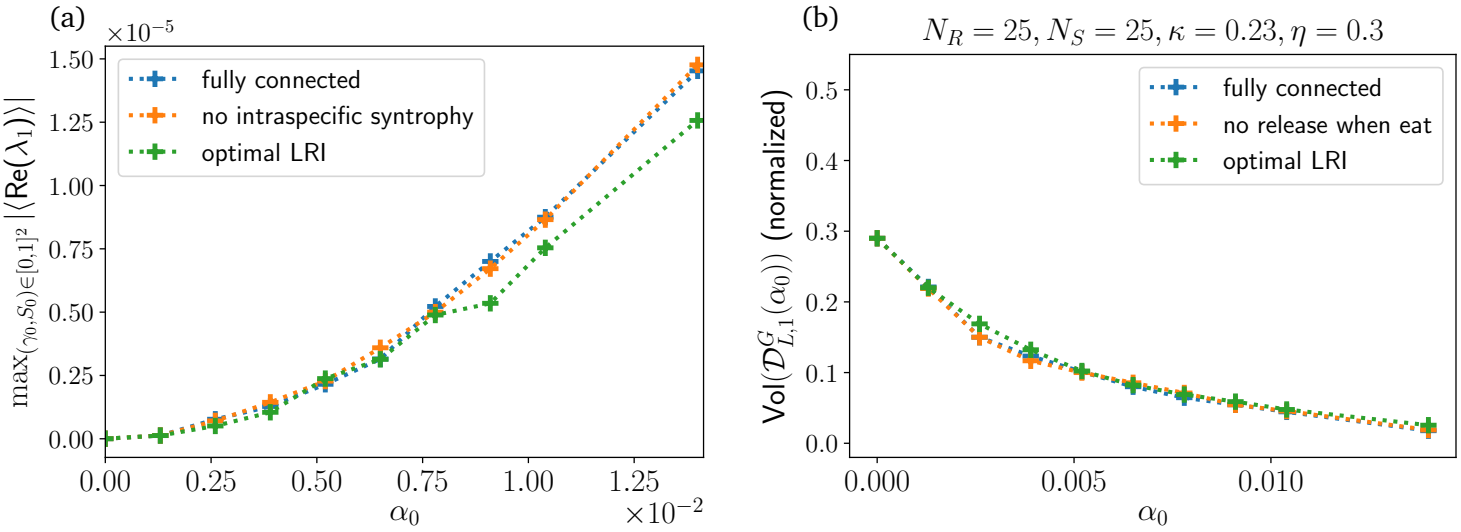

Figure 3.21: For a consumption matrix G with $\eta_G = 0.3$ and $\kappa_G = 0.23$. (a) Evolution of the maximal $|\langle \mathrm{Re}\,(\lambda_1)\rangle|$ observed in the $(\gamma_0, S_0) \in [0,1]^2$ region. The maximal eigenvalue increases in magnitude, making the system more dynamically stable, as syntrophy increases. That trend is true for all matrices we considered. (b) Volume of $\mathcal{D}^G_{L,1}(\alpha_0)$. As syntrophy increases, fewer and fewer points become fully dynamically stable. For both figures, the different lines show the different stand for the different structure of the syntrophy matrix that we considered.

3.2.4 The influence of the matrix dimension

As said above, because of Eq.(**??**), we expect stability to increase when the number of resources is increased for a fixed number of consumers. The following subsection shows what happens when the number of resources is doubled $N_R = 25 \rightarrow N_R = 50$ and every other metaparameter, as well as the number of consumers, keeps the same value as before.

Figure 3.22 shows that the effect of adding resources can be quite dramatic on the stability of the system. For that specific matrix for instance, adding resources allowed for a way larger $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ at each α_0. It even allows full dynamical stability in the feasibility region for larger syntrophic interactions than before. Even though more syntrophy can be sustained, it seems that the volume of $\mathcal{D}_{L,1}^{G,A}(\alpha_0)$ is smaller at $N_R = 50$ than at $N_R = 25$ (compare for instance Fig.3.23b with Fig.3.21b). This is compensated by the fact that way larger eigenvalues are observed at $N_R = 50$: although there are (a bit) fewer equilibrium points, these are more stable (compare Fig.3.23a and Fig.3.21a). This is a trend that we believe holds for all the matrices considered but a more thorough investigation should be conducted before claiming those results to be absolutely true.

Since matrices are individually more locally dynamically stable, the *common locally dynamically stable* region is larger with $N_R = 50$ resources than with $N_R = 25$. The difference between Figure 3.19 and Figure 3.24 is very striking: when the number of resources is increased not only is $\mathcal{D}_{L,1}^{S_M}(\alpha_0)$ larger, it also can overall bear a larger syntrophy. Remember that for $N_R = 25$, the critical common syntrophy was between 0 and 1.3×10^{-3}, and for $N_R = 50$, it is way larger, between 3.9×10^{-3} and 5.2×10^{-3}. **Is there anything more to add? I have some plots but I am not sure if they are the most relevant**

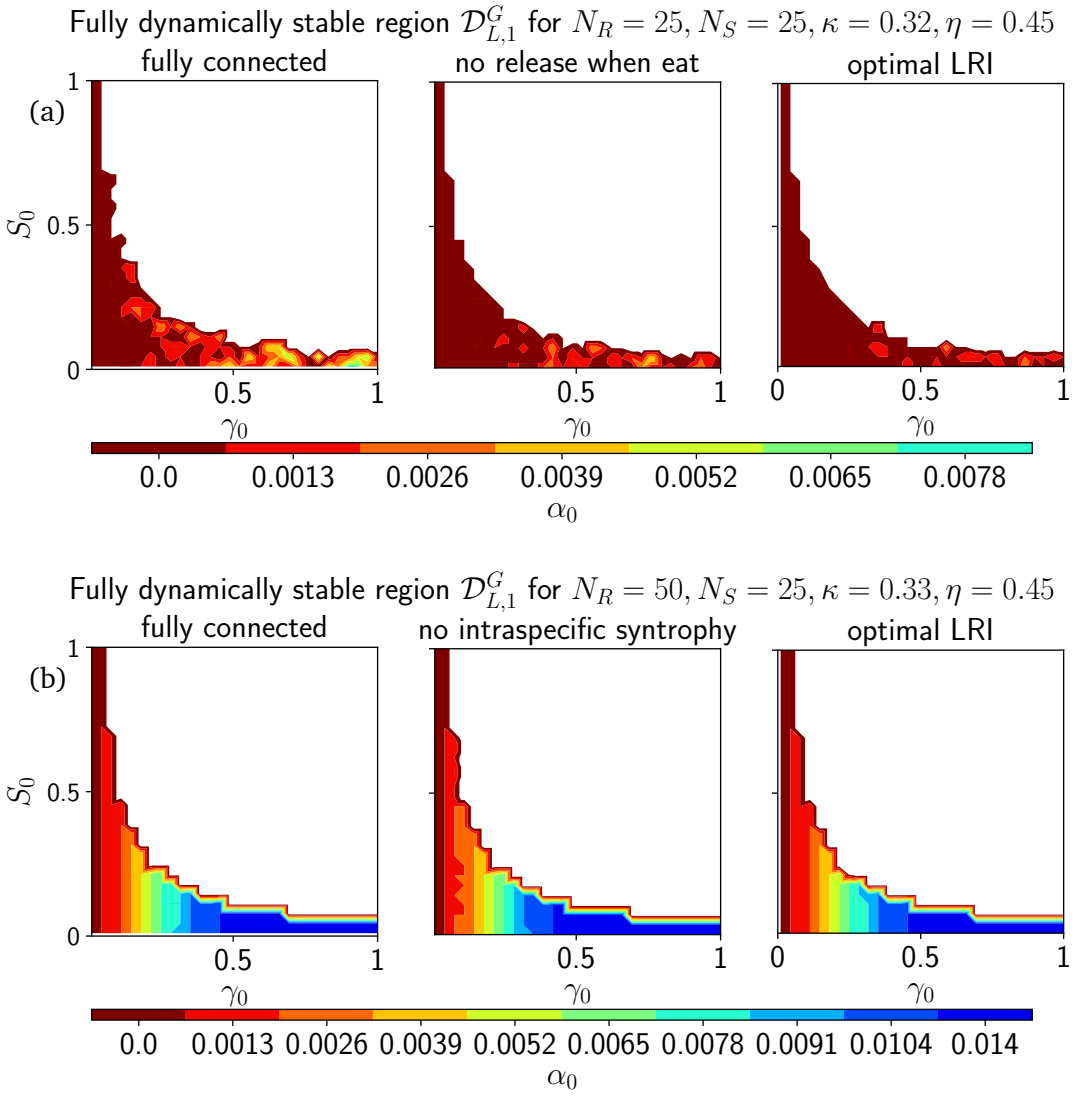

Figure 3.22: Fully dynamically stable region $\mathcal{D}^{G,A}_{L,1}$ with the three different structures of A considered: fully connected (left), no intraspecific syntrophy (middle) and LRI matrix (right). The two matrices have the same ecological overlap and connectance, only the number of resources changes. (a) G has $N_R = 25$, $N_S = 25$ and $\kappa_G = 0.32$ and $\eta_G = 0.45$. (b) G has $N_R = 50$, $N_S = 25$ and $\kappa_G = 0.33$ and $\eta_G = 0.45$. The fact that more points can sustain an increased syntrophy is a trend for most of the matrices of the set. **still check this: the others are more less the same or a tad less.**

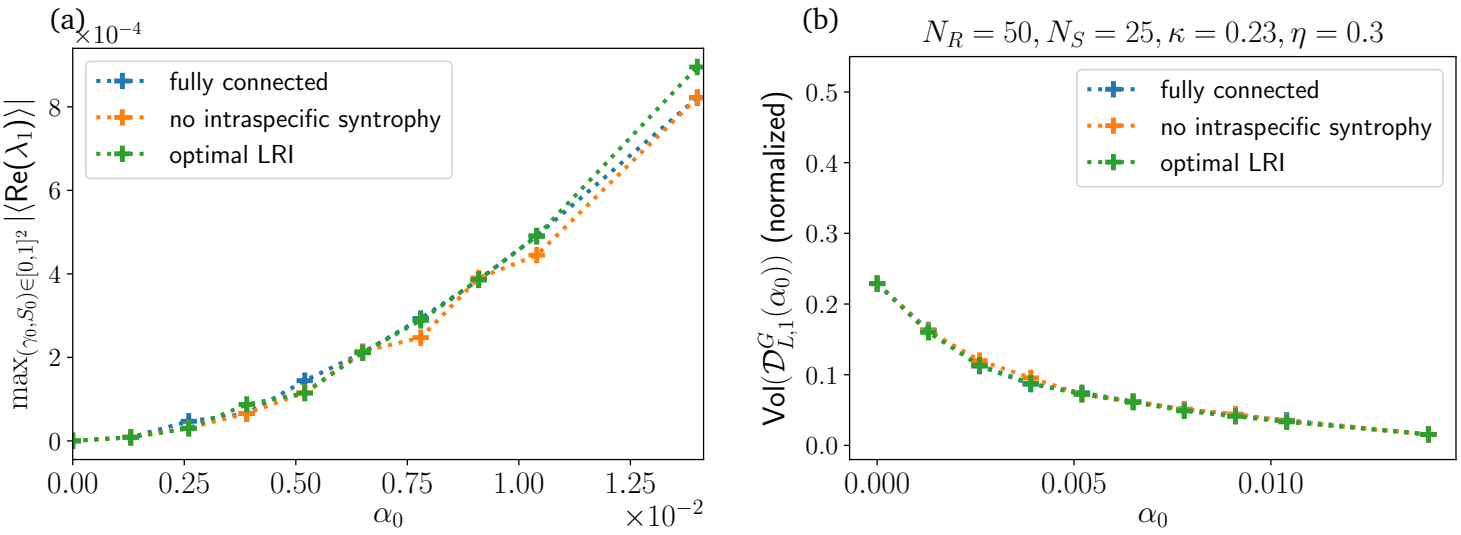

Figure 3.23: To be compared with Fig.3.21. The consumption matrix G considered here has $\eta_G = 0.3$ and $\kappa_G = 0.23$. (a) Maximal average $|\mathrm{Re}\,(\lambda_1)|$ observed in the unit square. (b) Percentage of the unit square occupied by the fully dynamically stable region of G as a function of syntrophy. The matrix considered has almost equal properties to the one in Fig.3.21, with the only difference that $N_R = 50$ here. Even though the size of $\mathcal{D}^{G,A}_{L,1}$ is smaller, the eigenvalues are larger in magnitude.

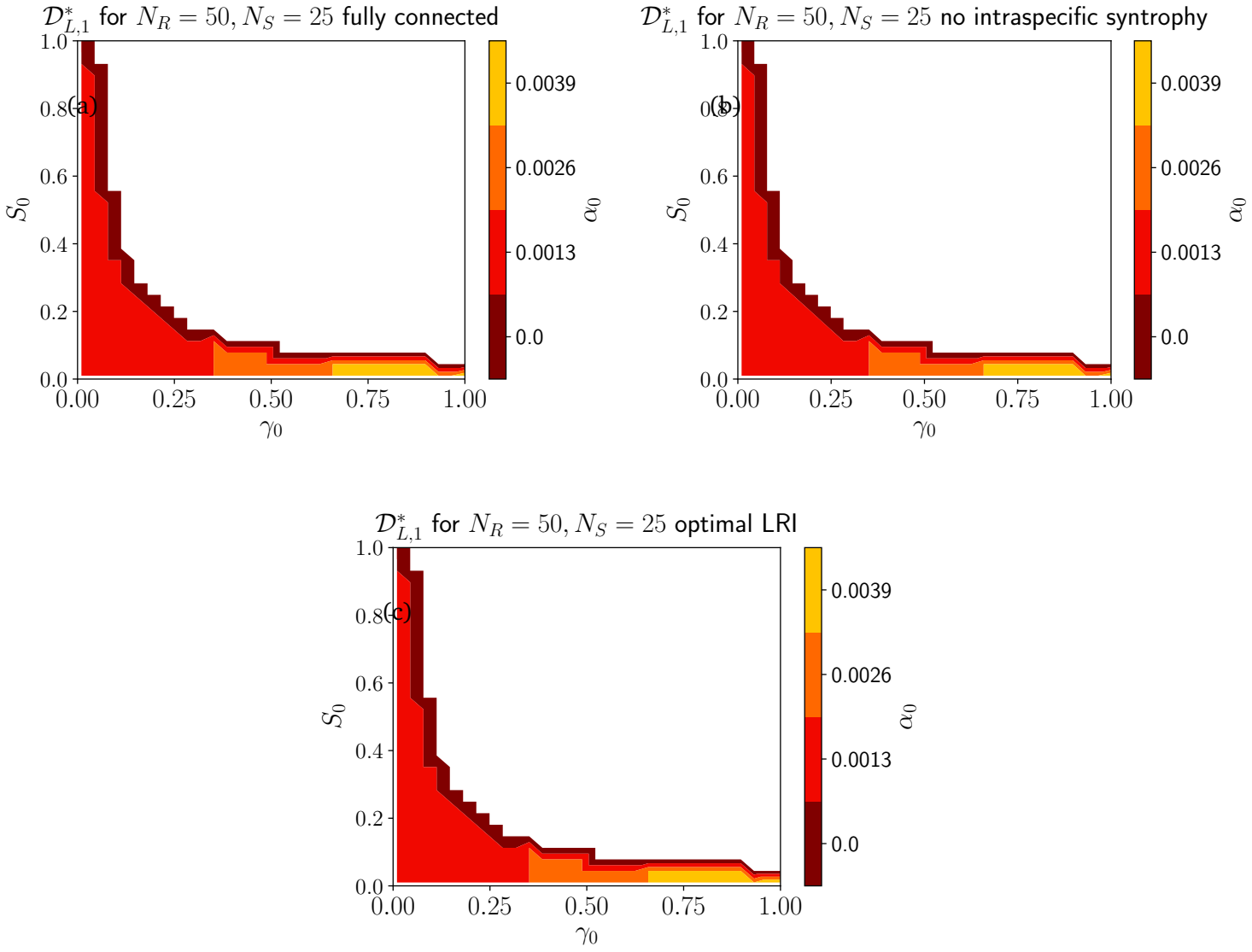

Figure 3.24: Common fully dynamically stable volume. It is larger with a larger number of resources -> even though individually it is not always better, it is better for the worse matrices (since the common volume can handle more syntrophy)

3.3 Structural stability

3.3.1 Estimating the critical structural perturbation

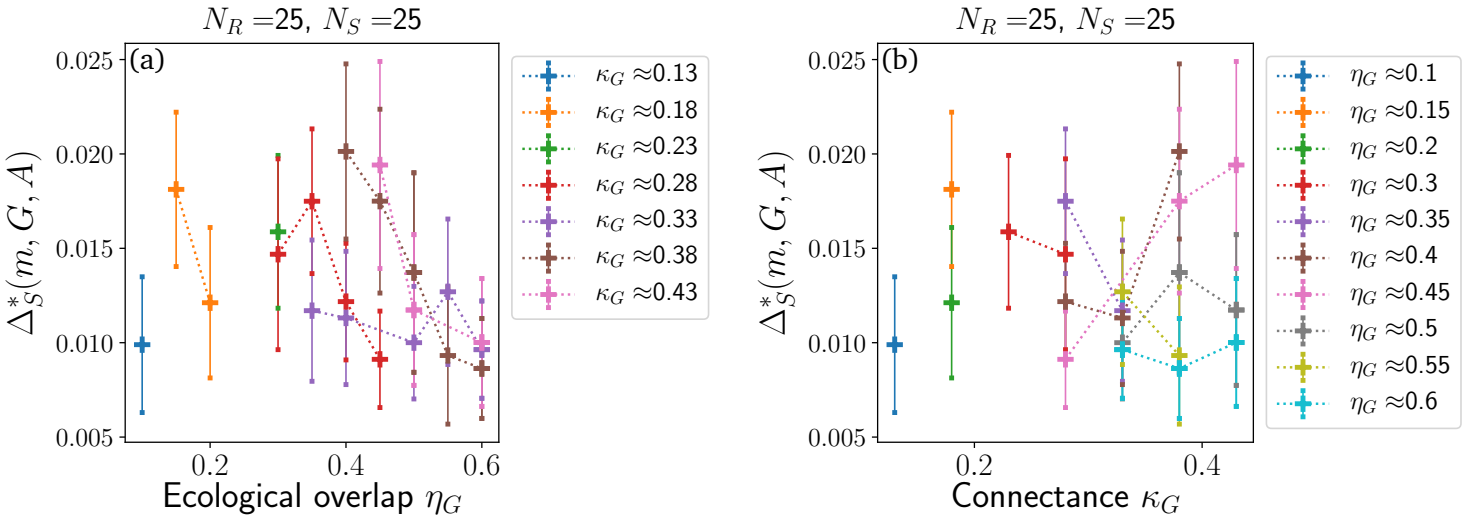

Figure 3.25: Critical structural perturbation $\Delta_S^*(m, G, A)$ (a) as a function of ecological overlap with fixed connectance and (b) as a function of connectance for a fixed ecological overlap. The critical perturbations are not very high, of the order of one percent. A clear trend may be observed, which is coherent with what was seen in Figure 3.18: for a given connectance, communities with a large ecological overlap are structurally less stable. Finding a trend for how Δ_S^* varies as for a given ecological overlap as the connectance varies is harder because of the large errors.

Chapter 4

Appendices

4.1 Demonstrations

4.1.1 Special determinant computation

We want to knwo when the determinant of the following N-dimensional square matrix is zero:

$$A_N = \begin{pmatrix} a & b & b \\ b & \ddots & b \\ b & b & a \end{pmatrix}, \textit{i.e. } A_{ij} = b + (a-b)\delta_{ij}. \tag{4.1}$$

The equation we want to solve is:

$$\det\left(A_N\right) = 0. \tag{4.2}$$

Note that, using Gaussian elimination, Eq.(4.2) can be transformed in:

$$\det \begin{pmatrix} a & b & \dots & b \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix} = 0 \tag{4.3}$$

Using Laplace's expansion, this can be written as:

$$a \det \begin{pmatrix} a-b & 0 & \dots & 0 \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix} + (a-b) \det \begin{pmatrix} b & b & \dots & b \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix} = 0 \tag{4.4}$$

Since the first term of the previous equation is a lower triangular matrix, its determinant is easily found:

$$a \det \begin{pmatrix} a-b & 0 & \dots & 0 \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix} = a\left(a-b\right)^{n-1}. \tag{4.5}$$

Finding an explicit equation for the left term is a bit more involving. Let us define the general n square matrix $F_n(a,b)$:

$$F_n(a,b) = \begin{pmatrix} b & b & \dots & b \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix}. \tag{4.6}$$

With a Laplace expansion one gets:

$$\det\left(F_n(a,b)\right) = b\det\begin{pmatrix} a-b & 0 & 0 \\ b-a & \ddots & 0 \\ 0 & b-a & a-b \end{pmatrix} + (a-b)\det\begin{pmatrix} b & b & \dots & b \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix}. \tag{4.7}$$

This means:

$$\det(F_n(a,b)) = b(a-b)^{n-1} + (a-b)\det(F_{n-1}(a,b)). \tag{4.8}$$

It is easy to check that the solution to the previous equation is:

$$\det(F_n(a,b)) = \left[(n-1)b + \det\left(F_1(a,b)\right)\right](a-b)^{n-1}. \tag{4.9}$$

Since $\det\left(F_1(a,b)\right) = 1$, we get:

$$\det\left(F_n(a,b)\right) = n(a-b)^{n-1}b \tag{4.10}$$

Inserting this in Eq.(4.4) yields:

$$\boxed{\det(A_N) = 0 \iff (a-b)^{N-1}\left[a + (N-1)b\right] = 0.} \tag{4.11}$$

4.1.2 The optimal S_0 for locally dynamically stable systems

How S_0 should be adjusted is a bit tricky because it is present in two terms that do not have the same behaviour: l_0^2/S_0 and $N_S^2\alpha_0^2 S_0$. So we need to compute the minimum value the sum of these two terms is and take S_0 as the minimum we found. The consumers equilibrium abundance S_0^* that yields the minimum value is given by

$$\frac{d}{dS_0}\left[\frac{l_0^2}{S_0} + N_S^2\alpha_0^2 S_0\right]_{S_0=S_0^*} = 0 \iff S_0^* = \frac{l_0}{N_S\alpha_0}. \tag{4.12}$$

One checks really easily that this point is indeed a minimum. So if $S_0 > S_0^*$ it should be decreased, and otherwise increased.

4.2 Supplementary material

4.2.1 Effective system

Models which involve the dynamics of species only are in general better known than consumers-resources models [**insert reference**]. In particular, a huge body of literature exists on the study of Lotka-Volterra systems [**insert reference**]. We may profit from this knowledge by transforming the effect of the resources dynamics into an effective consumers-only system.

This can be done by assuming that the resources reach an equilibrium way faster than the consumers. Mathematically, that is equivalent to

$$\frac{dR_\mu}{dt} \approx 0, \forall \mu. \tag{4.13}$$

Using Eq.(1.1a), we get an explicit value for the resources:

$$R_\mu \approx \frac{l_\mu + \sum_j \alpha_{\mu j} S_j}{m_\mu + \sum_k \gamma_{k\mu} S_k}. \tag{4.14}$$

This expression can be used in Eq.(1.1b) to get an effective system which describes the dynamics of the N_S consumers:

$$\frac{dS_i}{dt} = \left(\sum_\nu \left(\frac{\sigma_{i\nu}\gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S_k} - \alpha_{\nu i} \right) - d_i + \sum_{\nu j} \frac{\sigma_{i\nu}\gamma_{i\nu}\alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S_k} S_j \right) S_i. \tag{4.15}$$

This can be rewritten in a more compact way:

$$\frac{dS_i}{dt} = p_i(S) S_i + \sum_j M_{ij}(S) S_i S_j \tag{4.16}$$

with

$$p_i(S) = -\left(d_i + \sum_\nu \alpha_{\nu i} \right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S_k} \text{ and } M_{ij}(S) = \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}\alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S_k}. \tag{4.17}$$

If we assume the species S_k are not too far away from their equilibrium values[24], *i.e.*

$$S_k \approx S_k^* \ \forall k, \tag{4.18}$$

then using Eq.(**??**) we can simplify p_i. Indeed,

$$m_\nu + \sum_k \gamma_{k\nu} S_k \approx m_\nu + \sum_k \gamma_{k\nu} S_k^* = \frac{l_\nu + \sum_k \alpha_{\nu k} S_k^*}{R_\nu^*} \tag{4.19}$$

[24]Note that this is very rarely true, especially in the context of the study of structural stability, where entire species sometimes die out.

Hence, the explicit dynamical dependence on S can be removed from p_i and M_{ij}:

$$p_i(S) \approx p_i \equiv -\left(d_i + \sum_\nu \alpha_{\nu i}\right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu} l_\nu R^*_\nu}{l_\nu + \sum_k \alpha_{\nu k} S^*_k}, \tag{4.20}$$

and

$$M_{ij}(S) \approx M_{ij} \equiv \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu} R^*_\nu \alpha_{\nu j}}{l_\nu + \sum_k \alpha_{\nu k} S^*_k}. \tag{4.21}$$

Perturbation analysis

We study a system that we put close to an equilibrium S^*, *i.e.*

$$S = S^* + \Delta S, \text{with } \Delta S \ll 1. \tag{4.22}$$

Written this way, the effective equations of motion Eq.(4.16) are equivalent to:

$$\frac{d\Delta S_i}{dt} = p_i(S^* + \Delta S)\left(S^*_i + \Delta S_i\right) + \sum_j M_{ij}(S^* + \Delta S)\left(S^*_i + \Delta S_i\right)\left(S^*_j + \Delta S_j\right). \tag{4.23}$$

Since the deviations from equilibrium $\Delta S_i \ll 1$, we can forget the terms in higher power than quadratic:

$$\frac{d\Delta S_i}{dt} = \tilde{p}_i \Delta S_i + \sum_j E_{ij} \Delta S_j + \mathcal{O}(\Delta S^2), \tag{4.24}$$

with

$$\tilde{p}_i \equiv p_i(S^*) + \sum_k M_{ik}(S^*) S^*_k, \tag{4.25}$$

and

$$E_{ij} \equiv \left(\left.\frac{\partial p_i}{\partial S_j}\right|_{S^*} + M_{ij}(S^*) + \sum_k \left.\frac{\partial M_{ik}}{\partial S_j}\right|_{S^*} S^*_k\right) S^*_i. \tag{4.26}$$

After some computations, we can get $\tilde{p}_i$ and E_{ij} in terms of the initial parameters. Indeed,

$$p_i(S^*) = -\left(d_i + \sum_\nu \alpha_{\nu i}\right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S^*_k} \tag{4.27}$$

and

$$M_{ik}(S^*) = \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}\alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S^*_k}. \tag{4.28}$$

Hence, using Eq.(4.25):

$$\tilde{p}_i = -\left(d_i + \sum_\nu \alpha_{\nu i}\right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}}{m_\nu + \sum_k \gamma_{k\nu} S^*_k}\left(l_\nu + \sum_j \alpha_{\nu j} S^*_j\right). \tag{4.29}$$

This can be simplified using Eq.(4.19) and Eq.(1.2b):

$$\tilde{p}_i = -d_i + \sum_\nu \sigma_{i\nu}\gamma_{i\nu} R^*_\nu = \sum_\nu \alpha_{\nu i}. \tag{4.30}$$

With a similar computation, one finds

$$E_{ij} = \sum_{\nu} \frac{\sigma_{i\nu}\gamma_{i\nu}S_i^*}{m_\nu + \sum_k \gamma_{k\nu}S_k^*} \left(\alpha_{\nu j} - \gamma_{j\nu}R_\nu^*\right). \tag{4.31}$$

Finally, Eq.(4.24) can be recast in

$$\frac{d\Delta S_i}{dt} = \sum_j (J_E)_{ij}\Delta S_j, \tag{4.32}$$

where the effective $N_S \times N_S$ jacobian matrix J_E is defined by:

$$(J_E)_{ij} = \sum_{\nu} \left[\frac{\sigma_{i\nu}\gamma_{i\nu}S_i^*}{m_\nu + \sum_k \gamma_{k\nu}S_k^*} \left(\alpha_{\nu j} - \gamma_{j\nu}R_\nu^*\right) + \alpha_{\nu i}\delta_{ij} \right]. \tag{4.33}$$

We see that we without surprise we find again the B, Γ and Δ matrices coming from the jacobian at equilibrium:

$$(J_E)_{ij} = \sum_{\nu} \left[\frac{\mathrm{B}_{i\nu}\Gamma_{\nu j}}{\Delta_\nu} + \alpha_{\nu i}\delta_{ij} \right] \tag{4.34}$$

This matrix determines the stability of the equilibrium. Namely if the largest eigenvalue of J_E is positive, the equilibrium is unstable. If it is negative, the equilibrium is stable. If it is zero, the equilibrium is marginal.

Bibliography

[1] Ashutosh Jogalekar. *Physicists in Biology; And Other Quirks of the Genomic Age*. Scientific American Blog Network. Library Catalog: blogs.scientificamerican.com. URL: https://blogs.scientificamerican.com/the-curious-wavefunction/physicists-in-biology-and-other-quirks-of-the-genomic-age/ (visited on 04/02/2020).

[2] Erwin Schrödinger. *What is Life? The Physical Aspect of the Living Cell*. Cambridge University Press, 1944.

[3] Hue Sun Chan and Ken A Dill. "The Protein Folding Problem". In: (1993), p. 10.

[4] C. K. Fisher and P. Mehta. "The transition between the niche and neutral regimes in ecology". In: *Proceedings of the National Academy of Sciences* 111.36 (Sept. 9, 2014), pp. 13111–13116. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1405637111.

[5] Alessandro Attanasi et al. "Information transfer and behavioural inertia in starling flocks". In: *Nature Physics* 10.9 (Sept. 2014), pp. 691–696. ISSN: 1745-2473, 1745-2481. DOI: 10.1038/nphys3035.

[6] Ruth E. Ley et al. "Worlds within worlds: evolution of the vertebrate gut microbiota". In: *Nature Reviews Microbiology* 6.10 (Oct. 2008), pp. 776–788. ISSN: 1740-1526, 1740-1534. DOI: 10.1038/nrmicro1978.

[7] Cristina Becerra-Castro et al. "Wastewater reuse in irrigation: A microbiological perspective on implications in soil fertility and human and environmental health". In: *Environment International* 75 (Feb. 2015), pp. 117–135. ISSN: 01604120. DOI: 10.1016/j.envint.2014.11.001.

[8] P. G. Falkowski. "Biogeochemical Controls and Feedbacks on Ocean Primary Production". In: *Science* 281.5374 (July 10, 1998), pp. 200–206. DOI: 10.1126/science.281.5374.200.

[9] A J Lotka. "Analytical Note on Certain Rhythmic Relations in Organic Systems". In: *Proceedings of the National Academy of Sciences of the United States of America* 6.7 (July 1920), pp. 410–415. ISSN: 0027-8424. DOI: 10.1073/pnas.6.7.410.

[10] Babak Momeni, Li Xie, and Wenying Shou. "Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions". In: *eLife* 6 (Mar. 28, 2017), e25051. ISSN: 2050-084X. DOI: 10.7554/eLife.25051.

[11] James P. O'Dwyer. "Whence Lotka-Volterra?: Conservation laws and integrable systems in ecology". In: *Theoretical Ecology* 11.4 (Dec. 2018), pp. 441–452. ISSN: 1874-1738, 1874-1746. DOI: 10.1007/s12080-018-0377-0.

[12] Robert MacArthur. “Species packing and competitive equilibrium for many species”. In: *Theoretical Population Biology* 1.1 (May 1970), pp. 1–11. ISSN: 00405809. DOI: 10.1016/0040-5809(70)90039-0.

[13] James D. Brunner and Nicholas Chia. “Metabolite mediated modeling of microbial community dynamics captures emergent behavior more effectively than species-species modeling”. In: *Journal of The Royal Society Interface* 16.159 (Oct. 31, 2019), p. 20190423. ISSN: 1742-5689, 1742-5662. DOI: 10.1098/rsif.2019.0423.

[14] Brandon E.L. Morris et al. “Microbial syntrophy: interaction for the common good”. In: *FEMS Microbiology Reviews* 37.3 (May 2013), pp. 384–406. ISSN: 1574-6976. DOI: 10.1111/1574-6976.12019.

[15] Jeffrey D Orth, Ines Thiele, and Bernhard Ø Palsson. “What is flux balance analysis?” In: *Nature Biotechnology* 28.3 (Mar. 2010), pp. 245–248. ISSN: 1087-0156, 1546-1696. DOI: 10.1038/nbt.1614.

[16] Ines Thiele et al. “Multiscale Modeling of Metabolism and Macromolecular Synthesis in E. coli and Its Application to the Evolution of Codon Usage”. In: *PLoS ONE* 7.9 (Sept. 28, 2012). Ed. by Tamir Tuller, e45635. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0045635.

[17] T W James. “Continuous Culture of Microorganisms”. In: *Annual Review of Microbiology* 15.1 (Oct. 1961), pp. 27–46. ISSN: 0066-4227, 1545-3251. DOI: 10.1146/annurev.mi.15.100161.000331.

[18] Alberto Pascual-García and Ugo Bastolla. “Mutualism supports biodiversity when the direct competition is weak”. In: *Nature Communications* 8.1 (Apr. 2017), p. 14326. ISSN: 2041-1723. DOI: 10.1038/ncomms14326.

[19] Stacey Butler and James P. O’Dwyer. “Stability criteria for complex microbial communities”. In: *Nature Communications* 9.1 (Dec. 2018), p. 2970. ISSN: 2041-1723. DOI: 10.1038/s41467-018-05308-z.

[20] Mimmo Iannelli and Andrea Pugliese. *An Introduction to Mathematical Population Dynamics*. Vol. 79. UNITEXT. Cham: Springer International Publishing, 2014. ISBN: 978-3-319-03025-8 978-3-319-03026-5. DOI: 10.1007/978-3-319-03026-5.

[21] Ugo Bastolla et al. “The architecture of mutualistic networks minimizes competition and increases biodiversity”. In: *Nature* 458.7241 (Apr. 2009), pp. 1018–1020. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature07950.

[22] Samuel Jonhson, Virginia Domínguez-García, and Miguel A. Muñoz. “Factors Determining Nestedness in Complex Networks”. In: *PLoS ONE* 8.9 (Sept. 19, 2013). Ed. by Yamir Moreno, e74025. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0074025.

[23] ROBERT M. MAY. “Will a Large Complex System be Stable?” In: *Nature* 238.5364 (Aug. 1, 1972), pp. 413–414. ISSN: 1476-4687. DOI: 10.1038/238413a0.

[24] Stefano Allesina and Si Tang. “Stability Criteria for Complex Ecosystems”. In: *Nature* 483.7388 (Mar. 2012), pp. 205–208. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature10832.

[25] Stefano Allesina et al. “Predicting the stability of large structured food webs”. In: *Nature Communications* 6.1 (Nov. 2015), p. 7842. ISSN: 2041-1723. DOI: 10.1038/ncomms8842.

[26] Matthieu Barbier and Jean-François Arnoldi. *The cavity method for community ecology*. preprint. Ecology, June 8, 2017. DOI: 10.1101/147728.

[27] A. R. Ives and S. R. Carpenter. “Stability and Diversity of Ecosystems”. In: *Science* 317.5834 (July 6, 2007), pp. 58–62. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1133258.

[28] Giulio Biroli, Guy Bunin, and Chiara Cammarota. “Marginally stable equilibria in critical ecosystems”. In: *New Journal of Physics* 20.8 (Aug. 31, 2018), p. 083051. ISSN: 1367-2630. DOI: 10.1088/1367-2630/aada58.

[29] OSCAR Perron. “Zur Theorie der Matrices”. In: (), p. 16.

[30] Philip D. Powell. “Calculating Determinants of Block Matrices”. In: *arXiv:1112.4379 [math]* (Dec. 15, 2011).

[31] Sergej Gerschgorin. “"Uber die Abgrenzung der Eigenwerte einer Matrix”. In: *Bulletin de l’Académie des Sciences de l’URSS. Classe des sciences mathématiques et na* 6 (1931), pp. 749–754.

[32] Daniel Delahaye, Supatcha Chaimatanan, and Marcel Mongeau. “Simulated Annealing: From Basics to Applications”. In: *Handbook of Metaheuristics*. Ed. by Michel Gendreau and Jean-Yves Potvin. Vol. 272. Series Title: International Series in Operations Research & Management Science. Cham: Springer International Publishing, 2019, pp. 1–35. ISBN: 978-3-319-91085-7 978-3-319-91086-4. DOI: 10.1007/978-3-319-91086-4_1.

[33] Jan R. Magnus and Heinz Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. Third edition. Wiley series in probability and statistics. Hoboken, NJ: Wiley, 2019. 1 p. ISBN: 978-1-119-54119-6 978-1-119-54116-5.

[34] G. Hardin. “The Competitive Exclusion Principle”. In: *Science* 131.3409 (Apr. 29, 1960), pp. 1292–1297. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.131.3409.1292.

[35] Xin Wang and Yang-Yu Liu. “Overcome Competitive Exclusion in Ecosystems”. In: *arXiv:1805.06002 [physics, q-bio]* (May 15, 2019).