Random Zoo with HOIs (Supplementary)

Pablo Lechon-Alonso

May 2023

1 Notation

Throughout this material we will be dealing with scalars, matrices, and tensors of order arbitrarily greater than 2. Scalars are denoted by non-capitalized letters with subindices. For example, the abundance of species i is a scalar, denoted by x_i. Vectors are denoted by dropping the index. For example, the abundance vector of species in a community is x. Finally, matrices are denoted in capital letters. For example, the matrix with pairwise interactions of a community is denoted by A. Tensors of order greater than 2 are also denoted with capital letters. To avoid confusion with matrices, their dimension is stated when they are introduced. Slices of tensors are denoted by capital letters with sub-indices. For example, if B is a tensor of order 3 (a 3 dimensional array $3 \times 3 \times 3$), then B_i is a slice of such tensor, that is, a 3×3 matrix, and B_{ij} is a 3×1 column vector.

Multi-index notation will also be used throughout. As such, we will denote multi-indices with bold font, and indices with regular font. For example, the multi-index $\boldsymbol{j}$ containing n different indices is denoted as $\boldsymbol{j} = (j_1, \dots j_n)$.

Derivatives with respect to time are represented with dots, i.e. $\frac{dx}{dt} = \dot{x}$

2 Generalized Lotka-Volterra model with higher-order interactions

Consider the Generalized Lotka-Volterra (GLV) model of a system with n species, interacting in singles, pairs, triplets, and up to groups of $h+1$ species (we keep the focal species separately for notational convenience). That is the GLV model with interactions of arbitrarily high order. We can write the abundance dynamics of species i in the system as

$$\begin{aligned}\dot{x}_i &= x_i F_i(x_1, \dots, x_n) \\ &= x_i \left(r_i + \sum_{j_1} A_{ij_1} x_{j_1} + \sum_{j_1 j_2} B_{ij_1 j_2} x_{j_1} x_{j_2} + \dots + \sum_{j_1 \dots j_{d+1}} C_{ij_1 \dots j_{h+1}} x_{j_1} \dots x_{j_{h+1}} \right), \end{aligned} \quad (1)$$

where index i represents the focal population of species, and indices $j_1, \dots, j_{h+1}$, (with $j_i \in \{1, \dots, n\}$) label which other species' interact with population i. The first term, r_i, is the growth rate of species i, which can be interpreted as an interaction of order 1. The

A_{ij_1} coefficients capture interactions of order 2 (between pairs of species x_i and x_{j_1}); $B_{ij_1j_2}$ coefficients three-way interactions (between triplets of species x_i x_{j_1} and x_{j_2}), and so on, up to order $h+1$.

The first step we are interested in is to analytically compute what are the feasible equilibria of this system. This amounts to finding the real and positive solutions of the following system of n polynomials of degree h

$$F_i(x_1, \ldots, x_n) = 0 \ \forall i. \tag{2}$$

which in general has a total of h^n complex roots (Bézout, 1779). However, this quickly becomes an impossible task as n and h increase. In fact, Abel's impossibility theorem (Abel, 2012) states that there is no solution in radicals to general polynomial equations of degree five or higher with arbitrary coefficients. As such, we can only calculate the roots of Eq. 2 numerically. Despite not being able to compute the feasible equilibria analytically, we can make statistical statements about them if we sample the coefficients of Eq. 1 at random from a Gaussian distribution centered at zero and with an appropriate variance structure. Under these assumptions, the polynomials in Eq. 2 become the well-studied Kostlan-Shub-Smale (KSS) system of random polynomials (Section 3). This system of polynomials has been studied exhaustively since the early 90's (Kostlan, 1993; Shub, Smale, 1993a,b; Edelman, Kostlan, 1995; Armentano et al., 2018; Azaïs, Wschebor, 2004; Wschebor, 2005; Dalmao, 2015), and bears results that are of ecological interest.

3 Kostlan-Shub-Smale system of random polynomials

Consider a system $\boldsymbol{P}$ of n polynomials with common degree $d > 1$, on n variables. The i^{th} equation of such system can be written as

$$P_i(x_1, \ldots x_n) = \sum_{|\boldsymbol{l}| \leq d} a_{\boldsymbol{l}}^{(i)} \boldsymbol{x}^{\boldsymbol{l}}, \tag{3}$$

where

1. $\boldsymbol{l} = (j_1, \ldots, j_n) \in \mathbb{N}^{[0,n]}$, and $|\boldsymbol{l}| = \sum_{k=1}^{n} l_k$,
2. $a_{\boldsymbol{l}}^{(i)} = a_{l_1,\ldots,l_n}^{(i)} \in \mathbb{R}$, $|\boldsymbol{l}| \leq d$,
3. $\boldsymbol{x} = (x_1, \ldots x_n)$ and $\boldsymbol{x}^{\boldsymbol{l}} = \prod_{k=1}^{n} x_k^{l_k}$.

We say that the system $\boldsymbol{P}$ has the Kostlan-Shub-Smale (KSS) distribution (**?**Armentano et al., 2018) if the coefficients $a_{\boldsymbol{l}}^{(i)}$ are i.i.d. random variables with mean 0 and variance

$$\mathrm{Var}(a_{\boldsymbol{l}}^{(i)}) = \binom{d}{\boldsymbol{l}} = \frac{d!}{l_1! \ldots l_n! (d - |\boldsymbol{l}|)!}. \tag{4}$$

Importantly, for this class of systems of polynomials, it has been shown (**?**) that the expected number of real roots is

$$\mathbb{E}[N] = \sqrt{d^n} \tag{5}$$

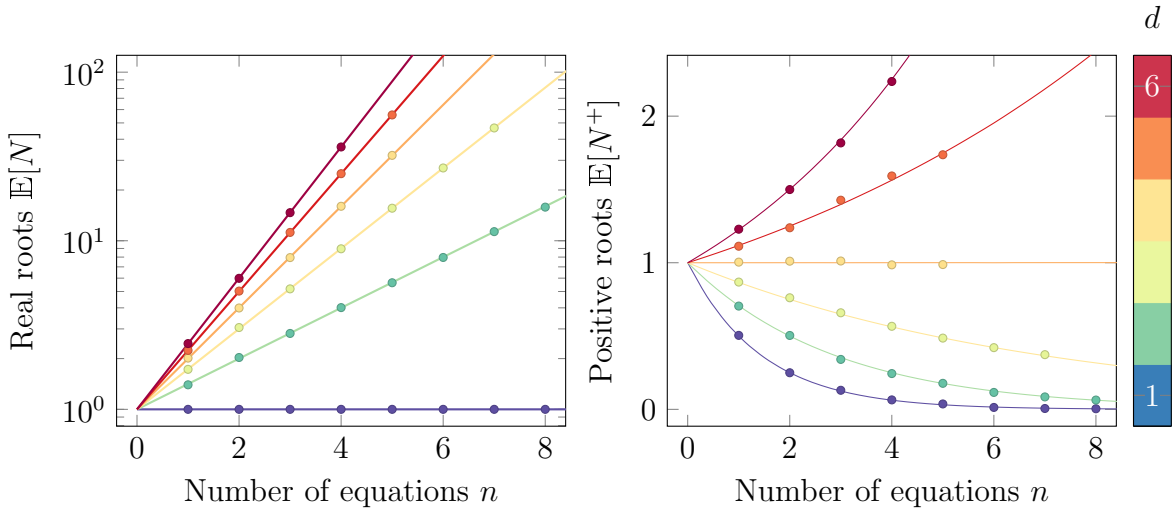

Figure 1: **Expected number of real roots and positive roots.** On the **left** panel we plot (semilog scale) the expected number of real roots (Eq. 5) for different values of d, and overlay the average value of real roots obtained from simulations of the KSS polynomial system. On the **right** panel we plot the expected number of positive solutions. Importantly, this number grows with n for $d > 5$

By symmetry, the expected number of positive roots is just

$$\mathbb{E}[N^+] = \frac{1}{2^n}\mathbb{E}[N] \tag{6}$$

Numerical simulations of the system in Eq. 3 using homotopy continuation (Breiding, Timme, 2018) (Fig. 1) track well both of the theoretical curves (Eqs. 5 and 6). Importantly, the expected number of positive roots grows with n for $d > 5$. In the next section we show establish the connection between this system and the GLV. Thereafter we make use of such connection to study the ecological consequences of the divergence of the number of positive roots.

4 GLV and KSS equivalence

Here we show that the system of polynomials arising when solving for feasible equilibria of Eq. 1 is precisely the KSS system of random polynomials when the coefficients in $r, A, B, \ldots$ are sampled with appropriate variances. Recently, this equivalence has been shown for the replicator equation with higher order interactions (Duong, Han, 2023).

First, we expand Eq. 3 as

$$\begin{aligned} P_i(x_1, \ldots x_n) &= \sum_{|\boldsymbol{l}|\leq d} a_{\boldsymbol{l}}^{(i)} \prod_{k=1}^{n} x_k^{l_k} \\ &= \sum_{|\boldsymbol{l}|=0} a_{0\ldots0}^{(i)} + \sum_{|\boldsymbol{l}|=1} a_{l_1\ldots l_n}^{(i)} \prod_{k=1}^{n} x_k^{l_k} + \cdots + \sum_{|\boldsymbol{l}|=d} a_{l_1\ldots l_n}^{(i)} \prod_{k=1}^{n} x_k^{l_k}. \end{aligned} \tag{7}$$

The first commonality between systems 7 and 2 is that they have the same total degree; that is $h = d$. We can now identify coefficients of Eq. 2 with the corresponding coefficients of Eq. 7 that multiply monomials of equal degree. For the first term in the sum of Eq. 7, that with $|\boldsymbol{l}| = 0$, there is only one monomial of degree 0; 1. As such, the corresponding coefficient in the GLV model (Eq. 1) is simply

$$r_i = a^{(i)}_{0\ldots0}$$

For r_i to be a coefficient of the KSS system, its variance should be that of Eq. 4

$$\mathrm{Var}[r_i] = \binom{h}{\boldsymbol{l}} = 1.$$

For the coefficients of the second term of the sum in Eq. 7 (those paired with monomials of degree $|\boldsymbol{l}| = 1$), the corresponding coefficients in the GLV model would be

$$\begin{aligned}
\sum_{j_1} A_{ij_1} x_{j_1} &= \sum_{|\boldsymbol{l}|=1} a^{(i)}_{l_1\ldots l_n} \prod_{k=1}^{n} x_k^{l_k} \\
A_{i1}x_1 + A_{i2}x_2 + \cdots + A_{in}x_n &= a^{(i)}_{10\ldots0}x_1 + a^{(i)}_{01\ldots0}x_2 + \ldots a^{(i)}_{00\ldots1}x_n \\
A_{ij} &= a^{(i)}_{\boldsymbol{l}_j},
\end{aligned}$$

where $\boldsymbol{l}_j$ denotes the multiindex of all zeros and a 1 in position j. As before, the variance of the A_{ij} coefficients should be

$$\mathrm{Var}[A_{ij}] = \binom{h}{h-1} = h,$$

Next, we identify with the GLV coefficients associated with interactions of order 3, B_{ijk}, with the coefficients of monomials of degree $|\boldsymbol{l}| = 2$ in Eq. 7

$$\begin{aligned}
\sum_{|\boldsymbol{l}|=2} a^{(i)}_{l_1\ldots l_n} \prod_{k=1}^{n} x_k^{l_k} &= \sum_{j_1 j_2} B_{ij_1j_2} x_{j_1} x_{j_2} \\
B_{i11}x_1^2 + (B_{i12} + B_{i21})x_1x_2 + \cdots + & \\
(B_{i1n} + B_{in1})x_1x_n + & \\
(B_{i2n} + B_{in2})x_2x_n + \cdots + B_{inn}x_n^2 &= a^{(i)}_{20\ldots0}x_1^2 + a^{(i)}_{11\ldots0}x_1x_2 + a^{(i)}_{10\ldots1}x_1x_n + \ldots \\
& \quad a^{(i)}_{10\ldots1}x_1x_n + a^{(i)}_{01\ldots1}x_2x_n + \cdots + a^{(i)}_{00\ldots2}x_n^2
\end{aligned}$$

The first thing to note from the above expression is that for $|\boldsymbol{l}| = 2$, we sometimes find more than one GLV parameter per KSS coefficient. In particular, when $j = k$ there correspondence is 1 to 1, but wehen $j \neq k$ it is 1 to 2. This is due to the commutative property of multiplication. Therefore, the variance of the B_{ijk} coefficients appears to be index dependent. Let us, apply Eq. 4 to each case, we have that

$$\mathrm{Var}[B_{ijj}] = \mathrm{Var}[a^{(i)}_{0\ldots2\ldots0}] = \frac{1}{2}h(h-1)$$

and

$$\mathrm{Var}[B_{ijk}] + \mathrm{Var}[B_{ikj}] = \mathrm{Var}[a^{(i)}_{0\ldots1\ldots1\ldots0}]$$

Without loss of generality we assume the variances of coefficients with indices $\{j, k\}$ that are permutations of each other to be the same. Plugging in Eq. 4, we get

$$\mathrm{Var}[B_{i\sigma(jk)}] = \frac{1}{2}h(h-1),$$

where $\sigma(jk)$ denotes permutations of the set $\{j, k\}$. Conveniently, the variance of the element B_{ijk} turns out to not depend on its position on the tensor. As we will see next, this is not an artifact specific to order 2, but a general feature.

Lastly, we consider the general case of monomials of order $|\boldsymbol{l}| = p$

$$\sum_{|\boldsymbol{l}|=p} a^{(i)}_{l_1\dots l_n} \prod_{k=1}^{n} x_k^{l_k} = \sum_{j_1 j_2} T_{ij_1\dots j_p} x_{j_1} \dots x_{j_p}.$$

Due to the commutative property, identifying coefficients on the same monomials leads to the following equality for any given monomial of degree p

$$a^{(i)}_{l_1\dots l_n} = \sum_{\sigma(j_1\dots j_p)} T_{ij_1\dots j_p}, \tag{8}$$

where the sum is taken over all the permutations of the set of elements $\mathcal{S} = \{j_1 \dots j_p\}$. Taking the variance on both sides leads to

$$\mathrm{Var}[a^{(i)}_{l_1\dots l_n}] = \mathrm{Var}\left[\sum_{\sigma(\mathcal{S})} T_{ij_1\dots j_p}\right] = r\mathrm{Var}[T_{ij_1\dots j_p}]. \tag{9}$$

Here we have stressed that the sum is not performed over different values of indices. Instead, given a set of values for each index, $\mathcal{S}$, we sum over all the permutations of such set. There are a total of r number of permutations of $\mathcal{S}$, accounting for repetitions. In fact, given a set of p objects such that there are p_1 identical objects of type 1, p_2 identical objects of type 2... and p_k identical objects of type k, the number of permutations of such set is

$$r = \frac{p!}{p_1! \cdots p_k!}$$

Note that since we are identifying coefficients, the numbers $p_1, \dots, p_k$ are indeed $l_1, \dots, l_k$. Making these substitutions and plugging in Eq. 4 into Eq. 9, we can solve for the variance of any coefficient of $T_{ij_1\dots j_p}$ arriving to

$$\mathrm{Var}[T_{i\mathrm{j}_1\dots j_p}] = \frac{l_1! \cdots l_n!}{h!} \frac{h!}{l_1! \dots l_n!(h-p)!} = \frac{h!}{p!(h-p)!}. \tag{10}$$

Surprisingly, this variance is independent of the multiindex structure, and only depends in the order of interaction order of the particular coefficient p, and h, which sets the overall order of HOIs in the ecosystem (Fig. 2).

We have shown that, as long as the coefficients of the GLV model are sampled from a Gaussian distribution with mean 0 and variance given by Eq. 10, then it holds that

$$F_i(x_1, \dots, x_n) \equiv P_i(x_1, \dots x_n),$$

with $h = d$. As such, all the results derived for the KSS system of random polynomials can be applied to the random GLV with higher-order interactions.

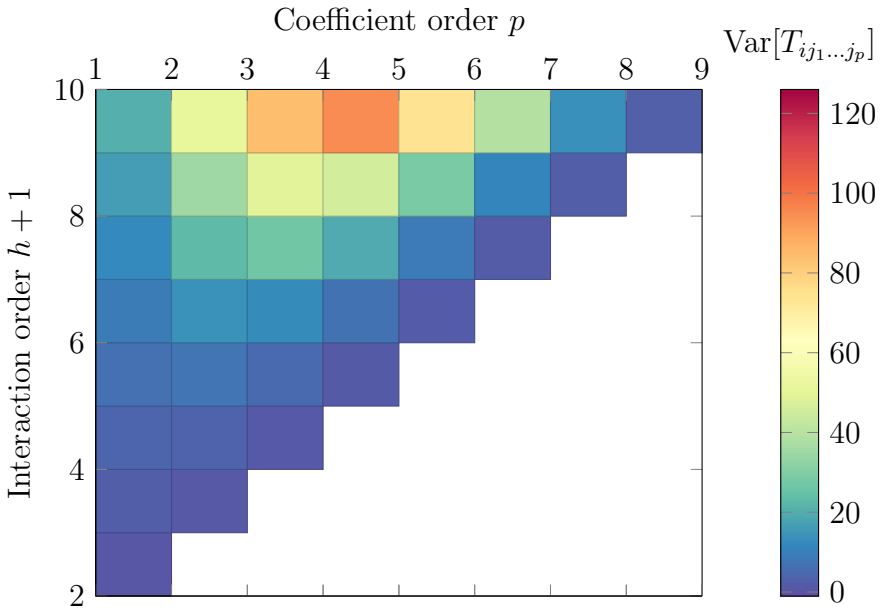

Figure 2: **Variance of the GLV coefficients**. We plot the variance obtained in Eq. 10 as a function of coefficient order p and overall interaction order of the system h. For a given coefficient order p (vertical stripes), its variance increases as the interaction order of the system increases. On the other hand, for a given interaction order h (horizontal stripes), the variance of the coefficient reaches a maximum for intermediate values of its order.

5 Probability of feasibility

We are concerned with the real positive roots; feasible equilibria of the GLV system with HOIs (Eq. 1. The probability to have at least one feasible equilibrium can be written as one minus the probability that no equilibria is feasible,

$$P_f = 1 - \sum_{n_i=0}^{h^n} p_i \left(1 - \frac{1}{2^n}\right)^{n_i} = 1 - \mathbb{E}[f(n_i)]. \tag{11}$$

Here, p_i is the probability that system 1 has n_i real roots, i.e., $P(N^+ = n_i)$, where N^+ is the random variable representing number of real roots. The function $f(n_i) = (1-2^{-n})^{n_i}$, is the probability that none of the n_i roots are feasible. We sum over all possible number of real roots; from 0 (all roots would be complex) to h^n (all roots would be real). Additionally, The case where $h = 1$ has been well characterized by Serván et al. (2018) (dashed line in Fig. 3). In this case, $h^n = 1$, and Eq. 11 reduces to

$$P_f = \frac{1}{2^n}.$$

When $h > 2$, Eq. 2 can have more than 1 real solutions. The full probability distribution of the number of real roots, X, is unknown, but we know its first moment $\mathbb{E}[X] = \bar{X}$ thanks to the GLV-KSS equivalence (Duong, Han, 2023). Thus, we can use it to

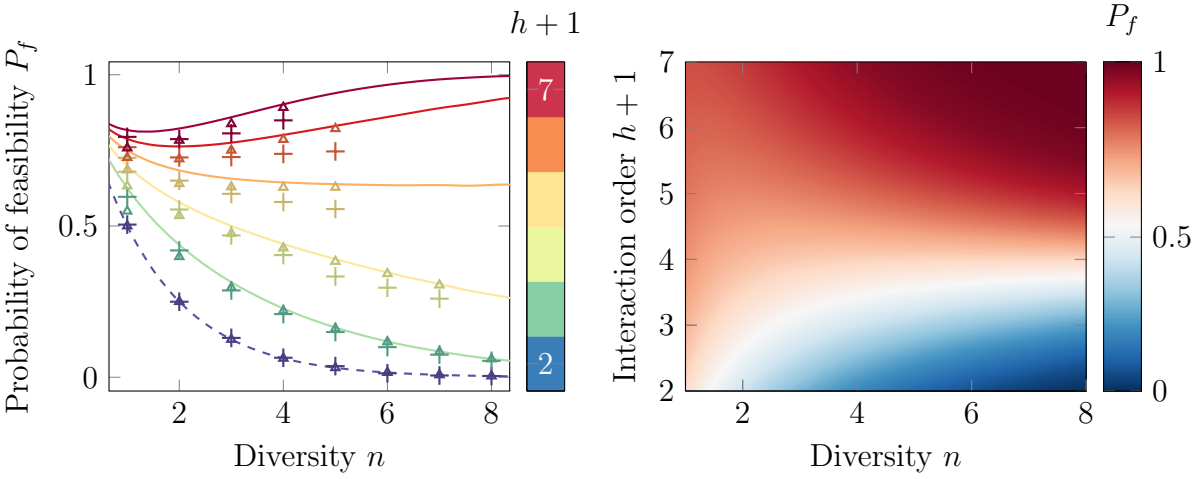

Figure 3: **Probability of feasibility as a function of diversity and interaction order.** Both panels show that the probability of feasibility drops to 0 exponentially with n for $h < 5$, but approximates a constant for $h = 5$, and converges to 1 for $h > 5$. In the **left** panel we plot the upper (solid lines) and lower (triangles) bounds of P_f (Eq. 12), and corresponding numerical simulations (crosses) for computationally tractable values of n, and h. The approximation to first order (lines) is an analytical upper bound (Eq. 14). The approximation to second order (triangles) requires numerical estimation of the variance. On the **right** panel we plot a heat map of the upper bound of P_f (computed from Eq. 12) as a function of both n and h.

approximate $\mathbb{E}[f(i)]$ through a Taylor expansion around the mean

$$
\begin{aligned}
P_f &= 1 - \left(f(\bar{X}) + \frac{df}{dn_i}\Big|_{n_i=\bar{X}} \mathbb{E}[n_i - \bar{X}] + \frac{1}{2}\frac{d^2 f}{dn_i^2}\Big|_{n_i=\bar{X}} \mathbb{E}[(n_i - \bar{X})^2] + \mathcal{O}\left(E[(n_i - \bar{X})^3]\right) \right) \\
&\approx 1 - \left(f(\bar{X}) + \frac{1}{2}\frac{d^2 f}{dn_i^2}\Big|_{n_i=\bar{X}} \mathrm{Var}[X] \right) \\
&= 1 - \left(\left(1 - \frac{1}{2^n}\right)^{h^{n/2}} + \frac{1}{2}\left(1 - \frac{1}{2^n}\right)^{h^{n/2}} \log^2\left(1 - \frac{1}{2^n}\right) \mathrm{Var}[X] \right) \\
&= 1 - \left(1 - \frac{1}{2^n}\right)^{h^{n/2}} \left(1 + \log^2\left(1 - \frac{1}{2^n}\right) \mathrm{Var}[X] \right) \\
&\approx 1 - \left(1 - \frac{1}{2^n}\right)^{h^{n/2}} \left(1 + \frac{1}{2^{2n+1}} \mathrm{Var}[X] \right). \qquad (12)
\end{aligned}
$$

Here we have truncated the expansion to second order, and made the approximation that $\log^2(1-x) \approx x^2$ when $x << 1$. Note that this approximation becomes valid early on as n increases.

Importantly, the first order approximation is an upper bound to P_f. We can see this using Jensen's inequality, which has the form

$$
\psi\left(\sum_i p_i x_i\right) \leq \sum_i p_i \psi(x_i), \qquad (13)
$$

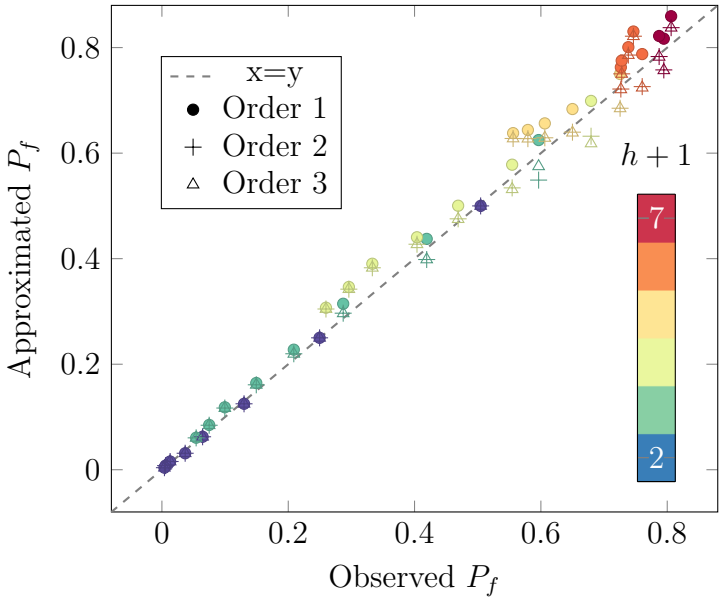

Figure 4: **Goodness of the Taylor expansion of Eq. 11.** Approximated probability of feasibility for Taylor expansions of orders 1, 2, and 3 versus the observed probability of feasibility in simulations. The one to one line (dashed grey) is plotted for reference. The second order expansion improves the approximation marginally compared to that of order 1, while the improvement of the third order is negligible.

where $p_i \geq 0$, $\sum_i p_i = 1$, and ψ is convex. In our case, $\psi(x_i) = f(n_i)$. This function is convex on n_i, since its second derivative

$$\frac{d^2 f}{dn^2} = f(n_i) \log^2 \left(1 - 2^{-n}\right) > 0$$

is always positive. We apply this inequality to Eq. 11 by multiplying Eq. 13 by -1, and adding 1, arriving to

$$1 - \sum_i p_i f(n_i) \leq 1 - f\left(\sum_i p_i n_i\right)$$
$$P_f \leq 1 - f(\bar{X}). \tag{14}$$

The second order approximation contains the variance of the number of real roots. Unfortunately, there is no analytical expression for $\mathrm{Var}[X]$. That is, if we truncate the expansion at second order, we need to compute the variance numerically from simulations. In Fig. 4 we check how good is the Jensen approximation by plotting the observed P_f versus approximations to different order. The first order approximation is already a very good one. The second order improvement is marginal, and the the third order one is negligible (Fig. 4). Additionally, this figure also shows the pattern reported in Fig. 3, since redder points (corresponding to high d) cluster around high probability of feasibility.

6 (Not) Recapitulating results in classical models

Consider the classical resource-competition model (Smith, Waltman, 1995; Posfai et al., 2017) where n species x_i (with $i \in \{1, ..., n\}$) compete for m resources y_j (with $j \in \{1, ..., m\}$). Each species x_i consumes resource j at some rate C_{ij}, $\in \mathbb{R}_+^{m\times 1}$, which specify the rate of utilization of each resource. We assume chemostat dynamics, such that the concentration of the resources is given by a resource supply rate $\vec{\kappa}$, the uptake of nutrients by the present species, and by a dilution rate μ_j. As such, the dynamics of the concentration of resource j are

$$\dot{y}_j = \kappa_j - y_j \sum_i C_{ij} x_i - y_j \mu_j.$$

The species in the chemostat grow according to their resource consumption, while allocating some energy for maintenance, such that we can write the dynamics of species i as

$$\dot{x}_i = x_i \left(\sum_j C_{ij} y_j - d_i \right).$$

If we assume that resource dynamics are much faster than species dynamics, we can perform a separation of timescales by setting $\dot{y}_j = 0$, such that aarriving at

$$y_j^\star = \frac{\kappa_j}{\mu_j + \sum_i C_{ij} x_i}.$$

Plugging the above equilibrium into the equation for the species we get

$$\dot{x}_i = x_i \left(\sum_j \frac{C_{ij}\kappa_j}{\mu_j + \sum_i C_{ij} x_i} - d_i \right). \tag{15}$$

In order to express this model as a GLV model with higher order interactions, we re-scale time by the function,

$$\tau = \int \frac{1}{S(t)} dt = \int \prod_k \left(\mu_k + \sum_i C_{ik} x_i(t) \right)^{-1} dt.$$

This is an appropriate time re-scale because $S(t) \geq 0$, such that τ will always increase. Applying the chain rule we have that

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{dx_i}{d\tau}\frac{d\tau}{dt} \\ &= \frac{1}{S(t)}\frac{dx_i}{d\tau} \end{aligned} \tag{16}$$

Substituting Eq. 16 in Eq. 15, we can re-write its dynamics in the re-scaled time as a GLV with HOIs

$$\begin{aligned} \frac{dx_i}{d\tau} &= S(t) x_i \left(\sum_j \frac{C_{ij}\kappa_j}{\mu_j + \sum_i C_{ij} x_i} - d_i \right) \\ &= x_i \left(\sum_j \prod_{k\neq j} \left(\mu_k + \sum_i C_{ik} x_i(t) \right) C_{ij}\kappa_j - d_i \prod_k \left(\mu_k + \sum_i C_{ik} x_i(t) \right) \right) \end{aligned}$$

That is, the order of the interactions in this model is the number of resources in the system, $h = m$. As such, models that are routinely used, are actually models with interactions of very high order in disguise.

7 Relaxing assumptions

In this section we separately relax the assumptions on the variance of the GLV coefficients, and on the distribution from which they sampled. In particular, we consider different biologically motivated variance structures for the GLV parameters, and compute the corresponding variance of the polynomial coefficients. In addition, we sample coefficients from the uniform distribution instead. Note that these relaxations yield system of polynomials that are not the KSS system anymore, so we lack theoretical expectations for the outcome.

7.1 Symmetric case

Consider the GLV with symmetric higher order interactions. For simplicity, we start by considering interaciton order $h+1 = 3$. The condition for matrix symmetry is well known; $A = A^T$. The condition for a tensor B of order 3 to be symmetric are

$$B_{ijk} = B_{ikj} = B_{kij} = B_{kji} = B_{jki} = B_{jik}.$$

That is, a symmetric tensor is invariant under a permutation of its arguments. This can be expressed in general as

$$T_{ij_1\ldots j_p} = T_{\sigma(ij_1\ldots j_p)}$$

From Eq. 8 we know that the order a coefficient of the resulting polynomial can be expressed as a function of GLV coefficients. Taking the variance of that expression leads to

$$\begin{aligned}
\mathrm{Var}[a^{(i)}_{l_1\ldots l_n}] &= \mathrm{Var}\left[\sum_{\sigma(\mathcal{S})} T_{ij_1\ldots j_p}\right] \\
&= \sum_{\sigma(\mathcal{S})} \mathrm{Var}\left[T_{ij_1\ldots j_p}\right] + \sum_{j>k} 2\mathrm{Cov}[T_{i\sigma_j}, T_{i\sigma_k}] \\
&= r\mathrm{Var}\left[T_{ij_1\ldots j_p}\right] + r(r-1)\mathrm{Var}[T_{ij_1\ldots j_p}] \\
&= r^2\mathrm{Var}[T_{ij_1\ldots j_p}]
\end{aligned}$$

where the set $\mathcal{S} = \{j_1, \ldots, j_p\}$, and the sum goes over all permutations of the set, and r is the number of different permutations of $\mathcal{S}$ (Eq. 4). Symmetry implies that the random variables $T_{i\sigma(j_1\ldots j_p)}$ are not independent, since they are equal, such that the covariance is not zero. Here σ_j denotes a specific permutation of all the possible permutations in $\sigma(\mathcal{S})$.

7.2 Sparse case

In this case we assume some of the coefficients of the GLV parameters are 0, such that the expression for the variance of the polynomial coefficients takes the form

$$\mathrm{Var}[a^{(i)}_{l_1\dots l_n}] = \mathrm{Var}\left[\sum_{\sigma(\mathcal{S})} T_{ij_1\dots j_p}\right] = sr\mathrm{Var}[T_{ij_1\dots j_p}],$$

where $s \in [0,1]$ is a sparsity parameter that controls what fraction of the $T_{i\sigma(\mathcal{S})}$ coefficients are 0.

7.3 Uniform distribution

In this set of simulations, we sample the parameters from a uniform distribution, with support in $[a,b]$. In order to reduce confounding factors and thus accurately attribute causality, to see what is the effect of changing the probability distribution, we determine a and b such that the first two moments of the uniform distribution match the those of the Gaussian distribution from which we sampled the GLV coefficients in Section 4. For the mean, this leads to the constraint

$$b = -a.$$

For the variance plugging in the above constrain, we obtain

$$\begin{aligned}
\mathrm{Var}[T_{ij_1\dots j_p}] &= \frac{1}{12}(b-a)^2 \\
\frac{h!}{p!(h-p)!} &= \frac{4a^2}{12} \\
a &= \pm\sqrt{\frac{3h!}{p!(h-p)!}}
\end{aligned}$$

As such, the support of the uniform distribution from which the GLV coefficients are sampled depends on the interaction order of the system $h+1$, as well as on the coefficient order p. Note that, one can check that sampling the GLV coefficients in this way ensures that the variance of the coefficients of the resulting polynomial systems is exaclty that of the KSS system (Eq. 4),

$$\begin{aligned}
\mathrm{Var}[a^{(i)}_{l_1\dots l_n}] &= \mathrm{Var}\left[\sum_{\sigma(\mathcal{S})} T_{ij_1\dots j_p}\right] \\
&= r\frac{1}{12}(b-a)^2 \\
&= \frac{1}{3}\frac{p!}{l_1!\cdots l_k!}\frac{3h!}{p!(h-p)!} \\
&= \frac{h!}{l_1!\cdots l_k!(h-p)!} = \mathrm{Var}[a^{(i)}_{l_1\dots l_n}]
\end{aligned} \tag{17}$$

Note that the multiindex of $a^{(i)}$ on the LHS of Eq. 17 runs up to n, $\boldsymbol{l} = (l_1,\dots,l_n)$, while on the R.H.S we have k of those terms. This is just a notational inconvenience, and has no numerical effect, since the missing $n-k$ terms are 0, and $0! = 1$.

References

Abel Niels Henrik. Mémoire sur les equations algébriques, où l'on démontre l'impossibilité de la résolution de l'équation générale du cinquième degré // Oeuvres complètes de Niels Henrik Abel: Nouvelle édition. 1. 2012. 28–33. (Cambridge Library Collection - Mathematics).

Armentano D., Azaïs J-M., Dalmao F., León J. Asymptotic variance of the number of real roots of random polynomial systems // Proceedings of the American Mathematical Society. IX 2018. 146, 12. 5437–5449.

Azaïs Jean-Marc, Wschebor Mario. On the Roots of a Random System of Equations. The Theorem of Shub and Smale and Some Extensions // Foundations of Computational Mathematics. XII 2004. 5, 2. 125–144.

Bézout Étttiene. Théorie générale des équations algébrique. 1779.

Breiding Paul, Timme Sascha. HomotopyContinuation.jl: A Package for Homotopy Continuation in Julia // International Congress on Mathematical Software. 2018. 458–465.

Dalmao Federico. Asymptotic variance and CLT for the number of zeros of Kostlan Shub Smale random polynomials // Comptes Rendus Mathematique. XII 2015. 353, 12. 1141–1145.

Duong Manh Hong, Han The Anh. Random evolutionary games and random polynomials. 2023.

Edelman Alan, Kostlan Eric. How many zeros of a random polynomial are real? // Bulletin of the American Mathematical Society. 1995. 32, 1. 1–37.

Kostlan E. On the Distribution of Roots of Random Polynomials // From Topology to Computation: Proceedings of the Smalefest. New York, NY: Springer US, 1993. 419–431.

Posfai Anna, Taillefumier Thibaud, Wingreen Ned S. Metabolic Trade-Offs Promote Diversity in a Model Ecosystem // Phys. Rev. Lett. Jan 2017. 118. 028103.

Serván Carlos A., Capitán José A., Grilli Jacopo, Morrison Kent E., Allesina Stefano. Coexistence of many species in random ecosystems // Nature Ecology & Evolution. Aug 2018. 2, 8. 1237–1242.

Shub Michael, Smale Steve. Complexity of Bezout's Theorem II Volumes and Probabilities // Computational Algebraic Geometry. 1993a. 267–285.

Shub Michael, Smale Steve. Complexity of Bézout's theorem. I. Geometric aspects // Journal of the American Mathematical Society. 1993b. 6, 2. 459–501.

Smith Hal L., Waltman Paul. The General Chemostat // The Theory of the Chemostat: Dynamics of Microbial Competition. 1995. 28–42. (Cambridge Studies in Mathematical Biology).

Wschebor Mario. On the Kostlan–Shub–Smale model for random polynomial systems. Variance of the number of roots // Journal of Complexity. XII 2005. 21, 6. 773–789.