Constrained optimization as ecological dynamics with applications to random quadratic programming in high dimensions

Pankaj Mehta,1,∗ Wenping Cui,1,2 Ching-Hao Wang,1 and Robert Marsland1

1Physics Department, Boston University, Boston, Massachusetts 02215, USA
2Physics Department, Boston College, Chestnut Hill, Massachusetts 02467, USA

(Dated: September 13, 2018)

Quadratic programming (QP) is a common and important constrained optimization problem. Here, we derive a surprising duality between constrained optimization with inequality constraints – of which QP is a special case – and consumer resource models describing ecological dynamics. Combining this duality with a recent ‘cavity solution’, we analyze high-dimensional, random QP where the optimization function and constraints are drawn randomly. Our theory shows remarkable agreement with numerics and points to a deep connection between optimization, dynamical systems, and ecology.

Optimization as ecological dynamics

We begin by deriving the duality between constrained optimization and ecological dynamics. Consider an optimization problem of the form

\[
\begin{align*}
\text{minimize} \quad & f(R) \\
\text{subject to} \quad & g_i(R) \leq 0, \; i = 1, \ldots, S, \quad (1) \\
& R_\alpha \geq 0, \; \alpha = 1, \ldots, M.
\end{align*}
\]

where the variables being optimized \( R = (R_1, R_2, \ldots, R_M) \) are constrained to be non-negative. We can introduce a ‘generalized’ Lagrange multiplier \( \lambda_i \) for each of the \( S \) inequality constraints in our optimization problem. In terms of the \( \lambda_i \), we can write a set of conditions collectively known as the Karush-Kuhn-Tucker (KKT) conditions that must be satisfied at any local optimum \( R_{\text{min}} \) of our problem [1–3]. We note that for this reason, in the optimization literature the \( \lambda_i \) are often called KKT-multipliers rather than Lagrange multipliers. The KKT conditions are:

\[
\begin{align*}
\text{Stationarity:} \quad & \nabla_R f(R_{\text{min}}) + \sum_j \lambda_j \nabla_R g_j(R_{\text{min}}) = 0 \\
\text{Primal feasibility:} \quad & g_i(R_{\text{min}}) \leq 0 \\
\text{Dual feasibility:} \quad & \lambda_i \geq 0 \\
\text{Complementary slackness:} \quad & \lambda_i (g_i(R_{\text{min}}) - m_i) = 0,
\end{align*}
\]

where the last three conditions must hold for all \( i = 1, \ldots, M \). The KKT conditions have a straightforward and intuitive explanation. At the optimum \( R_{\text{min}} \), either \( g_i(R_{\text{min}}) = 0 \) and the constraint is active \( \lambda_i \geq 0 \), or \( g_i(R_{\text{min}}) \leq 0 \) and the constraint is inactive \( \lambda_i = 0 \). In our problem, the KKT conditions must be supplemented with the additional requirement of positivity \( R_\alpha \geq 0 \).

One can easily show that the four KKT conditions and positivity are also satisfied by the steady states of the following set of differential equations restricted to the
space \( \lambda_i, R_\alpha \geq 0 \):

\[
\frac{d\lambda_i}{dt} = \lambda_i g_i(R) \\
\frac{dR_\alpha}{dt} = [-\partial_{R_\alpha} f(R) - \sum_j \lambda_j \partial_{R_\alpha} g_j(R)]R_\alpha
\]  

(2)

The first of these equations just describes exponential growth of a “species” \( i \) with a resource-dependent “growth rate” \( g_i(R) \). Species with \( g_i(R_{\text{min}}) \leq 0 \) correspond to constraints that are inactive and go extinct in the ecosystem (i.e. \( \lambda_{i\text{min}} = 0 \)), whereas species with \( g_i(R_{\text{min}}) = 0 \) survive at steady state and correspond to active constraints with \( \lambda_{i\text{min}} \neq 0 \) (see Figure 1 for a simple two-dimensional example). The second equation in (2) performs a “generalized gradient descent” on the optimization function \( f(R) + \sum_j \lambda_j g_j(R) \) (note the extra factor of \( R_\alpha \) in our dynamics compared to the usual gradient descent equations). In the context of ecology, these equations describe the dynamics of a set of resources \( \{R_\alpha\} \) produced at a rate \(-\partial_{R_\alpha} f(R)R_\alpha\) and consumed by individuals of species \( j \) at a rate \( \lambda_j \partial_{R_\alpha} g_j(R)R_\alpha \).

This suggests a simple dictionary for constructing systems dual to optimization problems with inequality constraints (see Figure 1). The variables are resources whose dynamics are governed by the gradient of the function being optimized. Each inequality is associated with a species through its corresponding Lagrange (KKT) multiplier. Species that survive in the ecosystem correspond to active constraints whereas species that go extinct correspond to inactive constraints. The steady-state values of the resource and species abundances correspond to the local optimum \( \{\lambda_{j\text{min}}\} \), respectively. Finally, the \( f(R_{\text{min}}) \) are closely related to Lyapunov functions known to exist in the literature for specific choices of resource dynamics [15, 18, 19].

**Ecological duals of Quadratic Programming (QP)**

For the rest of the paper, we focus on QP where the optimization function is quadratic, \( f(R) = \frac{1}{2}R^TQR + b^TR \), with \( Q \) a positive semi-definite matrix, and linear inequality constraints. By going to the eigenbasis of \( Q \), we can always rewrite the QP problem as minimizing a square distance

\[
\text{minimize} \quad \frac{1}{2}||R - K||^2 \\
\text{subject to} \quad \sum_\alpha c_{i\alpha} R_\alpha \leq m_i, \ i = 1, \ldots, S. \quad \text{(3)} \\
R_\alpha \geq 0, \ \alpha = 1, \ldots, M.
\]

Using (2), we can construct the dual ecological model:

\[
\frac{d\lambda_i}{dt} = \lambda_i(\sum_\alpha c_{i\alpha} R_\alpha - m_i) \\
\frac{dR_\alpha}{dt} = R_\alpha(K_\alpha - R_\alpha) - \sum_j \lambda_j c_{j\alpha} R_\alpha.
\]  

(4)

The is the famous MacArthur Consumer Resource Model (MCRM) which was first introduced by Robert MacArthur and Richard Levins in their seminal papers [18, 20] and has played an extremely important role in theoretical ecology [21, 22].

In optimization problems, one often works with the Lagrangian dual of an optimization problem. We show in the appendix that the dual to (3) is just

\[
\text{maximize} \quad \sum_i \lambda_i [\kappa_i - \frac{1}{2} \sum_j \alpha_{ij} \lambda_j] \\
\text{subject to} \quad \lambda_i \geq 0,
\]

with \( \kappa_i = \sum_\alpha K_\alpha(c_{i\alpha} - m_i) \), \( \alpha_{ij} = \sum_\alpha c_{i\alpha} c_{j\alpha} \), and the sum restricted to \( \alpha \) for which \( R_{\alpha\text{min}} \neq 0 \). It is once again straightforward to check that the local minima of this problem are in one-to-one correspondence with steady states of the Generalized Lotka-Volterra Equations (GLVs) of the form:

\[
\frac{d\lambda_i}{dt} = \lambda_i(\kappa_i - \sum_j \alpha_{ij} \lambda_j)
\]  

(6)
As with the primal problem, the species in the GLV have a natural interpretation as Lagrange multipliers enforcing inequality constraints. This GLV can also be directly obtained from the MCRM in (4) in the limit where the resource dynamics are extremely fast by setting $\frac{dR}{dt} = 0$ in the second equation and plugging in the steady-state resource abundances into the first equation [18, 19] (see Appendix). This shows the Lagrangian dual of QP maps to a dynamical system described by a GLV – which itself can be derived from the MCRM which is the dynamical dual to the primal optimization problem!

Random Quadratic Programming (RQP)

Recently, the MCRM was analyzed in the high-dimensional limit where the number of resources and species in the regional species pool is large $(S, M \gg 1)$. In this limit, the resource dynamics were extremely complex, with many resources deviating significantly from their unperturbed values and a large fraction of species in the regional pool going extinct [12]. In terms of the corresponding optimization problem, this suggests that $f(R_{\text{min}})$ will generically be far from zero and many of the constraints will be inactive.

To better understand this, we analyzed Random quadratic programming (RQP) problems in high dimension. In RQP, the parameters in (3) are drawn from random distributions (see Figure 2A). We focus on the case where the $K_\alpha$ and $m_i$ are independent random normal variables drawn from Gaussians with mean $K$ and $m$ and variances $\sigma^2_K$ and $\sigma^2_m$, respectively. The elements of the constraint matrix $c_{i\alpha}$ are also drawn from Gaussians with mean $\mu_c/M$ and variance $\sigma^2_c/M$ [27]. This scaling with $M$ is necessary to ensure that the sum that appears in the inequality constraints in (3) has a good thermodynamic limit when $M, S \to \infty$ with $M/S = \gamma$ held fixed.

We are especially interested in understanding the statistical properties of solutions to the RQP (see Fig. 2A). Among the quantities we examine are the expectation value of the optimized function at the minima $\langle f(R_{\text{min}}) \rangle / M$, the fraction of active constraints, $S^*/S$, the fraction of variables that are non-zero at the optimum, $M^*/M$, as well the first two moments of $R_{\alpha\text{min}}$ and $\lambda_{j\text{min}}$ (see Appendix for details).

It is possible to derive mean-field theory (MFT) for the statistical properties of the optimal solution in the RQP – or correspondingly the steady-states of the MCRM – using the cavity method. The basic idea behind the cavity method is to derive self-consistency equations that relate the optimization problem (ecosystem) with $M + 1$ variables (resources) and $S + 1$ inequality constraints (species) to a problem where a constraint (species) and variable (resource) have been removed: $(M + 1, S + 1) \to (M, S)$ [12]. The need to remove both a constraint and variable is important for keeping all order one terms in the thermodynamic limit [23, 24]. In what follows, we focus on the replica-symmetric solution.

The cavity equation exploits the observations the constraint $\sum_{\alpha=1}^M c_{i\alpha} R_\alpha$ is a sum of many random variables, $c_{i\alpha}$. When $M \gg 1$, due to the law of large numbers, we can model such a sum by a random variable drawn from a Gaussian whose mean and variance involve the statistical quantities described above. Less obvious from the perspective of QP is that we need to introduce a second mean-field quantity $K^{\alpha\text{eff}}_\alpha$ (see Appendix and [12]). After introducing the Lagrange multipliers that enforce the inequality constraints, the optimization function to be minimized takes the form

$$1/2 \|R - K\|^2 + \sum_j \lambda_j (c_{j\alpha} R_\alpha - m_j)$$

$$= 1/2 \sum_\alpha \left\{ R_\alpha [R_\alpha - K^{\alpha\text{eff}}_\alpha(\lambda)] + K_\alpha [K_\alpha - R_\alpha] \right\},$$

where we have defined the mean-field variable

$$K^{\alpha\text{eff}}_\alpha(\lambda) = K_\alpha - \sum_{j=1}^S \lambda_j c_{j\alpha}.$$

Since $K^{\alpha\text{eff}}_\alpha(\lambda)$ is also a sum of many terms containing $c_{i\alpha}$, it can also be approximated as a random variable drawn from a Gaussian whose mean and variance are calculated self-consistently.

The full derivation of the replica symmetric mean-field equations is identical to that in [12] and is given in the
Appendix. The resulting self-consistent mean-field cavity equations can be solved numerically in Mathematica. Figure 2 shows the results of our mean-field equations and comparisons to numerics where we directly optimize the RQP problem over many independent realizations using the CVXOPT package in Python [25]. Notice the remarkable agreement between our MFT and results from direct optimization even for moderate system sizes with $M = 100$. In the Appendix, we show that the cavity solution can also accurately describe the dual MCRM.

Figure 2 also shows that the statistical properties of the QP solutions change as we vary the number of constraints $S$ and the variance of the constraint matrix $c_{i\alpha}$. When $S \ll M$, the expectation value of the optimization function $f(R_{\text{min}})/M$ approaches zero – the minimum for the unconstrained problem. In this limit, the few constraints that are present are also active. As $S/M$ is increased, the fraction of active constraints quickly drops, $f(R_{\text{min}})/M$ quickly increases, after which both quantities reach a plateau where they vary very slowly with $S$. The value of the the plateau depends on $\sigma_c$. Increasing the variance of the constraints results in more active constraints and a larger value of $f(R_{\text{min}})$ at the optimum.

These results about RQP can be naturally understood using ideas from ecology. Intuitively, a smaller $\sigma_c$ means more “redundant” constraints. In ecology, this is the principle of limiting similarity: species with large niche overlaps (similar $c_{i\alpha}$) competitively exclude each other [18–22]. In the language of optimization, this ecological intuition suggests that when constraints are similar enough, only the most stringent of these will be active due to an effective competitive exclusion between constraints. Thus, in RQP competitive exclusion becomes a statement about the geometry of how random planes in high dimension repel each other at the corners of simplices. In all about the geometry of how random planes in high dimension geometrical intuition suggests that when constraints are similar enough, only the most stringent of these will be active due to an effective competitive exclusion between constraints. Thus, in RQP competitive exclusion becomes a statement about the geometry of how random planes in high dimension --- alike systems of common resources. By combining this mapping with a recent ‘cavity solution’ to the MCRM, we constructed a mean-field theory for the statistical properties of RQP that showed remarkable agreement with numerical simulations. Intuitions from ecology suggest that the geometry of constrained optimization can be described using a competitive exclusion between constraints which in our case correspond to random high-dimensional hyperplanes. This work suggests that the deep connection between geometry, ecology, and high-dimensional random ecosystems is a generic property of a large class of generalized consumer resource models [26]. Our works also gives a natural explanation of the existence of Lyapunov functions in these models.

**ACKNOWLEDGMENTS**

The work was supported by NIH NIGMS grant 1R35GM119461, Simons Investigator in the Mathematical Modeling of Living Systems (MMLS) to PM, and the Scialog Program sponsored jointly by Research Corporation for Science Advancement (RCSA) and the Gordon and Betty Moore Foundation.

---

[1] S. Boyd and L. Vandenberghe, *Convex optimization* (Cambridge university press, 2004).
[2] D. P. Bertsekas, *Nonlinear programming* (Athena scientific Belmont, 1999).
[3] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
[4] M. Mezard and A. Montanari, *Information, physics, and computation* (Oxford University Press, 2009).
[5] L. Zdeborová, Acta Physica Slovaca. Reviews and Tutorials 59, 169 (2009).
[6] M. Mézard, G. Parisi, and R. Zecchina, Science 297, 812 (2002).
[7] C. Moore and S. Mertens, *The nature of computation* (OUP Oxford, 2011).
[8] C. K. Fisher and P. Mehta, Proceedings of the National Academy of Sciences 111, 13111 (2014).
[9] D. A. Kessler and N. M. Shnerb, Physical Review E 91, 042705 (2015).
[10] B. Dickens, C. K. Fisher, and P. Mehta, Physical Review E 94, 022423 (2016).
[11] G. Bunin, Physical Review E 95, 042414 (2017).
[12] M. Advani, G. Bunin, and P. Mehta, Journal of Statistical Mechanics: Theory and Experiment 2018, 033406 (2018).
[13] M. Barbier, J.-F. Arnoldi, G. Bunin, and M. Loreau, Proceedings of the National Academy of Sciences 201710352 (2018).
[14] G. Birolì, G. Bunin, and C. Cammarota, New Journal of Physics (2018).
[15] M. Tikhonov and R. Monasson, Physical Review Letters 118, 048103 (2017).
[16] R. Marsland III, W. Cui, J. Goldford, A. Sanchez, K. Korolev, and P. Mehta, arXiv preprint arXiv:1805.12516 (2018).
[17] R. Macarthur and R. Levins, The American Naturalist 101, 377 (1967).
[18] R. MacArthur, Theoretical population biology 1, 1 (1970).
[19] P. Chesson, Theoretical Population Biology 37, 26 (1990), ISSN 0040-5809.
[20] R. Macarthur and R. Levins, The American Naturalist 101, 377 (1967), ISSN 0003-0147.
[21] P. Chesson, Annual review of Ecology and Systematics pp. 343–366 (2000).
[22] D. Tilman, Resource competition and community structure, vol. 17 (Princeton University Press, 1982).
[23] M. Mezard, Journal of Physics A: Mathematical and General 22, 2181 (1989).
[24] M. Ramezanali, P. P. Mitra, and A. M. Sengupta, arXiv preprint arXiv:1509.08995 (2015).
[25] M. Andersen, J. Dahl, and L. Vandenberghe, abel. ee. ucla. edu/cvxopt (2013).
[26] S. Landmann and A. Engel, arXiv preprint arXiv:1806.11358 (2018).
[27] We note that this scaling is slightly different from that in [12] where the elements where chosen to scale with $S$ not $M$. This choice does not change the results, but results in slightly different expressions.
Derivation of Lagrangian dual for QP

In this section, we derive the Lagrangian dual to our primal Quadratic Programming (QP) problem

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2}||R - K||^2 \\
\text{subject to} & \quad \sum_{\alpha} c_{i\alpha} R_{\alpha} \leq m_i, \quad i = 1, \ldots, S. \\
& \quad R_{\alpha} \geq 0, \quad \alpha = 1, \ldots, M.
\end{align*}
\]

(7)

We start by introducing Lagrange (KKT) multipliers \( \lambda_i \) dual to each of the \( S \) constraints and Langrange KKT (multipliers) \( \mu_\alpha \) that enforce positivity. Then, the function to be optimized is

\[
\begin{align*}
\text{maximize} & \quad \min_R \frac{1}{2} \sum_{\alpha} (R_{\alpha}^2 - 2K_\alpha R_{\alpha} + K_\alpha^2) + \sum_{j,\alpha} \lambda_j (c_{j\alpha} R_{\alpha} - m_i) - \mu_\alpha R_{\alpha} \\
\text{subject to} & \quad \lambda_j \geq 0, \quad j = 1, \ldots, S
\end{align*}
\]

(8)

We take the derivative with respect to \( R_{\alpha} \) and note that

\[
R_{\alpha*} = \max[0, K_\alpha - \sum_j c_{j\alpha} \lambda_j]
\]

(9)

where we have used the KKT condition \( \mu_\alpha R_{\alpha*} = 0 \).

Plugging this back into (8), we find that the function to be maximized with respect to the \( \lambda_i \) is

\[
\sum_i \lambda_i [\kappa_i - \frac{1}{2} \sum_j \alpha_{ij} \lambda_j]
\]

(10)

with

\[
\kappa_i = \sum_{\alpha, R_{\alpha*} \neq 0} K_\alpha c_{i\alpha} - m_i
\]

(11)

and

\[
\alpha_{ij} = \sum_{\alpha, R_{\alpha*} \neq 0} c_{i\alpha} c_{j\alpha}.
\]

(12)

Derivation of Lotka Volterra Equations form MCRM

We start from the MCRM dynamical equations

\[
\begin{align*}
\frac{d\lambda_i}{dt} & = \lambda_i (\sum_{\alpha} c_{i\alpha} R_{\alpha} - m_i) \\
\frac{dR_{\alpha}}{dt} & = R_{\alpha} [(K_\alpha - R_{\alpha}) - \sum_j \lambda_j c_{j\alpha}] R_{\alpha}.
\end{align*}
\]

(13)

Notice that setting the second equation to zero we get

\[
R_{\alpha*} = \max[0, K_\alpha - \sum_j c_{j\alpha} \lambda_j].
\]

(14)

Plugging this into the first equation in (13) gives

\[
\frac{d\lambda_i}{dt} = \lambda_i (\kappa_i - \sum_j \alpha_{ij} \lambda_j)
\]

(15)

with \( \alpha_{ij} \) and \( \kappa_i \) defined as in the last appendix.
In this section, we supplement Figure 2 in main text with an additional figure showing a comparison of the Cavity solution, optimization of RQP, and steady-state values of the MCRM dual to the RQP. For each choice of parameters, the RQP were solved using the CVXOPT package in Python 3. The dual MCRM was constructed as outlined in main text and then integrated to steady-state using standard ODE solvers in Python. See supplementar

FIG. 3: Comparison of Cavity Solution (solid line), RQP (long dash line), and dual MCRMs (short dash line). The simulations represent averages from 50 independent realizations and parameters as in Figure 2 of main text.

Derivation of cavity solution

Model setup

In this section, we derive the cavity solution to the MCRM (Eq. (4) in the main text)

\[
\begin{align*}
\frac{d\lambda_i}{dt} &= \lambda_i \left( \sum_\alpha c_{i\alpha} R_\alpha - m_i \right) \\
\frac{dR_\alpha}{dt} &= R_\alpha (K_\alpha - R_\alpha) - \sum_j \lambda_j c_{j\alpha} R_\alpha.
\end{align*}
\]  

(16)

Note that here we follow closely the derivation in [12]. The only difference is that here we consider the consumer preference \( c_{i\alpha} \) as random variables drawn from a Gaussian distribution with mean \( \mu_c/M \) and variance \( \sigma_c^2/M \), as opposed to the choices \( \mu_c/S \) and \( \sigma_c^2/S \) used in that work. With these definitions, we can decompose the consumer preference into \( c_{i\alpha} = \mu_c/M + \sigma_c d_{i\alpha} \), where the fluctuating part \( d_{i\alpha} \) obeys

\[
\langle d_{i\alpha} \rangle = 0 \quad \langle d_{i\beta} d_{j\beta} \rangle = \frac{\delta_{ij} \delta_{\alpha\beta}}{M}.
\]

(17)
\[\text{(18)}\]

We also assume that both the carrying capacity \( K_\alpha \) and the minimum maintenance cost \( m_i \) are independent Gaussian random variables with mean and covariance given by

\[
\begin{align*}
\langle K_\alpha \rangle &= K \\
\text{Cov}(K_\alpha, K_\beta) &= \delta_{\alpha\beta} \sigma_K^2 \\
\langle m_i \rangle &= m \\
\text{Cov}(m_i, m_j) &= \delta_{ij} \sigma_m^2.
\end{align*}
\]

(19)
\[\text{(20)}\]
\[\text{(21)}\]
\[\text{(22)}\]
Let \( \langle R \rangle = \frac{1}{M} \sum R_\alpha \) and \( \langle \lambda \rangle = \frac{1}{S} \sum \lambda_i \) be the average resource and average species abundance, respectively. With all these defined, we can re-write Eq. (16) as

\[
\frac{d\lambda_i}{dt} = \lambda_i \left\{ \mu_c \langle R \rangle - m_i + \sigma_c \sum_\alpha d_{i\alpha} R_\alpha - \delta m_i \right\} \tag{23}
\]

\[
\frac{dR_\alpha}{dt} = R_\alpha \left\{ [K - \mu_c \gamma^{-1} \langle \lambda \rangle] - R_\alpha - \sigma_c \sum_j d_{j\alpha} \lambda_j + \delta K_\alpha \right\} , \tag{24}
\]

where \( \delta K_\alpha = K_\alpha - K, \delta m_i = m_i - m \) and \( \lambda = M/S \). We can interpret the bracketed terms in these equations as population mean growth rate and effective resource capacity, respectively, viz.

\[
g \equiv \mu_c \langle R \rangle - m \tag{25}
\]

\[
K^{\text{eff}} \equiv K - \mu_c \gamma^{-1} \langle \lambda \rangle . \tag{26}
\]

As noted in the main text, the basic idea of cavity method is to relate an ecosystem with \( M + 1 \) resources (variables) and \( S + 1 \) species (inequality constraints) to that with \( M \) resources and \( S \) species. Following Eq.(23)(24), one can write down the ecological model for the \((M + 1, S + 1)\) system where resource \( R_0 \) and species \( \lambda_0 \) are introduced to the \((M, S)\) system as:

\[
\frac{d\lambda_i}{dt} = \lambda_i \left\{ g + \sigma_c \sum_\alpha d_{i\alpha} R_\alpha + \sigma_c d_{i0} R_0 - \delta m_i \right\} \tag{27}
\]

\[
\frac{dR_\alpha}{dt} = R_\alpha \left\{ K^{\text{eff}} - R_\alpha - \sigma_c \sum_j d_{j\alpha} \lambda_j - \sigma_c d_{0\alpha} \lambda_0 + \delta K_\alpha \right\} , \tag{28}
\]

where all sums from now on are understood to be over the indices \( \alpha, j > 0 \) from the \((M, S)\) system. The equations for the newly introduced species \((i = 0)\) and resource \((\alpha = 0)\) are given by

\[
\frac{d\lambda_0}{dt} = \lambda_0 \left\{ g + \sigma_c \sum_\alpha d_{0\alpha} R_\alpha + \sigma_c d_{00} R_0 - \delta m_0 \right\} \tag{29}
\]

\[
\frac{dR_0}{dt} = R_0 \left\{ K^{\text{eff}} - R_0 - \sigma_c \sum_j d_{j0} \lambda_j - \sigma_c d_{00} \lambda_0 + \delta K_0 \right\} , \tag{30}
\]

**Deriving the self-consistency equations with cavity method**

Following the same procedure in [12], we introduce the following susceptibilities:

\[
\chi^{(\lambda)}_{i\beta} = \frac{\partial \lambda_i}{\partial K_\beta} \tag{31}
\]

\[
\chi^{(R)}_{\alpha\beta} = \frac{\partial R_\alpha}{\partial K_\beta} \tag{32}
\]

\[
\nu^{(\lambda)}_{ij} = \frac{\partial \lambda_i}{\partial m_j} \tag{33}
\]

\[
\nu^{(R)}_{\alpha j} = \frac{\partial R_\alpha}{\partial m_j} \tag{34}
\]
where we denote $\overline{X}$ as the steady-state value of $X$. Recall that the goal is to derive a set of self-consistency equations that relates the ecological system (optimization problem) characterized by $M+1$ resources (variables) and $S+1$ species (constraints) to that with the new species and new resources removed: $(S+1, M+1) \rightarrow (S, M)$. To simplify notation, denote $\overline{X}_{\lambda 0}$ be the steady-state value of quantity $X$ in the absence of the new resource and new species. Since the introduction of a new species and resource represents only a small (order 1/$M$) perturbation to the original ecological system, we can express the steady-state species and resource abundances in the $(S,M)$ values. We note that the new terms $\sigma_c d_{0a} R_0$ in Eq. (27) and $\sigma_c d_{0a} \lambda_0$ in Eq. (28) can be treated as perturbations to $m_i$, and $K_i$, respectively, yielding:

$$\lambda_i = \overline{\lambda}_{\lambda 0} - \sigma_c \sum_{\beta} \lambda^{(\beta)} d_{0\beta} \overline{\lambda}_0 - \sigma_c \sum_j \nu^{(\beta)}_{ij} d_{j0} R_0$$  \hspace{1cm} (35)$$

$$R_{\alpha} = \overline{R}_{\alpha \lambda 0} - \sigma_c \sum_{\beta} \lambda^{(R)} d_{0\beta} \overline{\lambda}_0 - \sigma_c \sum_j \nu^{(R)}_{\alpha j} d_{j0} R_0.$$  \hspace{1cm} (36)$$

The next step is to plug Eq. (35) (36) into Eq. (29) (30) and solve for the steady-state value of $\lambda_0$ and $R_0$. For the new species, setting Eq. (29) to zero and plugging in Eq. (36) gives:

$$0 = \overline{\lambda}_0 \left[ g + \sigma_c \sum_{\alpha} d_{0a} \overline{R}_{\alpha \lambda 0} - \sigma_c^2 \sum_{\beta} \lambda^{(R)} d_{0\beta} \overline{\lambda}_0 - \sigma_c^2 \sum_{\alpha j} \nu^{(R)}_{\alpha j} d_{j0} R_0 - \delta m_0 + \sigma_c d_{00} R_0 \right].$$  \hspace{1cm} (37)$$

We now note that each of the sums in this equation is the sum over a large number of uncorrelated random variables, and can therefore be well approximated by Gaussian random variables for large enough $M$ and $S$. It is a straightforward exercise to show that the mean and variance of the third sum as well as the variance of the second sum are all order 1/$M$ or higher, and can be ignored in comparison to the order 1 terms. The mean of the second sum is:

$$\sum_{\alpha \beta} \langle \lambda^{(R)}_{\alpha \beta} \rangle (d_{0a} d_{0\beta}) = \frac{1}{M} \sum_{\alpha} \langle \lambda^{(R)}_{\alpha \alpha} \rangle = \chi$$  \hspace{1cm} (38)$$

where we have used the statistics of $d_{i\alpha}$ as defined in Eqs. (17) (18), and have defined $\chi \equiv \langle \lambda^{(R)}_{i\alpha} \rangle$.

Using these observations about the second and third sums, we obtain:

$$0 = \overline{\lambda}_0 \left[ g - \sigma_c^2 \chi \overline{\lambda}_0 + \sigma_c \sum_{\alpha} d_{0a} \overline{R}_{\alpha \lambda 0} - \delta m_0 \right] + O(M^{-1/2}),$$  \hspace{1cm} (39)$$

Since the $m_i$ come from a Gaussian distribution, we can model the combination of the remaining sum with $\delta m_i$ by a single Gaussian random variable with zero mean and variance $\sigma_g^2$ given by:

$$\sigma_g^2 \equiv \text{Var} \left( \sigma_c \sum_{\alpha} d_{0a} \overline{R}_{\alpha \lambda 0} - \delta m_0 \right)$$  \hspace{1cm} (40)$$

$$= \text{Var} \left( \sigma_c \sum_{\alpha} d_{0a} \overline{R}_{\alpha \lambda 0} \right) + \text{Var} (\delta m_0)$$  \hspace{1cm} (41)$$

$$= \sigma_c^2 \frac{1}{M} \sum_{\alpha} \overline{R}_{\alpha \lambda 0}^2 + \sigma_m^2$$  \hspace{1cm} (42)$$

$$= \sigma_c^2 q_R + \sigma_m^2,$$  \hspace{1cm} (43)$$

where

$$q_R = \frac{1}{M} \sum_{\alpha} \overline{R}_{\alpha \lambda 0}^2.$$  \hspace{1cm} (44)$$

Denoting $z_\lambda$ as a random variable with zero mean and unit variance, we can express Eq. (39) in terms of the quantities just defined:

$$0 = \overline{\lambda}_0 \left( g - \sigma_c^2 \chi \overline{\lambda}_0 + \sigma_g z_\lambda \right).$$  \hspace{1cm} (45)$$
Inverting this equation one gets

\[ \bar{\lambda}_0 = \frac{\max[0, g + \sigma_g z_\lambda]}{\sigma^2 \chi}, \] (46)

which is a truncated Gaussian.

We can follow the same procedure to solve for the steady state of the resource. Setting Eq.(30) to zero and plugging in Eq.(35) gives

\[ 0 = \overline{R}_0 \left( K^{\text{eff}} - \overline{R}_0 - \sigma \sum_j d_{j0} \overline{\lambda}_{j,0} + \sigma^2 \sum_j \chi_j^{(\lambda)} d_{j0} d_{0j} \overline{\lambda}_0 + \sigma^2 \sum \nu_{jk} d_{j0} d_{k0} \overline{R}_0 + \delta K_0 - \sigma c d_{00} \overline{\lambda}_0 \right). \] (47)

Keeping only the leading order terms one arrives at

\[ 0 \approx \overline{R}_0 \left( K^{\text{eff}} - \overline{R}_0 + \delta K_0 - \sigma \sum_j d_{j0} \overline{\lambda}_{j,0} + \sigma^2 \gamma^{-1} \nu \overline{R}_0 \right). \] (48)

where \( \nu \equiv \langle \nu_{jj}^{(\lambda)} \rangle \) is the average susceptibility. As before, \( \delta K_0 - \sigma \sum_j d_{j0} \overline{\lambda}_{j,0} \) is a Gaussian random variable with zero mean and variance \( \sigma^2 K^{\text{eff}} \) given by

\[ \sigma^2 K^{\text{eff}} = \text{Var}\left( \delta K_0 - \sigma \sum_j d_{j0} \overline{\lambda}_{j,0} \right) \] (49)

\[ = \text{Var}(\delta K_0) + \text{Var}\left(\sigma \sum_j d_{j0} \overline{\lambda}_{j,0}\right) \] (50)

\[ = \sigma^2_K + \sigma^2 c \frac{1}{M} \sum \overline{\lambda}_{j,0}^2 \] (51)

\[ = \sigma^2_K + \sigma^2 c \gamma^{-1} q_\lambda, \] (52)

where

\[ q_\lambda = \frac{1}{S} \sum_j \overline{\lambda}_{j,0}^2. \] (53)

Denoting \( z_R \) as a random variable with zero mean and unit variance, we can express Eq.(48) in terms of the quantities just defined:

\[ 0 = \overline{R}_0 \left( K^{\text{eff}} - \overline{R}_0 + \sigma K^{\text{eff}} z_R + \sigma^2 \gamma^{-1} \nu \overline{R}_0 \right). \] (54)

Finally, inverting this equation gives the steady-state distribution of the resource

\[ \overline{R}_0 = \frac{\max(0, K^{\text{eff}} + \sigma K^{\text{eff}} z_R)}{1 - \gamma^{-1} \sigma^2 c \nu} \] (55)

Next let’s examine the self-consistency equations for the fraction of non-zero species and resources, \( \phi_\lambda \) and \( \phi_R \), respectively. Note that the goal is to find the values of \( \{\phi_\lambda, \phi_R, \langle \lambda \rangle, \langle R \rangle, q_\lambda, q_R, \chi, \nu\} \) with given sets of parameters \( \{K, \sigma_K, m, \sigma_m, \mu_c, S, M\} \). By variable counting, we’ll need eight equations to solve for these eight unknowns but so far we’ve only got two, Eq.(46) and Eq.(55). To find the remaining six equations, let’s define some quantities (c.f. Eq.(25)(26)):

\[ \Delta_g \equiv \frac{g}{\sigma_g} = \frac{\mu_c \langle R \rangle - m}{\sigma_g}, \] (56)

\[ \Delta_{K^{\text{eff}}} \equiv \frac{K^{\text{eff}}}{\sigma_{K^{\text{eff}}}} = \frac{K - \mu_c \gamma^{-1} \langle \lambda \rangle}{\sigma_{K^{\text{eff}}}}, \] (57)
as well as the function

$$w_j(\Delta) = \int_{-\Delta}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} (z + \Delta)^j,$$  \hspace{1cm} (58)$$

which will simplify our notation later. First let’s derive the self-consistency equation for the susceptibilities. This is done by taking the derivative of Eq.(55) with respect to $K$ and of Eq.(46) with respect to $m$ while noting the definition of $\phi_\lambda$ and $\phi_R$:

$$\nu = -\frac{\phi_\lambda}{\sigma^2_\chi},$$  \hspace{1cm} (59)$$

$$\chi = \frac{\phi_R}{1 - \gamma^{-1}\sigma^2_\nu}.$$  \hspace{1cm} (60)$$

Since Eq.(46) and Eq.(55) imply that the species and resource distributions are truncated Gaussians, it will be useful to note the following:

Let $y = \max(0, \frac{a}{b} + \frac{c}{b}z)$, with $z$ being a Gaussian random variable with zero mean and unit variance. Then its $j$-th moment is given by

$$\langle y^j \rangle = \left( \frac{b}{c} \right)^j \int_{-\frac{b}{a}}^{\infty} dz \sqrt{2\pi} e^{-\frac{z^2}{2}} (z + \frac{b}{a})^j.$$  \hspace{1cm} (61)$$

With this we can easily write down the self-consistency equations for the fraction of non-zero species and resources as well as the moments of their abundances (c.f. Eq.(46) and Eq.(55)):

$$\phi_\lambda = w_0(\Delta_g)$$  \hspace{1cm} (62)$$

$$\phi_R = w_0(\Delta_{K^{\text{eff}}})$$  \hspace{1cm} (63)$$

$$\langle \lambda \rangle = \frac{\sigma_g}{\sigma^2_\chi} w_1(\Delta_g)$$  \hspace{1cm} (64)$$

$$\langle R \rangle = \frac{\sigma_{K^{\text{eff}}}}{1 - \gamma^{-1}\sigma^2_\nu} w_1(\Delta_{K^{\text{eff}}})$$  \hspace{1cm} (65)$$

$$q_\lambda = \langle \lambda^2 \rangle = \left( \frac{\sigma_g}{\sigma^2_\chi} \right)^2 w_2(\Delta_g)$$  \hspace{1cm} (66)$$

$$q_r = \langle R^2 \rangle = \left( \frac{\sigma_{K^{\text{eff}}}}{1 - \gamma^{-1}\sigma^2_\nu} \right)^2 w_2(\Delta_{K^{\text{eff}}}).$$  \hspace{1cm} (67)$$

Note that we only write down the first and the second moments since these six equations, along with Eq.(46) and Eq.(55), complete the equations required to solve for the eight variables.

Cavity solution to the optimization function

Here we derive the cavity solution to the optimization function $f(R)$ defined as

$$\langle f(R) \rangle = \frac{1}{2} \langle ||R - K||^2 \rangle$$  \hspace{1cm} (68)$$

$$= \frac{1}{2} \sum_\alpha \langle R^2_\alpha \rangle - 2 \langle K_\alpha R_\alpha \rangle + \langle K^2_\alpha \rangle.$$  \hspace{1cm} (69)$$

The first term is given by Eq.(67) while the last term is just $K^2 + \sigma^2_K$. What remains to be solved is $\langle K_\alpha R_\alpha \rangle$. From Eq.(55), one can write

$$R_\alpha(K_\alpha) = \frac{\max(0, K_\alpha - \mu \gamma^{-1} \langle \lambda \rangle + z_\lambda \sqrt{\sigma^2_\chi \gamma^{-1} q_\lambda})}{1 - \gamma^{-1}\sigma^2_\nu}.$$  \hspace{1cm} (70)$$
Now let variable $k$ be drawn from the same distribution as $K_{\alpha}$, namely, Gaussian with mean $K$ and variance $\sigma_{K}^{2}$, one gets

$$R(k) = \frac{\max(0, k - \mu_c \gamma^{-1}(\lambda) + z_\lambda \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda})}{1 - \gamma^{-1} \sigma^{2}_\nu}.\quad (71)$$

Therefore, we compute

$$
\langle kR(k) \rangle_{z_\lambda, k} = \frac{1}{\sqrt{2\pi}} \left( \int dk kR(k)e^{-\frac{(k-K)^{2}}{2\delta}} \right)_{z_\lambda} \quad (72)
$$

$$
= \frac{1}{1 - \gamma^{-1} \sigma^{2}_\nu} \left( \int_{-\infty}^{\infty} dk k \int_{-\infty}^{\infty} d\sigma \left[ 0, k - \mu_c \gamma^{-1}(\lambda) + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right] e^{-\frac{(k-K)^{2}}{2\delta}} \right)_{z_\lambda} \quad (73)
$$

$$
= \frac{1}{1 - \gamma^{-1} \sigma^{2}_\nu} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dk d\sigma}{2\pi \sqrt{\sigma_K}} k \max \left[ 0, k - \mu_c \gamma^{-1}(\lambda) + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right] e^{-\frac{(k-K)^{2}}{2\delta}} e^{-\frac{\sigma^{2}_\nu}{2}} \quad (74)
$$

To simplify the calculation, let us introduce another Gaussian variable $z_K$ with zero mean and unit variance. The integral part can now be written as:

$$
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dz_K d\lambda}{2\pi} e^{-\frac{z_K^2 + \lambda^2}{2}} (K + \sigma_K z_K) \max \left[ 0, K + \sigma_K z_K - \mu_c \gamma^{-1}(\lambda) + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right] \quad (75)
$$

$$
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dz_K d\lambda}{2\pi} e^{-\frac{z_K^2 + \lambda^2}{2}} K \max \left[ 0, K - \mu_c \gamma^{-1}(\lambda) + \sigma_K z_K + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right] \quad (76)
$$

+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dz_K d\lambda}{2\pi} e^{-\frac{z_K^2 + \lambda^2}{2}} \sigma_K z_K \max \left[ 0, K - \mu_c \gamma^{-1}(\lambda) + \sigma_K z_K + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right]

Using $z_R \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} = \sigma_K z_K + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} z_\lambda$, the first term of Equation (76) can be written as

$$
\int_{-\infty}^{\infty} \frac{dz_R}{\sqrt{2\pi}} e^{-\frac{z_R^2}{2}} K \max \left[ 0, K - \mu_c \gamma^{-1}(\lambda) + z_R \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right] = \sqrt{\sigma_{K}^{2} + \sigma_{c}^{2} \gamma^{-1} q_\lambda \Delta}, \quad (77)
$$

where

$$
\Delta = \frac{K - \mu_c \gamma^{-1}(\lambda)}{\sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda}}. \quad (78)
$$

Using integration by parts in the $z_K$ integral, we find that the second term of Equation (76) is

$$
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{z_K^2 + \lambda^2}{2}} d\sigma_K z_K \max \left[ 0, K + \sigma_K z_K - \mu_c \gamma^{-1}(\lambda) + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right] \quad (79)
$$

where $\Theta(x)$ equals 0 for $x < 0$, and equals 1 for $x \geq 0$. It arises from taking the derivative of

$$
\max \left[ 0, K + \sigma_K z_K - \mu_c \gamma^{-1}(\lambda) + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right]
$$

with respect to $z_K$ in the integration by parts. As in the first integral, we can now change variables to $z_R$, and use the $\Theta$ function to set the lower limit of integration:

$$
\sigma_{K}^{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{z_K^2 + \lambda^2}{2}} d\sigma_K d\lambda \Theta \left( K + \sigma_K z_K - \mu_c \gamma^{-1}(\lambda) + \sqrt{\sigma_{c}^{2} \gamma^{-1} q_\lambda} \right) \quad (80)
$$

$$
= \sigma_{K}^{2} \int_{-\Delta}^{\infty} e^{-\frac{z_R^2}{2}} d\sigma_K \Delta \quad (81)
$$

where $\Delta$ is the same quantity defined in Equation (78) above.

Putting Equations (77) and (81) back into Equation (74), we finally find:

$$
\langle kR(k) \rangle_{z_\lambda, k} = \frac{1}{\sqrt{2\pi}} \frac{\sigma_{K}^{2} w_0(\Delta) + \sqrt{\sigma_{K}^{2} + \sigma_{c}^{2} \gamma^{-1} q_\lambda \Delta}}{\sqrt{2\pi}}. \quad (82)
$$