Consumer resources model

Léo Buchenel

March 27, 2020

Contents

1 Introduction **3**
1.1 Why Consumers Resources Models in microbial ecology? 3
1.2 Establishing the model and the goals 3
1.2.1 Equilibria of the model 4
1.2.2 Attack strategy and important notions 4

2 Methods **8**
2.1 Feasibility 8
2.1.1 Basic concepts 8
2.1.2 The feasibility volume 10
2.1.3 Estimating the fully feasible volume $\mathcal{V}_1^{G,A}$ 11
2.1.4 Algorithmic procedure 13
2.2 Dynamical stability 15
2.2.1 Definitions 15
2.2.2 The quest for a full solution 18
2.2.3 Bounds on the eigenvalues 20
2.2.4 Low intra resources interaction (LRI) regime 22
2.2.5 Effective system 26
2.2.6 Flux analysis - a way to get a sense of scales 29
2.2.7 Application : fully connected consumption and syntrophy networks . 29
2.3 Structural stability 34

3 Results **35**
3.1 Feasibility 35
3.1.1 The feasibility volume $\mathcal{V}_x^{G,A}$ 35
3.1.2 Evolution of the feasible volume with syntrophy 37
3.2 Dynamical stability 43
3.2.1 LRI regime – Outcome of the Monte Carlo algorithm 43
3.2.2 Fully dynamically stable volume 43
3.2.3 Largest eigenvalue of the jacobian 43
3.3 Structural stability 48

4 Appendices **49**
4.1 Demonstrations 49
4.1.1 Special determinant computation 49

4.1.2 The optimal S_0 for locally dynamically stable systems 50

Chapter 1

Introduction

1.1 Why Consumers Resources Models in microbial ecology?

1.2 Establishing the model and the goals

We want to write down a consumers-resources model (CRM) which describes the coupled evolution between N_S different species of biomass (denoted S_i with $i = 1, \ldots, N_S$) and the N_R resources they feed off (denoted R_μ with $\mu = 1, \ldots, N_R$).

The coupled evolution of the set of variables $\{R_\mu, S_i\}$ is given by:

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

We will call the variables $\{l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i\}$ *parameters of the model* **add R^* and** S^*. On the other hand, because they may dynamically evolve, $\{R_\mu, S_i\}$ will be refered to as *dynamical variables* or simply *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 abundance R_μ). Quantities related to species have subscripts in latin alphabet (*e.g.* the species i 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 death rate m_μ).
- Matrices (*i.e.* quantities with two indices, usually relating resources and species) are written with the greek alphabet (*e.g.* the rate $\gamma_{i\mu}$ at which species i consumes resource μ).

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

- Constant input from an outsider experimenter : this corresponds to the constant $+l_\mu$ term,
- Natural diffusion/deterioration at rate m_μ : this corresponds to the $-m_\mu R_\mu$ term,
- Consumption by the biomass species S_j at a rate $\gamma_{j\mu}$. In total this corresponds to the Lotka-Volterra style term [**insert ref**] $-\sum_j \gamma_{j\mu} R_\nu S_j$,
- Production coming from the species S_j at a rate $\alpha_{\mu j}$: $+\sum_j \alpha_{\mu j} S_j$. This is essentially what makes this model different from traditional CRMs [**insert ref**].

On the other hand, biomass of species S_i changes because of the following processes [**insert figure**]:

- 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.
- Release of resource R_ν at rate $\alpha_{\nu i}$ (this is the syntrophic interaction). In total $-\sum_\nu \alpha_{\nu i} S_i$.

We will mostly focus on the case where no resource coming from syntrophy is lost, *i.e.* $\tau_{\mu i} = \alpha_{\mu i}$.

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 are interested in studying the stability of the equilibrium points of our model Eqs.(1.1). We say that $\{R^*_\mu, S^*_j\}$ is an *equilibrium*[1] of our model if they are fixed points of it, that means if the following equations are fulfilled :

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

As said above, our main goal is to study the stability of such equilibria. Before introducing different notions around equilibria, we focus on simplifying the problem first as much as we can. Note that we consider R^*_μ and S^*_i as parameters of the model.

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 **??** is available at the address `https://gitlab.ethz.ch/palberto/consumersresources.git`.

[1] For the sake of brevity, we will sometimes drop the μ and j subscripts when we write $\{R^*_\mu, S^*_j\}$.

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 $\forall i = 1, \ldots, N_S; \mu = 1, \ldots, N_R$. 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 $\mathcal{P} \simeq \mathbb{R}_+^{3N_R+2N_S+4N_R N_S}$. The difficulty of task becomes clear : $\mathcal{P}$ is a huge, complicated space so we need to find a way to simplify it.

Indeed, the goal is to study microbial communities with a large number of consumers and resources, typically $N_R, N_S \approx 25, 50, 100, \ldots$ *i.e.* $\sim 2'500$ or more free parameters and so it is clear that a precise study on each element of each matrix is way too tenuous of a job. It's easier to look at it from a statistical point of view, *i.e.* we will write a matrix $q_{i\mu}$ as [1] :

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

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 can write :

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

Note that we did not put 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 for S^* and R^*, we require positive-valued equilibria [2]).

In order to make computations analytically tractable, we will require a small standard deviation on the parameters involved in the problem. In that regime, every random variable $\mathcal{Q}$ is well approximated by its average value Q_0, that we call a *metaparameter*. While studying

things analytically we will hence often come back to the following approximation:

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

This 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, what we did was simply designing 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 will mostly work in the collapsed space $\mathcal{C}(\mathcal{P})$ because it reduces the number of parameters from $3N_R + 2N_S + 3N_R N_S$ (continous) to 8 (continous) + $3N_R N_S$ (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**: this measures how "nested" the system is, *i.e.* if there are clusters grouped together[2]. It is known [4, 1] that nestedness can play a profound role in the dynamics of ecological communities. Although it is somewhat controversed [5], we will keep the definition of the nestedness $\eta(M)$ of a binary matrix M as it was used in [4]:

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

  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

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

[2]In typical Lotka-Volterra models, where only species-species interactions are considered, *e.g.* [3], measuring the nestedness of the γ consumption matrix would be in the same spirit as counting how many niches there are in the community.

- Its **nestedness**[3] : this measure, simply defined as the ratio of non-zero links in a matrix, is central in the study of plants-and-animals systems [1]. 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 will make us lose information. Losing information – and hence complexity – is convenient when doing analytical computations but it is not when we want to produce precise 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 $\mathcal{A} : \mathcal{M} \times \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S} \to \mathcal{P}$ which brings us from the collapsed space to the parameter space[4]. Numerically what we will do is : from a set of metaparameters $m \in \mathcal{M}$ and binary matrices $B = (G, A) \in \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$, 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.

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

[4]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.(1.2) are fulfilled.

2.1.1 Basic concepts

Since its very inception [6], the study of ecological interactions has been and still is tightly close to the one of random matrices [7, 8, 9]. 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[5]. 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:

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

[5]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*[6]. 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) and calculating the number of feasible ones :

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

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

2.1.2 The feasibility volume

The algorithmic procedure detailed in Section 2.1.4 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 [9], we notice that our system (1.1) 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[7], 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 $\mathrm{Vol}\left(\mathcal{V}_x^{G,A}\right) \leqslant \mathrm{Vol}\left(\mathcal{V}_y^{G,A}\right)\ \forall x > y, G$. 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:

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

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 :

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

[7]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.

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 :

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

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[8] 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[9], we will take a zero-variance à la Butler and O'Dwyer in [2], *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:

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

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

6. With all of these parameters drawn, we can solve Eq.(1.2b) for the species death rate d_i.

7. We solve Eq.(1.2a) 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 2.1, 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 :

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

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

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

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 [1] 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:

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

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 [10, 1]. 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* [11].

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 [12] cannot be applied. The eigenvalues of J^* are obtained through the eigenvalue problem:

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

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[10] 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[11]:

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

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[12], *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 [13] :

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

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 dynamics is essentially dictated by the $\Gamma\mathrm{B}$ N_R-dimensional square matrix, which is given by :

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

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

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

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

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 [14] 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 :*

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

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

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:

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

and the "consumers" discs :

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

The circle's theorem Eq.(2.64) tells us that all eigenvalues will be in the union of these circles, *i.e.* there exists $\forall \lambda \in \sigma\,(J^*)$ 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 $\mathrm{Im}\,(\Delta_\mu) = 0$, 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 :

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

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 :

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

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* $\mathrm{Re}\,(\lambda_1) < 0$, *i.e.* local dynamical stability is guaranteed. We inspire ourselves from the general idea of the mathematical proofs of [2]. We first rewrite Eq.(2.61) as[13] :

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

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 $\mathrm{Re}\,(\lambda) > 0$. 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 $\mathrm{Re}\,(\lambda) \leqslant 0$. As this is a contradiction to the hypothesis that the regime is unstable, we must conclude that the regime is stable[14].

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.

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:

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

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

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

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[15] 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

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

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[16] 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. □

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

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

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), (2.22) and (2.25).
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:

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

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

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

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 :

$$O_{\mu\nu} \equiv \sum_k A_{\mu k} G_{k\nu}. \quad (2.95)$$

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.
Similarly, the consumption overlap matrix is defined as :

$$C_{\mu\nu} = \sum_k G_{k\mu} G_{k\nu}. \quad (2.96)$$

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.(**??**) :

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

Combining this with the approximation of $\Gamma\mathrm{B}$ above we get an approximative LRI regime condition on the metaparameters :

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

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.(2.25) 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 [15] : 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:

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

  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):

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

  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.(2.28) : $\alpha_0 = \min(1 - \sigma_0, \sigma_0)\gamma_0 R_0 N_R$. 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 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

$$\frac{dR_\mu}{dt} \approx 0, \forall \mu. \tag{2.103}$$

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

$$R_\mu \approx \frac{l_\mu + \sum_j \alpha_{\mu j} S_j}{m_\mu + \sum_k \gamma_{k\mu} S_k}. \tag{2.104}$$

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

$$\frac{dS_i}{dt} = \left(\sum_\nu \left(\frac{\sigma_{i\nu}\gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S_k} - \alpha_{\nu i} \right) - d_i + \sum_{\nu j} \frac{\sigma_{i\nu}\gamma_{i\nu}\alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S_k} S_j \right) S_i. \tag{2.105}$$

This can be rewritten in a more compact way:

$$\frac{dS_i}{dt} = p_i(S) S_i + \sum_j M_{ij}(S) S_i S_j \tag{2.106}$$

with

$$p_i(S) = -\left(d_i + \sum_\nu \alpha_{\nu i} \right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S_k} \text{ and } M_{ij}(S) = \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}\alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S_k}. \tag{2.107}$$

If we assume the species S_k are not too far away from their equilibrium values[17], *i.e.*

$$S_k \approx S_k^* \ \forall k, \tag{2.108}$$

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

$$m_\nu + \sum_k \gamma_{k\nu} S_k \approx m_\nu + \sum_k \gamma_{k\nu} S_k^* = \frac{l_\nu + \sum_k \alpha_{\nu k} S_k^*}{R_\nu^*} \tag{2.109}$$

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

$$p_i(S) \approx p_i \equiv -\left(d_i + \sum_\nu \alpha_{\nu i} \right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu} l_\nu R_\nu^*}{l_\nu + \sum_k \alpha_{\nu k} S_k^*}, \tag{2.110}$$

and

$$M_{ij}(S) \approx M_{ij} \equiv \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu} R_\nu^* \alpha_{\nu j}}{l_\nu + \sum_k \alpha_{\nu k} S_k^*}. \tag{2.111}$$

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

Perturbation analysis

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

$$S = S^* + \Delta S, \text{with } \Delta S \ll 1. \tag{2.112}$$

Written this way, the effective equations of motion Eq.(2.106) are equivalent to:

$$\frac{d\Delta S_i}{dt} = p_i(S^* + \Delta S)\left(S_i^* + \Delta S_i\right) + \sum_j M_{ij}(S^* + \Delta S)\left(S_i^* + \Delta S_i\right)\left(S_j^* + \Delta S_j\right). \tag{2.113}$$

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

$$\frac{d\Delta S_i}{dt} = \tilde{p}_i \Delta S_i + \sum_j E_{ij}\Delta S_j + \mathcal{O}(\Delta S^2), \tag{2.114}$$

with

$$\tilde{p}_i \equiv p_i(S^*) + \sum_k M_{ik}(S^*)S_k^*, \tag{2.115}$$

and

$$E_{ij} \equiv \left(\left.\frac{\partial p_i}{\partial S_j}\right|_{S^*} + M_{ij}(S^*) + \sum_k \left.\frac{\partial M_{ik}}{\partial S_j}\right|_{S^*} S_k^*\right) S_i^*. \tag{2.116}$$

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

$$p_i(S^*) = -\left(d_i + \sum_\nu \alpha_{\nu i}\right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}l_\nu}{m_\nu + \sum_k \gamma_{k\nu}S_k^*} \tag{2.117}$$

and

$$M_{ik}(S^*) = \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}\alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu}S_k^*}. \tag{2.118}$$

Hence, using Eq.(2.115) :

$$\tilde{p}_i = -\left(d_i + \sum_\nu \alpha_{\nu i}\right) + \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}}{m_\nu + \sum_k \gamma_{k\nu}S_k^*}\left(l_\nu + \sum_j \alpha_{\nu j}S_j^*\right). \tag{2.119}$$

This can be simplified using Eq.(2.109) and Eq.(1.2b) :

$$\tilde{p}_i = -d_i + \sum_\nu \sigma_{i\nu}\gamma_{i\nu}R_\nu^* = \sum_\nu \alpha_{\nu i}. \tag{2.120}$$

With a similar computation, one finds

$$E_{ij} = \sum_\nu \frac{\sigma_{i\nu}\gamma_{i\nu}S_i^*}{m_\nu + \sum_k \gamma_{k\nu}S_k^*}\left(\alpha_{\nu j} - \gamma_{j\nu}R_\nu^*\right). \tag{2.121}$$

Finally, Eq.(2.114) can be recast in

$$\frac{d\Delta S_i}{dt} = \sum_j (J_E)_{ij}\Delta S_j, \tag{2.122}$$

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

$$(J_E)_{ij} = \sum_\nu \left[\frac{\sigma_{i\nu}\gamma_{i\nu}S_i^*}{m_\nu + \sum_k \gamma_{k\nu}S_k^*} \left(\alpha_{\nu j} - \gamma_{j\nu}R_\nu^*\right) + \alpha_{\nu i}\delta_{ij} \right]. \tag{2.123}$$

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

$$(J_E)_{ij} = \sum_\nu \left[\frac{\mathrm{B}_{i\nu}\Gamma_{\nu j}}{\Delta_\nu} + \alpha_{\nu i}\delta_{ij} \right] \tag{2.124}$$

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.

2.2.6 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 :

$$C_{\text{tot}} = \sum_{i,\nu} \sigma_{i\nu}\gamma_{i\nu}R_\nu S_i. \tag{2.125}$$

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

$$S_{\text{tot}} = \sum_{i,\nu} \alpha_{\nu i} S_i. \tag{2.126}$$

A natural order parameter O is then

$$O \equiv \frac{S_{\text{tot}}}{C_{\text{tot}}} = \frac{\sum_{i,\nu} \alpha_{\nu i} S_i}{\sum_{i,\nu} \sigma_{i\nu}\gamma_{i\nu}R_\nu S_i} \approx \frac{N_S N_R \alpha_0 S_0}{\sigma_0 R_0 S_0 N_S N_R \gamma_0} = \frac{\alpha_0}{\sigma_0 R_0 \gamma_0}. \tag{2.127}$$

2.2.7 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.*

$$G_{i\mu} = A_{\mu i} = 1 \ \forall \mu = 1, \ldots, N_R \quad \forall i = 1, \ldots, N_S. \tag{2.128}$$

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 [9] and perform a *standard deviation expansion*. They showed that the variance of the interaction matrix drives the strength of the interactions in their model [9], 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 [9] is to write the elements of the q matrix in terms of new variables $\tilde{q}_{i\mu}$:

$$q_{i\mu} = \langle q \rangle + \sigma_q \tilde{q}_{i\mu}. \tag{2.129}$$

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

$$\langle q \rangle \equiv \frac{\sum_{\mu,i} q_{i\mu}}{N_S N_R}, \tag{2.130}$$

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

$$\sigma_q \equiv \sqrt{\langle q^2 \rangle - \langle q \rangle^2} \text{ with } \langle q^2 \rangle \equiv \frac{\sum_{\mu,i} q_{i\mu}^2}{N_S N_R}. \tag{2.131}$$

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 [9]:

$$\langle \tilde{q} \rangle = 0 \text{ and } \langle \tilde{q}^2 \rangle = 1. \tag{2.132}$$

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 $N_R, N_S \gg 1$:

$$\langle q \rangle \approx \langle \mathfrak{Q} \rangle \equiv q_0. \tag{2.133}$$

We can then rewrite the free parameters of our model[18]:

$$\begin{cases} l_\nu \approx l_0 + \sigma_l \tilde{l}_\nu & (2.134a) \\ R^*_\nu \approx R_0 + \sigma_R \tilde{r}_\nu & (2.134b) \\ S^*_i \approx S_0 + \sigma_S \tilde{s}_i & (2.134c) \\ \gamma_{i\nu} \approx \gamma_0 + \sigma_\gamma \tilde{g}_{i\nu} & (2.134d) \\ \alpha_{\nu i} \approx \alpha_0 + \sigma_\alpha \tilde{\alpha}_{\nu i} & (2.134e) \\ \sigma_{i\nu} \approx \sigma_0 + \sigma_\sigma \tilde{\sigma}_{i\nu} & (2.134f) \end{cases}$$

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 (2.40) can be written with the new variables :

$$\begin{cases} \dfrac{l_\mu + \sum_j \alpha_{\mu j} S^*_j}{R^*_\mu} = \dfrac{l_0 + N_S \alpha_0 S_0 + \sigma_l \tilde{l}_\mu + \sigma_\alpha S_0 \sum_j \tilde{\alpha}_{\mu j} + \sigma_\alpha \sigma_S \sum_j \tilde{\alpha}_{\mu j} \tilde{s}_j}{R_0 + \sigma_R \tilde{r}_\mu} & (2.135) \\ -\gamma_{j\mu} R^*_\mu + \alpha_{\mu j} = -\gamma_0 R_0 + \alpha_0 + \sigma_\gamma R_0 \tilde{\gamma}_{j\mu} + \sigma_R \gamma_0 \tilde{r}_\mu + \sigma_\alpha \tilde{\alpha}_{\mu j} + \sigma_\gamma \sigma_R \tilde{\gamma}_{j\mu} \tilde{r}_\mu & (2.136) \\ \sigma_{i\nu} \gamma_{i\nu} S^*_i = \sigma_0 \gamma_0 S_0 + \sigma_\sigma \gamma_0 S_0 \tilde{\sigma}_{i\nu} + \sigma_\gamma \sigma_0 S_0 \tilde{\gamma}_{i\nu} + \sigma_S \sigma_0 \gamma_0 \tilde{s}_i + \sigma_\sigma \sigma_\gamma S_0 \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \\ \qquad + \sigma_\sigma \sigma_S \gamma_0 \tilde{\sigma}_{i\nu} \tilde{s}_i + \sigma_\gamma \sigma_S \sigma_0 \tilde{\gamma}_{i\nu} \tilde{s}_i + \sigma_\sigma \sigma_\gamma \sigma_S \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \tilde{s}_i & (2.137) \end{cases}$$

[18]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 :

$$\sigma_P \equiv \epsilon_P \langle P \rangle,\ \forall P \in \{l_\nu, R^*_\nu, S^*_i, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \tag{2.138}$$

The previous relations then become :

$$\begin{cases} \dfrac{l_\mu + \sum_j \alpha_{\mu j} S^*_j}{R^*_\mu} = \dfrac{l_0\left(1+\epsilon_l \tilde{l}_\mu\right) + \alpha_0 S_0 \left(N_S + \epsilon_\alpha \sum_j \tilde{\alpha}_{\mu j} + \epsilon_\alpha \epsilon_S \sum_j \tilde{\alpha}_{\mu j} \tilde{s}_j\right)}{R_0\left(1+\epsilon_R \tilde{r}_\mu\right)} & (2.139) \\ -\gamma_{j\mu} R^*_\mu + \alpha_{\mu j} = -\gamma_0 R_0 \left(1 + \epsilon_\gamma \tilde{\gamma}_{j\mu} + \epsilon_R \tilde{r}_\mu + \epsilon_\gamma \epsilon_R \tilde{\gamma}_{j\mu} \tilde{r}_\mu\right) + \alpha_0 \left(1 + \epsilon_\alpha \tilde{\alpha}_{\mu j}\right) & (2.140) \\ \sigma_{i\nu} \gamma_{i\nu} S^*_i = \sigma_0 \gamma_0 S_0 \left(1 + \epsilon_\sigma \tilde{\sigma}_{i\nu} + \epsilon_\gamma \tilde{\gamma}_{i\nu} + \epsilon_S \tilde{s}_i + \epsilon_\sigma \epsilon_\gamma \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} + \epsilon_\sigma \epsilon_S \tilde{\sigma}_{i\nu} \tilde{s}_i \right. & \\ \qquad \left. + \epsilon_\gamma \epsilon_S \tilde{\gamma}_{i\nu} \tilde{s}_i + \epsilon_\sigma \epsilon_\gamma \epsilon_S \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \tilde{s}_i \right) & (2.141) \end{cases}$$

Standard deviation expansion at first order

We can then assume the relative standard deviations are small :

$$\epsilon_P \ll 1,\ \forall P \in \{l_\nu, R^*_\nu, S^*_i, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \tag{2.142}$$

The previous equations can be rewritten as :

$$\begin{cases} \dfrac{l_\mu + \sum_j \alpha_{\mu j} S^*_j}{R^*_\mu} = \dfrac{l_0 + N_S \alpha_0 S_0}{R_0} - \epsilon_R \dfrac{l_0 + N_S \alpha_0 S_0}{R_0} \tilde{r}_\mu + \epsilon_l \dfrac{l_0}{R_0} \tilde{l}_\mu + \epsilon_\alpha \dfrac{\alpha_0 S_0}{R_0} \sum_j \tilde{\alpha}_{\mu j} & (2.143) \\ -\gamma_{j\mu} R^*_\mu + \alpha_{\mu j} = -\gamma_0 R_0 + \alpha_0 - \epsilon_\gamma \gamma_0 R_0 \tilde{\gamma}_{j\mu} - \epsilon_R \gamma_0 R_0 \tilde{r}_\mu + \epsilon_\alpha \alpha_0 \tilde{\alpha}_{\mu j} & (2.144) \\ \sigma_{i\nu} \gamma_{i\nu} S^*_i = \sigma_0 \gamma_0 S_0 \left(1 + \epsilon_\sigma \tilde{\sigma}_{i\nu} + \epsilon_\gamma \tilde{\gamma}_{i\nu} + \epsilon_S \tilde{s}_i\right) & (2.145) \end{cases}$$

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 :

$$\epsilon_P \approx \epsilon \ll 1,\ \forall P \in \{l_\nu, R^*_\nu, S^*_i, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \tag{2.146}$$

This allows us to rewrite the jacobian at equilibrium as :

$$J^* = J_0 + \epsilon \tilde{J}, \tag{2.147}$$

with

$$J_0 = \begin{pmatrix} -\Delta_0 & \Gamma_0 \\ B_0 & 0 \end{pmatrix} \tag{2.148}$$

where

$$(\Delta_0)_{\mu\nu} \equiv \Delta_0 \delta_{\mu\nu} = \frac{l_0 + N_S \alpha_0 S_0}{R_0} \delta_{\mu\nu} \tag{2.149}$$

$$(\Gamma_0)_{\mu i} \equiv \Gamma_0 = -\gamma_0 R_0 + \alpha_0 \tag{2.150}$$

$$(\mathrm{B}_0)_{i\nu} \equiv \mathrm{B}_0 = \sigma_0 \gamma_0 S_0 \tag{2.151}$$

and

$$\tilde{J} = \begin{pmatrix} \tilde{\Delta} & \tilde{\Gamma} \\ \tilde{\mathrm{B}} & 0 \end{pmatrix} \tag{2.152}$$

with

$$\tilde{\Delta}_{\mu\nu} \equiv \left(-\Delta_0 \tilde{r}_\mu + \frac{l_0}{R_0}\tilde{l}_\mu + \frac{\alpha_0 S_0}{R_0}\sum_j \tilde{\alpha}_{\mu j}\right)\delta_{\mu\nu} \tag{2.153}$$

$$\tilde{\Gamma}_{\mu i} \equiv \alpha_0 \tilde{\alpha}_{\mu i} - \gamma_0 R_0 \left(\tilde{\gamma}_{i\mu} + \tilde{r}_\mu\right) \tag{2.154}$$

$$\tilde{B}_{i\mu} \equiv B_0 \left(\tilde{\sigma}_{i\mu} + \tilde{\gamma}_{i\mu} + \tilde{s}_i\right) \tag{2.155}$$

Using Jacobi's formula [16], the equation $\det\left(J^* - \lambda\right) = 0$ can be rewritten as :

$$\boxed{\det(J_0 - \lambda + \epsilon J^*) = \det(J_0 - \lambda) + \epsilon \ \mathrm{Tr}\left(\mathrm{adj}(J_0 - \lambda)\tilde{J}\right) = 0.} \tag{2.156}$$

where $\mathrm{adj}\left(\dots\right)$ 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.(2.156) becomes:

$$\det(J_0 - \lambda) = \det\begin{pmatrix} -\Delta_0 - \lambda & \Gamma_0 \\ B_0 & -\lambda \end{pmatrix} = 0 \tag{2.157}$$

If we assume that $\lambda \neq 0$, using a reasoning similar to Section 2.2.2 we can write the previous equation as:

$$\det\left(\lambda^2 + \Delta_0\lambda - \Gamma_0 \mathrm{B}_0\right) = 0 \tag{2.158}$$

Component-wise, we have :

$$\left(\lambda^2 + \Delta_0\lambda - \Gamma_0 \mathrm{B}_0\right)_{\mu\nu} = \left(\lambda^2 + \Delta_0\lambda\right)\delta_{\mu\nu} - \Gamma_0 \mathrm{B}_0. \tag{2.159}$$

Using Eq.(4.11), the non-zero solutions of Eq.(2.158) are immediately found:

$$\left(\lambda + \Delta_0\right)^{N_R - 1}\left(\lambda^2 + \Delta_0\lambda - N_R\Gamma_0 \mathrm{B}_0\right) = 0. \tag{2.160}$$

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:

$$\lambda^2 + \Delta_0\lambda - N_R\Gamma_0 \mathrm{B}_0 = 0. \tag{2.161}$$

In the end, the spectrum is given by:

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

$$\sigma(J_0) = \left\{0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2}\left(1 \pm i\sqrt{-\left(1 + \frac{4N_R\Gamma_0 \mathrm{B}_0}{\Delta_0^2}\right)}\right)\right\} \tag{2.162}$$

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

$$\sigma(J_0) = \left\{0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2}, -\frac{\Delta_0}{2}\right\} \tag{2.163}$$

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

$$\sigma(J_0) = \left\{0, \ldots, 0, -\Delta_0, \ldots, -\Delta_0, -\frac{\Delta_0}{2}\left(1 \pm \sqrt{1 + \frac{4N_R\Gamma_0\mathrm{B}_0}{\Delta_0^2}}\right)\right\} \tag{2.164}$$

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:

$$\boxed{\text{The non-variance system is dynamically unstable} \iff \Gamma_0 > 0.} \tag{2.165}$$

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 :

$$\alpha_0 \lessapprox \min(1-\sigma_0, \sigma_0)\gamma_0 R_0 \iff \Gamma_0 \lessapprox [\min(1-\sigma_0, \sigma_0) - 1]\,\gamma_0 R_0 < 0, \tag{2.166}$$

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

$$4N_R\Gamma_0\mathrm{B}_0 + \Delta_0^2 \ll 0 \iff 4N_R\sigma_0\gamma_0(\alpha_0 - \gamma_0 R_0) + \frac{l_0^2}{S_0} + 2N_S\alpha_0 l_0 + N_S^2\alpha_0^2 S_0 \ll 0. \tag{2.167}$$

2.3 Structural stability

Very similarly to dynamical stability, where the abundances of the resources and species are changed, one can perturb the parameters of a model at an equilibrium point. Namely the idea of *structural stability* is the following : one takes a given equilibrium point [**Check whether it needs to be linearly stable or not**] $\{R^*, S^*\}$ and changes some or all the parameters of the model. We will focus here on perturbing the external feeding rate of the resources[19] :

$$l_\mu \to \tilde{l}_\mu \equiv l_\mu \left(1 + \Delta_S\right). \tag{2.168}$$

The system is then time evolved under the equations of evolution (1.1), except that $l_\mu \to \tilde{l}_\mu$, until an equilibrium $\{\tilde{R}^*, \tilde{S}^*\}$ is reached. The same metrics as before can be used to quantify the structural stability.

[19]This choice may seem a bit arbitrary at first hand. It stems from the idea that we wanted to keep things simple and hence decided to change only one parameter of the model. We chose the external feeding rate of the resources because it simply is the easiest one to control in an actual experimental setup.

Chapter 3

Results

3.1 Feasibility

As explained in Methods 2.1, 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 2.1.1). 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 2.1.2). 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.(2.29). 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 2.1.2), 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.

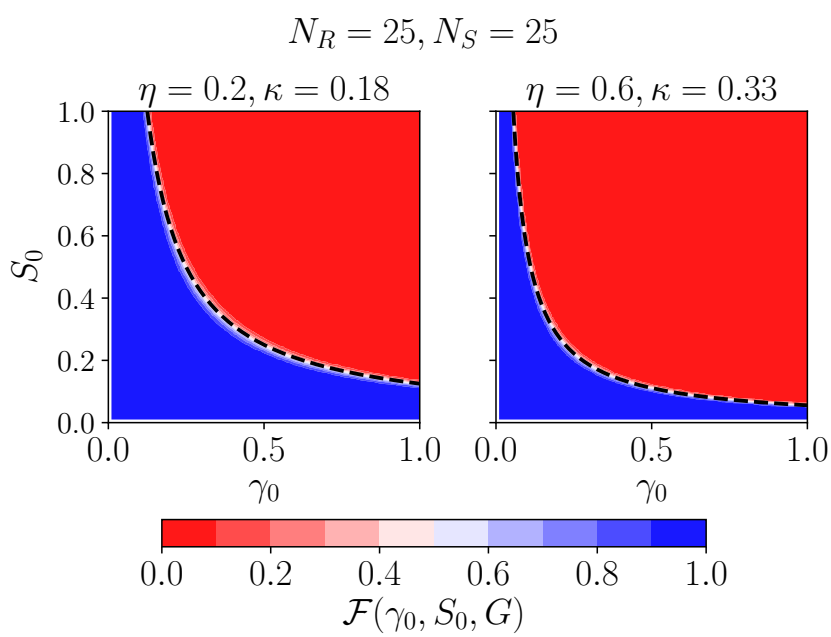

Figure 3.1: Plot of the feasability region. The color curve indicicates the feasibility function $\mathcal{F}(\gamma_0, S_0, \alpha_0 = 0, G)$ for G_1 (left) and G_2 (right). We observe a steep descent which marks a very clear transition from a totally feasible regime to a totally unfeasible regime, 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%.

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.

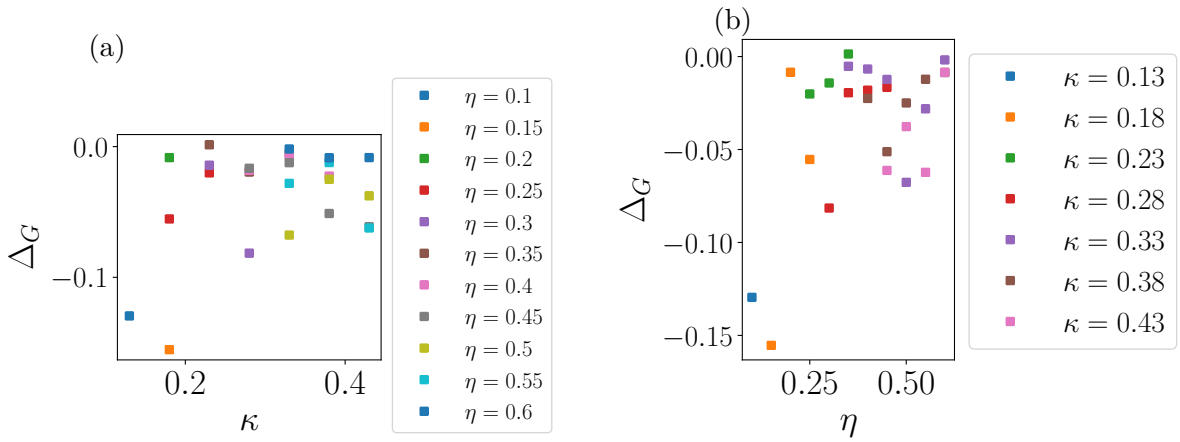

Figure 3.2: Relative error in the determination of the boundary of $\mathcal{V}_1^{G,0}$ (a) varying connectance at fixed nestedness and (b) varying nestedness at fixed connectance. The theoretical prediction tends to overestimate the measured value. The larger the nestedness and connectance, the better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

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 $S_0 = (0.043 \pm 10^{-8})/\gamma_0 - (4.6 \times 10^{-3} \pm 3 \times 10^{-8})$. The theoretical prediction is not as good as before with an error of ~ 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**.

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

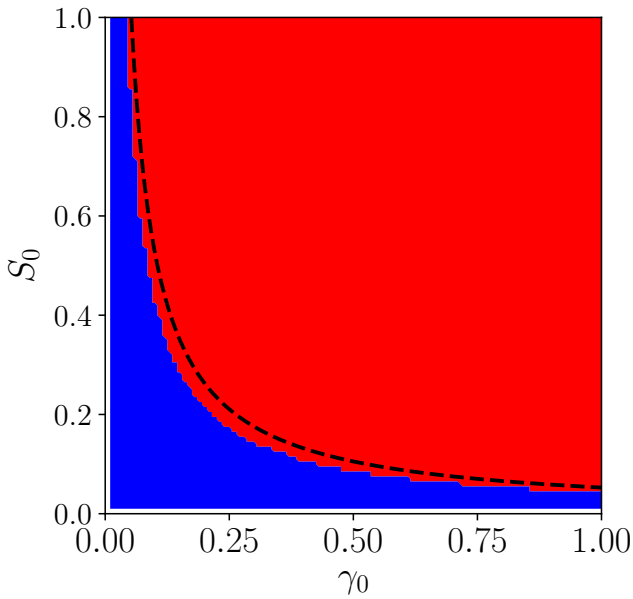

Figure 3.3: Plot of the common feasability region. The blue area indicates the common feasibility volume, computed numerically, while the dashed line shows the analytical prediction. Although the match is not as good as before, the relative error is only of the order of 20%. The red part is the area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

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 coprophagy" : the structure of A is such that consumers are not allowed to release what they consume.
- "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.

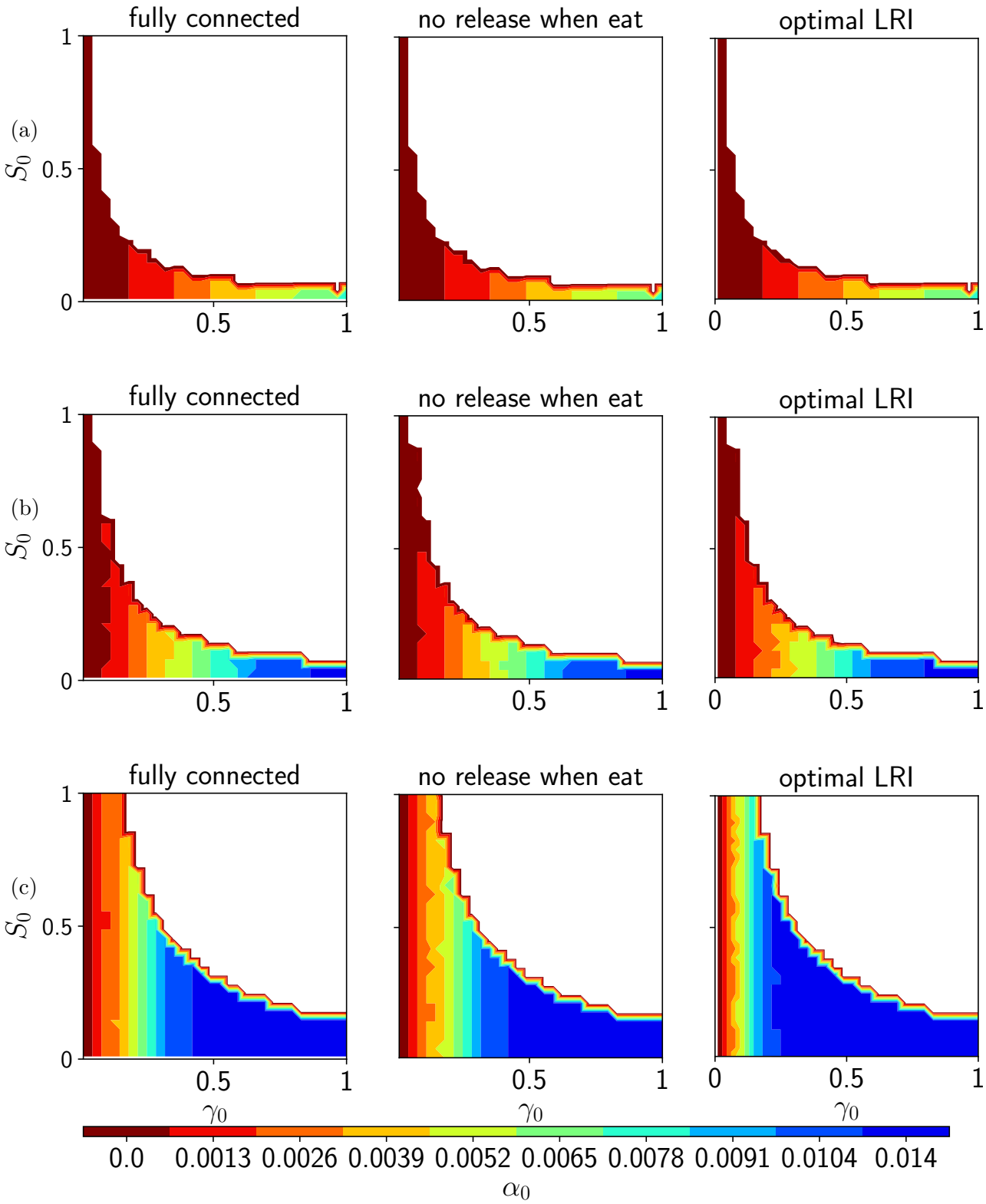

Figure 3.4: Fully feasible region in the $(\gamma_0, S_0) \in [0,1] \times [0,1]$ unit square as a function of syntrophy for different consumption matrices G : (a) $\eta_G = 0.6$, $\kappa_G = 0.32$, (b) $\eta_G = 0.35$, $\kappa_G = 0.22$ and (c) $\eta_G = 0.15$, $\kappa_G = 0.18$. The white zone corresponds 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.

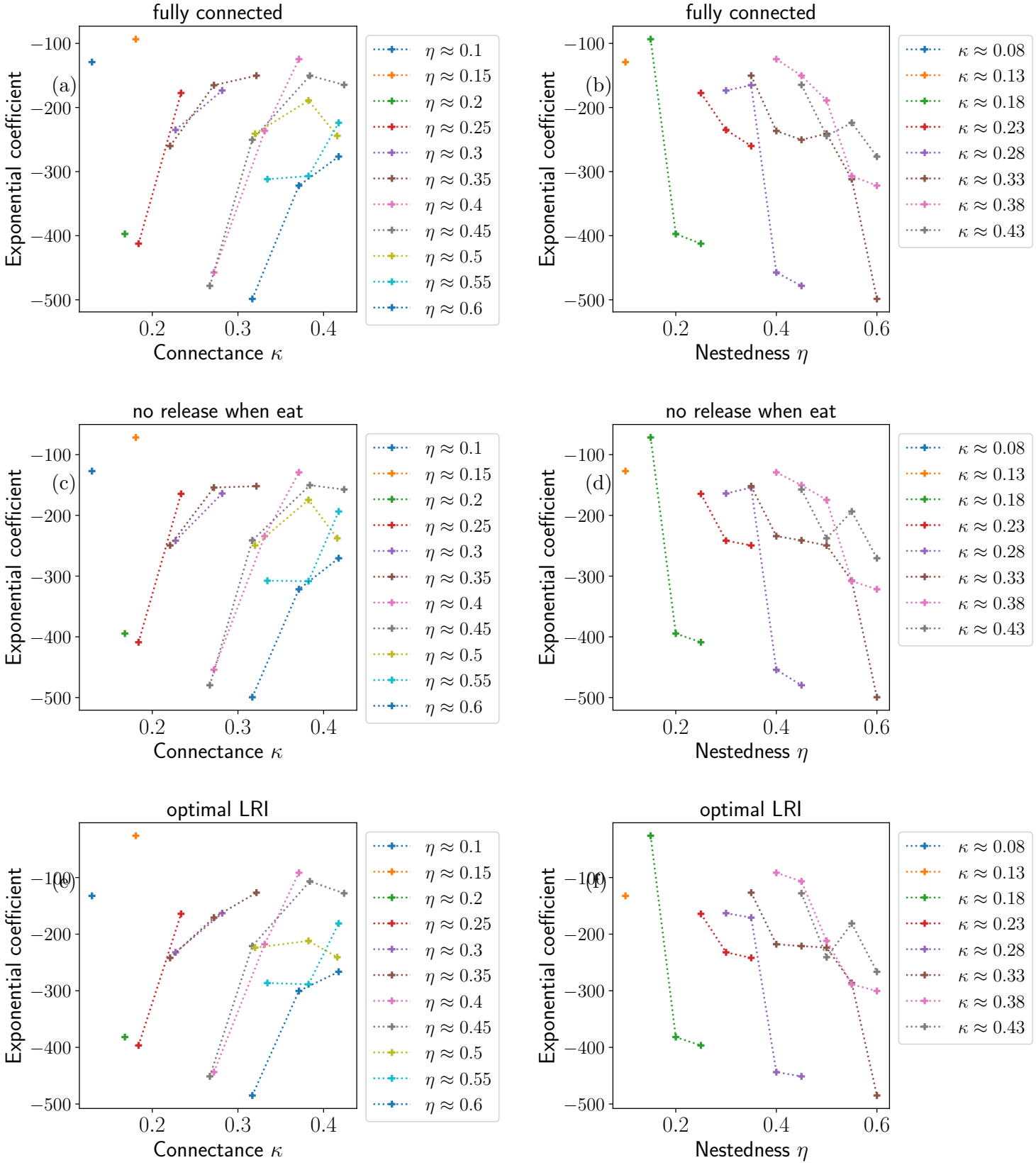

Figure 3.5: Feasibility decay rate. Plots on the first column (a)-(c)-(e) show how λ_F changes with connectance for a given ecological overlap, while plot on the second column (b)-(d)-(f) show how it evolves when ecological overlap is changed and connectance is kept fixed. Different structures of the A matrix are considered : (a)-(b) fully connected, (c)-(d) without intraspecific syntrophy and (e)-(f) outcome of the MC algorithm.

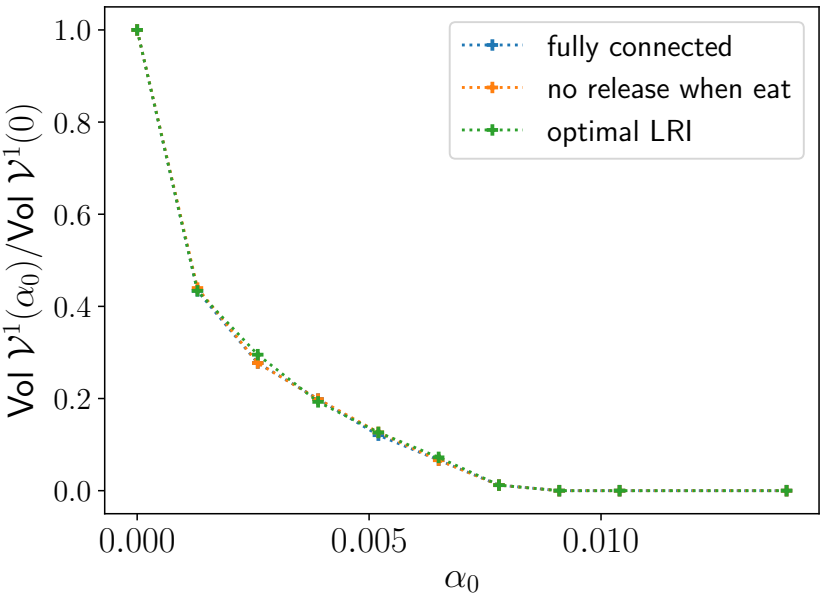

Figure 3.6: Volume of the common feasibility region $\mathcal{V}^1(\alpha_0)$ as a function of syntrophic interaction strength α_0. The curves indicate the different structures used for the syntrophy network A.

The influence of matrix structure

Common fully feasible volume

We looked at ten different α_0 values[21] from 0 to 0.015. Figure 3.6 shows that the volume of the common feasibility region shrinks down pretty fast until it reaches zero. The maximal feasible α_0 can be estimated with the help of Eq.(2.29):

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

so the maximal feasible syntrophy is well predicted by theory.

[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.

[21]This choice was made because after $\alpha_0 \approx 0.01$ the common feasibility volume shrinks down to zero.

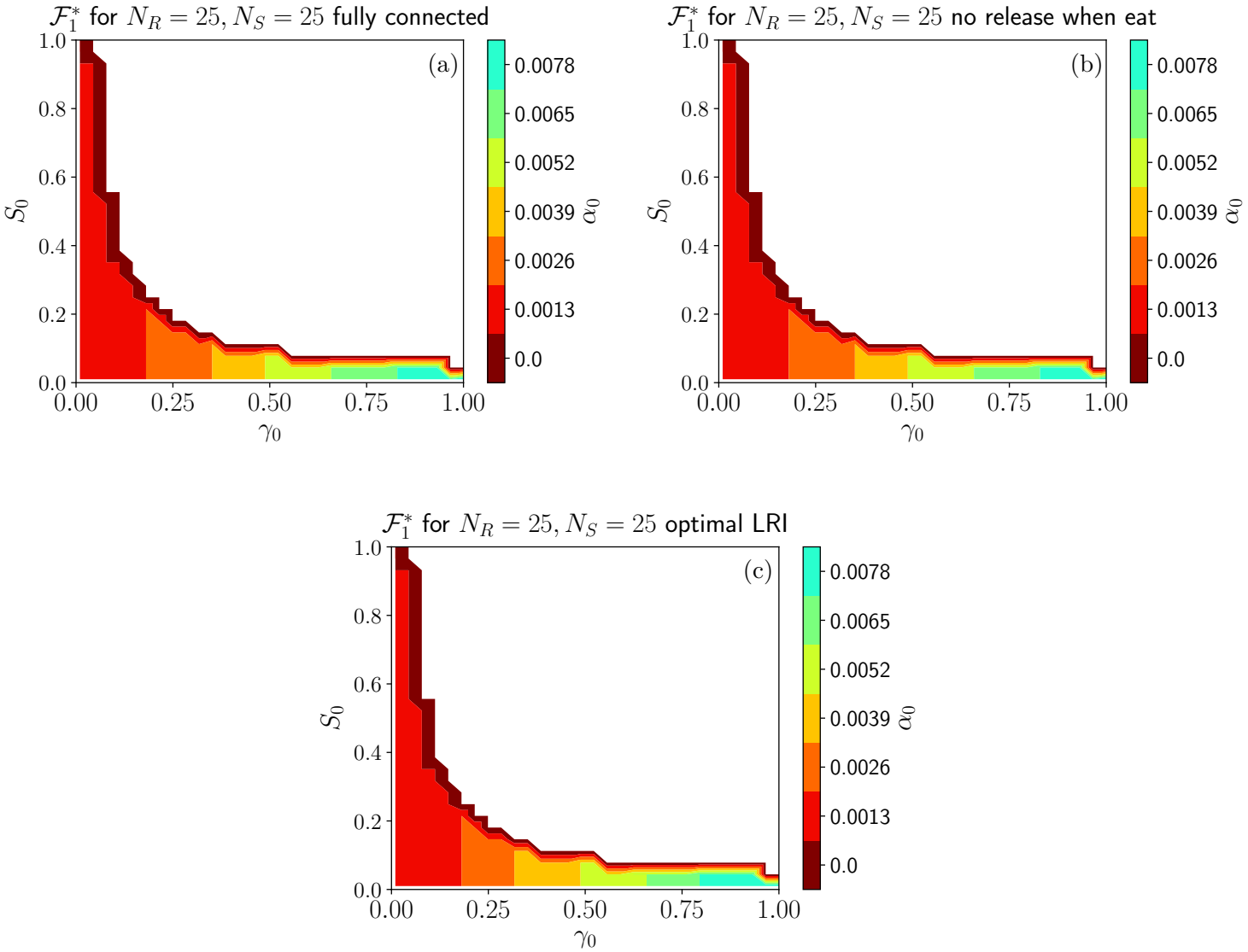

Figure 3.7: Surface plot of the fully feasible volume $\mathcal{V}^1(\alpha_0)$. The color bar on the side indicates the value of α_0 to which the surface corresponds. The white part of the plot corresponds to points that *never* are fully feasible. Note that even though it is not very clear on the figure $\mathcal{V}^1(\alpha_0^+) \subset \mathcal{V}^1(\alpha_0^-)\ \forall \alpha_0^+ > \alpha_0^-$, *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.

3.2 Dynamical stability

3.2.1 LRI regime – Outcome of the Monte Carlo algorithm

3.2.2 Fully dynamically stable volume

3.2.3 Largest eigenvalue of the jacobian

Equation 2.167 from Methods 2.2.7 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.(2.29) 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 3.10 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 on the contrary increases the maximal $|\mathrm{Re}\,(\lambda_1)|$ observed as shows Fig.3.11.

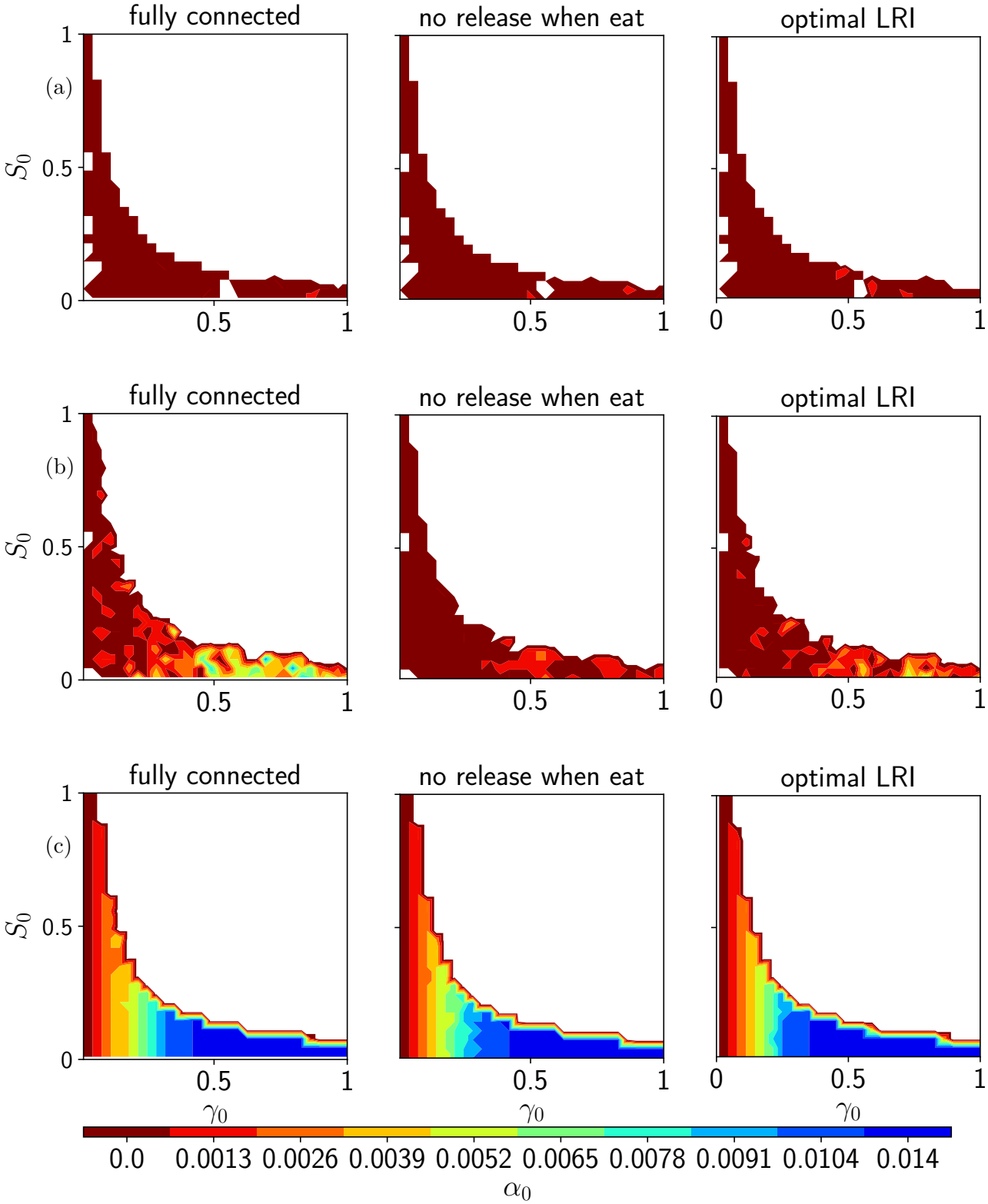

Figure 3.8: Locally fully dynamically stable region $\mathcal{D}_L^{G,1}$ as a function of syntrophy for different matrices G. The white zone corresponds to points that are never fully locally dynamically stable. The colour of a given point tells until which syntrophy that point is fully locally dynamically 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.

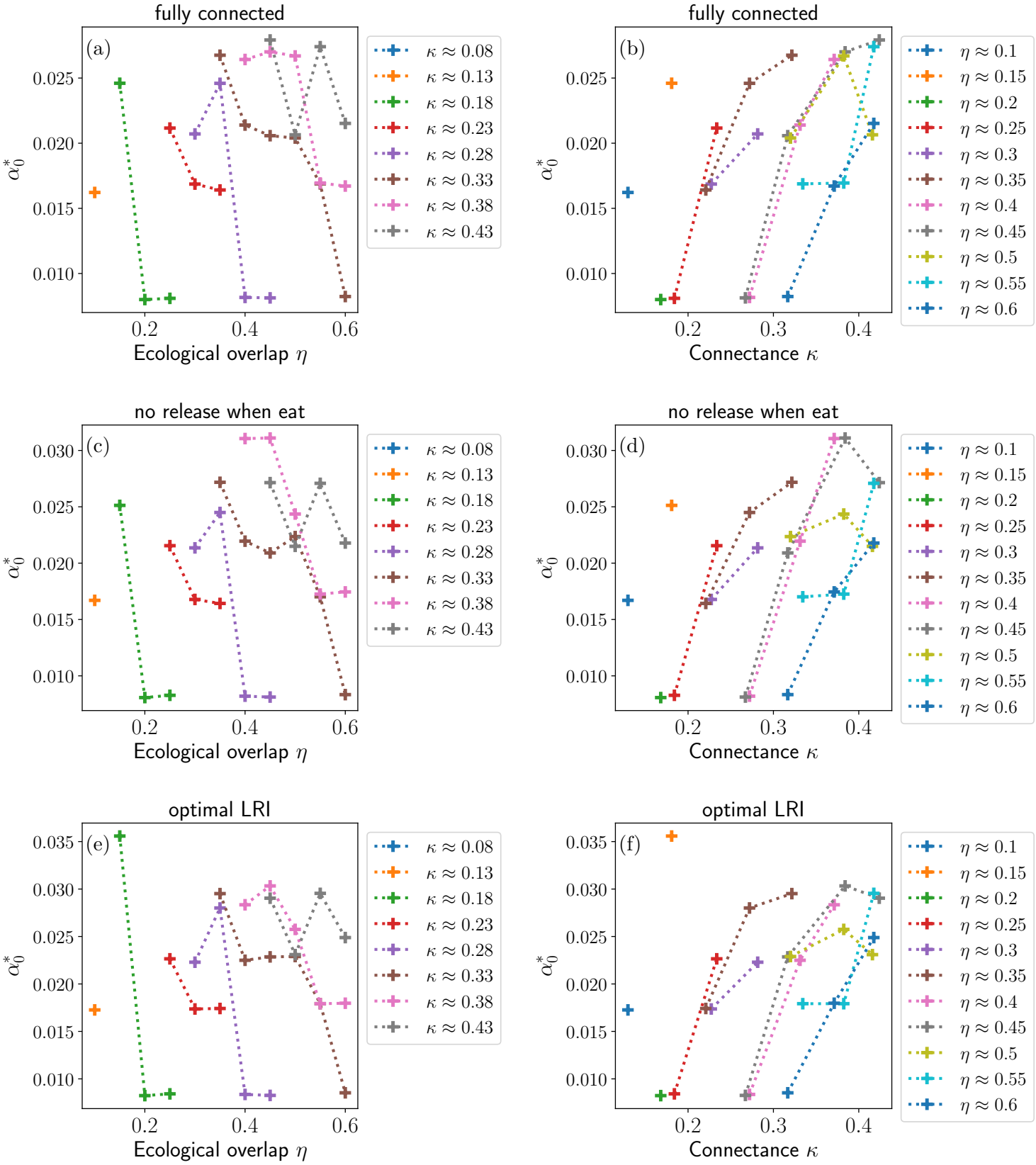

Figure 3.9: Critical syntrophy α_0^D, defined as the smallest syntrophy for which we can still find metaparameters that will give rise to fully dynamically stable systems. How α_0^D is estimated is explained in the main text. Errors on α_0^D are not plotted but are at most around 10%. (a)(c)(e) Evolution 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.

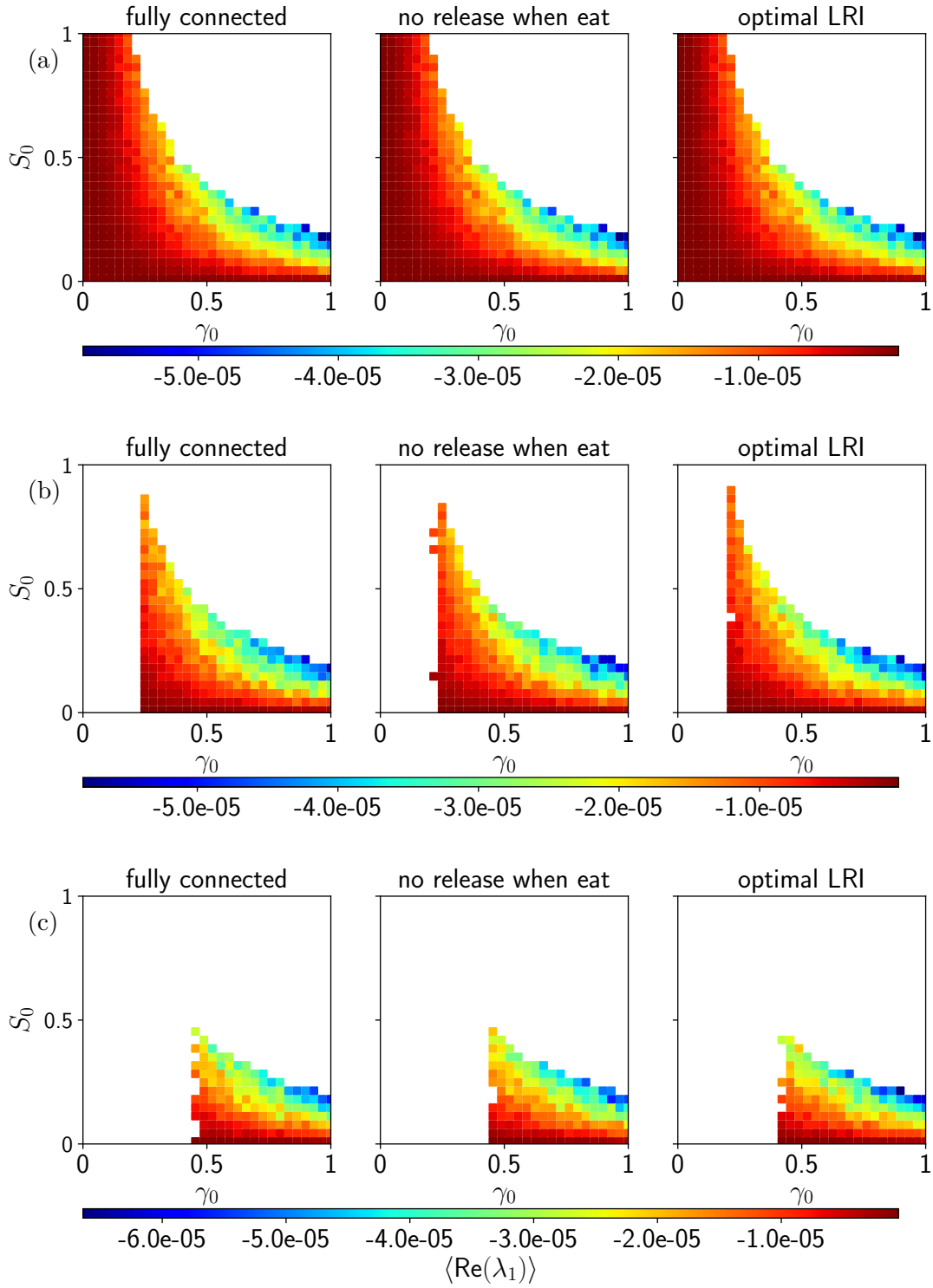

Figure 3.10: Largest real eigenvalue $\mathrm{Re}(\lambda_1)$ averaged over 300 realisations for each (γ_0, S_0) points for the consumption matrix G with consumers overlap $\eta_G = 0.1$ and connectance $\kappa_G = 0.13$. The white points correspond to not fully dynamically stable systems. Each row corresponds 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.

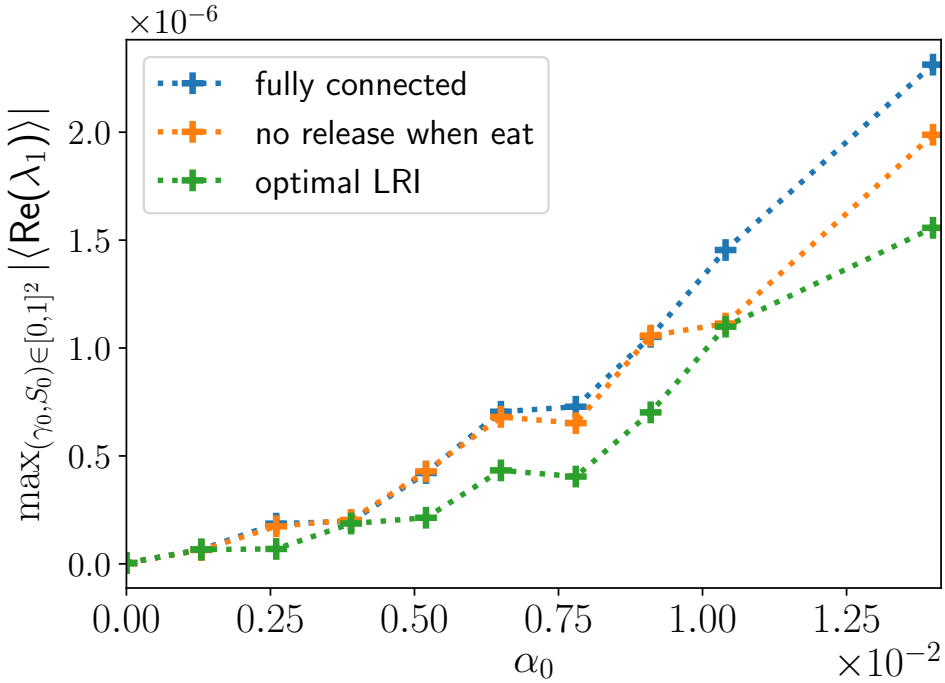

Figure 3.11: Evolution of the maximal $|\mathrm{Re}\,(\lambda_1)|$ observed in the $(\gamma_0, S_0) \in [0,1]^2$ region for a consumption matrix G with $\eta_G = 0.4$ and $\kappa_G = 0.38$. The different lines show the different stand for the different structure of the syntrophy matrix that we considered. The maximal eigenvalue increases in magnitude, making the system more dynamically stable, as syntrophy increases. That trend is true for all matrices we considered.

3.3 Structural stability

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:

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

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.

Bibliography

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

[12] OSCAR Perron. "Zur Theorie der Matrices". In: (), p. 16.

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

[14] 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.

[15] 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.

[16] 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.