## 0.1 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 $i = 1, \dots, N_S$) and the $N_R$ resources they feed off (denoted $R_\mu$ with $\mu = 1, \dots, N_R$).

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

$$\begin{cases} \dfrac{dR_\mu}{dt} = l_\mu - m_\mu R_\mu - \displaystyle\sum_j \gamma_{j\mu} R_\mu S_j + \sum_j \alpha_{\mu j} S_j & \text{(1a)} \\ \dfrac{dS_i}{dt} = \displaystyle\sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu S_i - d_i S_i - \sum_\nu \tau_{\nu i} S_i & \text{(1b)} \end{cases}$$

We will call the variables $\{l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i, \tau_{\mu i}\}$ *parameters of the model*. 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 $\mu$ abudance $R_\mu$). 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 $\mu$ death rate $m_\mu$).
- 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 $\mu$).

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 $R_\mu$ 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 $m_\mu$ : 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 $R_\nu$ 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 $R_\nu$ at rate $\tau_{\nu i}$ (this is the syntrophic interaction). In total $-\sum_\nu \tau_{\nu i} S_i$.

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.

### 0.1.1 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`.

#### 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 : $l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i, \tau_{\mu i}, R^*_\mu$ and $S^*_i$ $\forall i = 1, \dots, N_S; \mu = 1, \dots, N_R$. We define the *parameters space* $\mathcal{P}$ as the space that contains all the parameters :

$$\mathcal{P} \equiv \left\{ p : p = (l_\mu, m_\mu, d_i, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, \tau_{\mu i}, R^*_\mu, S^*_i) \right\} \tag{2}$$

Without taking into account the constraints on these parameters, there are $N_R + N_R + N_R N_S + N_R N_S + N_R N_S + N_S + N_R N_S + N_R + N_S = 3N_R + 2N_S + 4N_R N_S$ free parameters, so $\mathcal{P} \simeq \mathbb{R}^{3N_R + 2N_S + 4N_R N_S}$. Note that $\mathcal{P}$ 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 $N_R, N_S \approx 25, 50, 100, \dots$ *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] :

$$q_{i\mu} = \mathfrak{Q} Q_{i\mu} \tag{3}$$

where $\mathfrak{Q}$ is a random variable of mean $Q_0$ and standard deviation $\sigma_Q$. $Q_{i\mu}$ 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 :

$$l_\mu = \mathfrak{L} \tag{4a}$$
$$m_\mu = \mathfrak{M} \tag{4b}$$
$$\gamma_{i\mu} = \mathfrak{G} G_{i\mu} \tag{4c}$$
$$\alpha_{\mu i} = \mathfrak{A} A_{\mu i} \tag{4d}$$
$$\sigma_{i\mu} = \mathfrak{S} \tag{4e}$$
$$d_i = \mathfrak{D} \tag{4f}$$
$$\tau_{\mu i} = \mathfrak{T} \tag{4g}$$
$$R^*_\mu = \mathfrak{R} \tag{4h}$$
$$S^*_i = \mathfrak{S} \tag{4i}$$

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:

$$l_\mu \approx l_0 \tag{5a}$$
$$m_\mu \approx m_0 \tag{5b}$$
$$\gamma_{i\mu} \approx \gamma_0 G_{i\mu} \tag{5c}$$
$$\alpha_{\mu i} \approx \alpha_0 A_{\mu i} \tag{5d}$$
$$\sigma_{i\mu} \approx \sigma_0 \tag{5e}$$
$$d_i \approx d_0 \tag{5f}$$
$$\tau_{\mu i} \approx \tau_0 T_{\mu i} \tag{5g}$$
$$R^*_\mu \approx R_0 \tag{5h}$$
$$S^*_i \approx S_0 \tag{5i}$$

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 $\mathcal{B}^2_{N_R \times N_S} \times \mathcal{B}_{N_S \times N_R}$, where $\mathcal{M}$ is the *metaparameters space* :

$$\mathcal{M} \equiv \{m : m = (l_0, m_0, d_0, \gamma_0, \alpha_0, \sigma_0, \tau_0, R_0, S_0)\} \tag{6}$$

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 $\mathcal{C} : \mathcal{P} \to \mathcal{M} \times \mathcal{B}^2_{N_R \times N_S} \times \mathcal{B}_{N_S \times N_R}$ 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 $3N_R + 2N_S + 4N_R N_S$ (continous) to 9 (continous) + $3N_R N_S$ (binary). And to make the problem even simpler, instead of looking at each entry of the binary matrices $G, A$ and $T$ 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[1]. 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]:

$$\eta(M) \equiv \frac{\sum_{i<j} n_{ij}}{\sum_{i<j} \min(n_i, n_j)} \tag{7}$$

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

$$n_i \equiv \sum_k M_{ik}, \tag{8}$$

  and $n_{ij}$ is the overlap matrix defined as

$$n_{ij} \equiv \sum_k M_{ik} M_{jk}. \tag{9}$$

- Its **connectance** : this measure, simply defined as the ratio of non-zero links in a matrix, is central in the study of plants-and-animals systems [1]. It is formally defined for a matrix $q_{ij}$ of size $N \times M$ as :

$$\kappa(q) \equiv \frac{\sum_{ij} Q_{ij}}{NM} \tag{10}$$

  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)? The answer is simple : we define a function $\mathcal{A} : \mathcal{M} \times \mathcal{B}^2_{N_R \times N_S} \times \mathcal{B}_{N_S \times N_R} \to \mathcal{P}$ which brings us from the collapsed space to the parameter space[2]. Numerically what we will do is : from a set of metaparameters $m \in \mathcal{M}$ and binary matrices $B = (G, A)$, produce a (or several) set(s) of parameters $p = \mathcal{A}(m, B) \in \mathcal{P}$ and study properties of it. Section **??** details how $\mathcal{A}$ is constructed.

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

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