The Minimum Environmental Perturbation Principle: A New Perspective on Niche Theory

Robert Marsland III[1,*]

Wenping Cui[1,2]

Pankaj Mehta[1]

1. Boston University, Boston, Massachusetts 02215;

2. Boston College, Chestnut Hill, Massachusetts 02467;

∗ Corresponding author; e-mail: marsland@bu.edu.

Abstract

Contemporary niche theory is a powerful conceptual framework for understanding how organisms interact with each other and with their shared environment. Here we show that, for a wide range of modeling assumptions, niche theory is equivalent to a Minimum Environmental Perturbation Principle (MEPP): ecosystems self-organize into a state that minimizes the collective impact of organisms on their environment. Different choices of environmental dynamics naturally give rise to distinct dissimilarity measures for quantifying environmental impact. The MEPP allows for the analysis of ecosystems with large numbers of species and environmental factors and provides a new avenue for analyzing ecological invasions. We show that the presence of environmental feedbacks, where organisms can produce new resources in addition to depleting them, violates the global MEPP. However, even in the presence of such feedbacks, a weaker, local version of the MEPP still applies in a limited region of resource space. We show that the MEPP framework is consistent with classic experiments on competition for substitutable resources in herbivorous zooplankton and propose new experiments for testing MEPP directly.

Introduction

The concept of a "niche" has long been central to ecological theory. Over the past fifty years, this concept has been significantly clarified and refined, beginning with the pioneering work of Levins, MacArthur and Tilman, and more recently consolidated by Chase and Leibold (Chase and Leibold 2003; Leibold 1995; MacArthur 1969, 1970; Tilman 1982). As illustrated in fig. 1*(a)*, this framework highlights two distinct aspects of an organism's niche: its requirements for survival and its impact on the environment (Leibold 1995). The state of the environment is mathematically represented by a vector $\mathbf{R}$ in an abstract space, whose components R_α ($\alpha = 1, 2 \ldots M$) correspond to concentrations of resources, populations of predators, stress intensity, or any other factors that affect an organism's growth rate (Levin 1970).

In this space, the "requirement niche" of each species i ($i = 1, 2 \ldots S$) can be encoded by a zero-net-growth isocline (ZNGI) (Leibold 1995; Tilman 1982). The ZNGI is a hypersurface that separates the environmental states where the growth rate $g_i(\mathbf{R})$ is positive from the states where it is negative. Environmental states along the ZNGI support reproduction rates that exactly balance death or dilution rates, leading to constant population sizes. The ZNGI thus represents one boundary of Hutchinson's "fundamental niche," which is the M-dimensional hypervolume of environmental conditions in which the species can survive indefinitely (Chase and Leibold 2003; Hutchinson 1957; Leibold 1995).

The "impact niche" is represented by an impact vector $\mathbf{q}_i(\mathbf{R})$, which specifies the magnitude and direction of the environmental change induced by a single individual of the species (Leibold 1995; Tilman 1982). This vector includes impacts on resource abundances, predator populations, pH, oxygen/$C0_2$ levels, and any other environmental factors that can be affected by the presence or activity of the organism. The impact vector thus encodes the "role" of an organism in the community, and roughly corresponds to Elton's niche concept (Elton 1927; Leibold 1995).

The environment also has its own dynamics, produced by external supplies of abiotic nutrients, regrowth of biotic resources like grasses or insects, natural death of predators, or any other

environmental process that is independent of the organisms whose niches are being analyzed. These intrinsic dynamics can be represented by a "supply vector" $\mathbf{h}(\mathbf{R})$ (Chase and Leibold 2003; Tilman 1982). For a given set of species defined by their ZNGI's and impact vectors, it is the supply vector that determines which species can survive, which must go extinct, and which ones stably coexist.

As long as all ecologically relevant interactions among organisms are mediated by the environment, these three quantities are sufficient to determine community properties (Chase and Leibold 2003). In particular, they define the conditions for stable coexistence, which can be represented graphically in the case of two environmental factors $M = 2$, as illustrated in Figure 1*(a)*. The graphical approach facilitates a complete enumeration of possible community structures in terms of the relative positions of the ZNGI's, impact vectors and supply vectors (Chase and Leibold 2003).

As the dimension of $\mathbf{R}$ grows, however, the space becomes harder to visualize, and the number of possibilities grows exponentially. The high-dimensional situation is especially complicated by the fact that the two aspects of the niche – requirements and impacts – can in principle be varied independently (Leibold 1995). In many situations, however, it is natural to assume that the requirement and impact niches of a given organism are related, so that knowledge of one immediately determines the other. In this article, we show that this assumption leads to a powerful new principle for finding uninvadable equilibrium states of ecosystems at any level of environmental complexity. As shown in table 2, our derivation is applicable to a wide variety of ecological models including those with biotic and abiotic resource dynamics, substitutable or essential resources, linear or saturating response functions, environmental feedbacks, and predation- and/or resource-limited ecosystems.

The ZNGI's, supply vectors and impact vectors all play a role in the construction of the optimization principle, as illustrated in Figure 1*(b)*. For a given regional species pool, the ZNGI's fix the boundaries of the uninvadable region Ω, which is the region of resource space where per-capita growth rates g_i of all species are either zero or negative. Outside of this region, the growth

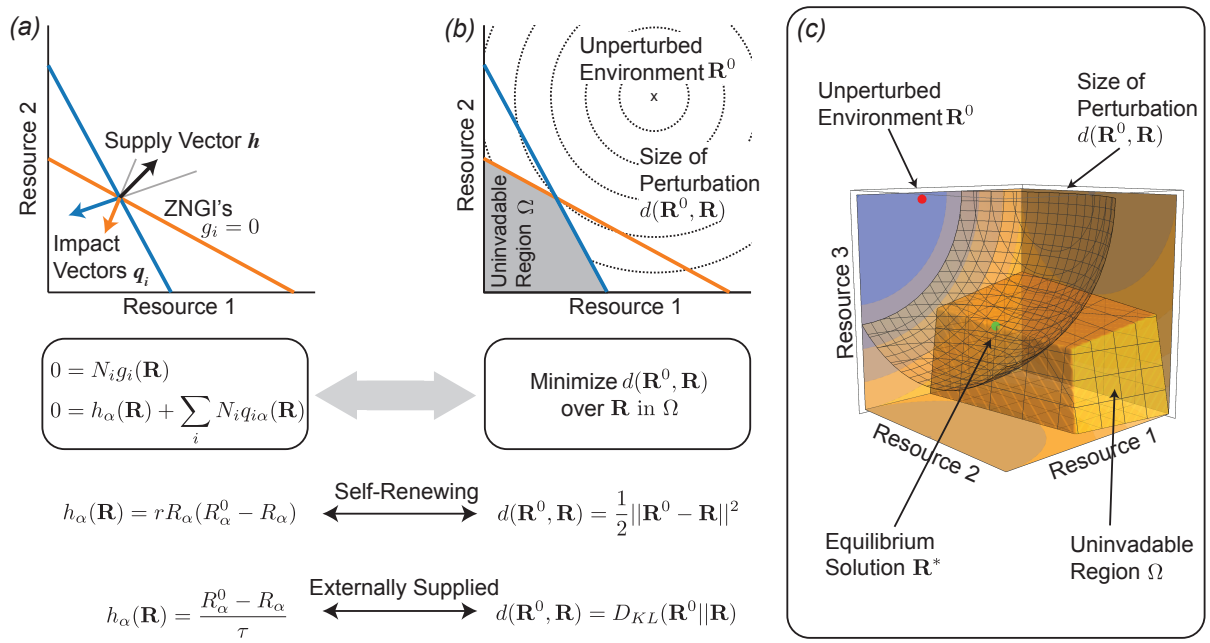

Figure 1: **Reinterpreting contemporary niche theory.** *(a)* Contemporary niche theory provides graphical and mathematical statements of equilibrium criteria in terms of the *impact vectors* (colored arrows) and *zero net-growth isoclines* (ZNGI's, colored lines) of each consumer species, and a *supply vector* (black arrow) that encodes the intrinsic dynamics of the environment. Two standard choices for the supply vector are logistic growth of self-renewing resources, and a linear model of externally supplied resource fluxes. *(b)* The Minimum Environmental Perturbation Principle (MEPP) provides an alternative procedure for finding uninvadable equilibrium states, by minimizing a dissimilarity measure $d(\mathbf{R}^0, \mathbf{R})$ (dotted contour lines) of the size of the environmental perturbation caused by the consumers. The minimization is performed over resource concentration vectors $\mathbf{R}$ within the shaded uninvadable region Ω where all the per-capita growth rates $g_i(\mathbf{R})$ are negative or zero. Different supply vector fields generate different measures $d(\mathbf{R}^0, \mathbf{R})$, with self-renewing and externally supplied resource dynamics giving rise to the Euclidean distance and Kullback-Leibler divergence respectively. *(c)* MEPP readily generalizes to larger numbers of resources. The orange box is the uninvadable region, and the circles projected on the axes are contours of the perturbation measure $d(\mathbf{R}^0, \mathbf{R})$. The contour with the smallest value of $d(\mathbf{R}^0, \mathbf{R})$ that still intersects the uninvadable region is shown as a curved mesh surface.

rate of at least one species from the regional pool is positive, guaranteeing that the system is susceptible to invasions.

Any point in the uninvadable region can be made into an uninvadable equilibrium state $\mathbf{R}^*$ of the ecological dynamics with a suitable choice of the supply vector field $\mathbf{h}(\mathbf{R})$. If the environment's own equilibrium state $\mathbf{R}^0$ lies in the interior of the uninvadable region, then all species in the regional pool will go extinct, and $\mathbf{R}^0 = \mathbf{R}^*$ is itself the uninvadable equilibrium state. If $\mathbf{R}^0$ is outside this region, however, $\mathbf{R}^*$ will lie somewhere on the boundary, and at least one consumer species will survive. The surviving species perturb the environmental state away from $\mathbf{R}^0$, and the magnitude of this perturbation can be quantified in terms of Euclidean distance, Kullback-Leibler divergence, or any number of other dissimilarity measures $d(\mathbf{R}^0, \mathbf{R})$.

To determine which point on the boundary is chosen by the ecological dynamics as the equilibrium state – where reproduction exactly balances death or dilution, and supply exactly balances impacts – we also need to know the impact vectors $\mathbf{q}_i$ of the surviving species. In many applications of niche theory, one is primarily interested in the conditions for coexistence of a set of similar species, such as the different species of warblers in MacArthur's original field work (MacArthur 1958), or the paramecia of Gause's competitive exclusion experiments (Gause and Witt 1935), with all other organisms treated as components of the environment. The functional similarity of the competitors justifies the assumption of a universal relationship between the ZNGI's and the impact vectors, such that for any new species, the direction of the impact vector can be inferred from the ZNGI.

In the case of resource competition, for example, if all the competing species are sufficiently similar in digestive and metabolic capacities, it is reasonable to assign a fixed set of relative nutritional values to different types of substitutable resources. If one warbler species is observed to require a larger quantity of crawling insects than another to substitute for a given reduction in the availability of flying insects, it can then be inferred that the relative foraging abilities of the two species for the two kinds of insects (which determine the impact vectors) differ in the same way.

This assumption significantly simplifies the analysis even of two-resource ecosystems, and is present in much of the early work on consumer-resource models (Chesson 1990; MacArthur 1969, 1970). Later studies showed that strong violations of this postulate lead to limit cycles, heteroclinic cycles and chaos when the number of resources is sufficiently large (Huisman and Weissing 1999, 2001*a*,*b*). The recent microbiome-related resurgence of interest in consumer-resource models has almost exclusively focused on cases where a strict ZNGI/impact relationship is maintained (Advani et al. 2018; Butler and O'Dwyer 2018; Posfai et al. 2017; Tikhonov and Monasson 2017, 2018).

Our main result is a mathematical proof that this same assumption converts niche theory into an optimization problem, as illustrated in Figure 1*(b)*. With the impact vectors no longer constituting an independent set of parameters, uninvadable equilibrium states can be found from the ZNGI's and the supply vectors alone. The Minimum Environmental Perturbation Principle (MEPP) then says that out of all the uninvadable environmental states, the equilibrium state $\mathbf{R}^*$ locally minimizes the magnitude of the environmental perturbation $d(\mathbf{R}^0, \mathbf{R})$ away from the intrinsic environmental equilibrium $\mathbf{R}^0$.

Methods

All simulations and data analysis were performed in Python using the Scipy scientific computing package (Jones et al. 2001–). Data and scripts to generate all figures can be downloaded from https://github.com/Emergent-Behaviors-in-Biology/mepp.

General Derivation

To derive MEPP, we begin from the full set of conditions for an uninvadable equilibrium to exist at environmental state $\mathbf{R}^*$:

Steady environment: $h_\alpha(\mathbf{R}^*) + \sum_i N_i q_{i\alpha}(\mathbf{R}^*) = 0$

Non-invasibility: $g_i(\mathbf{R}^*) \leq 0$

Feasible populations: $N_i \geq 0$

Steady populations: $N_i g_i(\mathbf{R}^*) = 0$,

where the first condition must hold for all $\alpha = 1, \ldots M$ and the last three conditions must hold for all $i = 1, \ldots, S$. These equations are very similar in form to the Karush-Kuhn-Tucker (KKT) conditions that must be satisfied at any local minimum $\mathbf{R}^*$ of a function $f(\mathbf{R})$ subject to inequality constraints $g_i(\mathbf{R}) \leq 0$ (Bertsekas 1999; Bishop 2006; Boyd and Vandenberghe 2004). The KKT conditions are:

Stationarity: $\nabla f(\mathbf{R}^*) + \sum_j \lambda_j \nabla g_j(\mathbf{R}^*) = 0$

Primal feasibility: $g_i(\mathbf{R}^*) \leq 0$

Dual feasibility: $\lambda_i \geq 0$

Complementary slackness: $\lambda_i g_i(\mathbf{R}^*) = 0$,

where the λ_i are generalized Lagrange multipliers (known in the literature as KKT multipliers) that enforce the inequality constraints. The KKT conditions have a straightforward and intuitive explanation: at the optimum $\mathbf{R}^*$, either $g_i(\mathbf{R}^*) = 0$ and the constraint is active $\lambda_i \geq 0$, or $g_i(\mathbf{R}^*) \leq 0$ and the constraint is inactive $\lambda_i = 0$.

The equilibrium conditions become identical to the KKT conditions when

$$\mathbf{q}_i = -\nabla g_i \tag{1}$$

$$\mathbf{h} = -\nabla f(\mathbf{R}) \tag{2}$$

for some function $f(\mathbf{R})$. We are free to add a constant to $f(\mathbf{R})$ while still satisfying the conditions, and can therefore always make $f(\mathbf{R}^0) = 0$ at its unconstrained minimum $\mathbf{R}^0$, which is the location of the unperturbed environmental state. Since f increases as we move away from the minimum, it can be used (at least locally, but usually globally) as a measure of the size of the perturbation

away from $\mathbf{R}^0$. The identity of the conditions for an uninvadable equilibrium with the KKT conditions then guarantees that the perturbation $d(\mathbf{R}^0, \mathbf{R}) \equiv f(\mathbf{R})$ is locally minimized in any uninvadable equilibrium state.

To apply MEPP to a given ecological model, one must transform the steady-state equation for the environment so that $\mathbf{q}_i = -\nabla g_i$. The transformed supply vector $\mathbf{h}$ then gives rise to the natural dissimilarity measure d. The central assumption stated in the main text, that the direction of $\mathbf{q}_i$ can be inferred from the ZNGI, guarantees that this is possible. For two resources, this assumption reduces to the simple statement that the angle between the ZNGI and the impact vector is species-independent (but can depend on the location in resource space). More generally, it means that the dependence of the impact vectors on $\mathbf{R}$ can be written as

$$q_{i\alpha} = -a_i(\mathbf{R}) b_\alpha(\mathbf{R}) \partial_\alpha g_i(\mathbf{R}). \tag{3}$$

The arbitrary positive-valued functions $a_i(\mathbf{R})$ change the magnitude of the impact vectors, but not the direction, while another set of arbitrary positive-valued functions $b_\alpha(\mathbf{R})$ provide a universal, species-independent mapping from the normal vector ∇g_i of the ZNGI to the impact vector. Whenever the relationship between requirements and impacts can be cast in this form, we can divide the steady environment condition by b_α to obtain

$$\frac{h_\alpha(\mathbf{R}^*)}{b_\alpha(\mathbf{R}^*)} - \sum_i N_i a_i(\mathbf{R}^*) \partial_\alpha g_i(\mathbf{R}) = 0. \tag{4}$$

If the vector h_α / b_α is curl-free, then we can also write

$$\frac{h_\alpha}{b_\alpha} = -\partial_\alpha d, \tag{5}$$

and find

$$\nabla d(\mathbf{R}^0, \mathbf{R}^*) + \sum_i N_i a_i(\mathbf{R}^*) \nabla g_i(\mathbf{R}^*) = 0. \tag{6}$$

This is identical to the stationarity equation in the KKT conditions, with KKT multipliers $\lambda_i = N_i a_i(\mathbf{R}^*)$, and objective function $f(\mathbf{R}) = d(\mathbf{R}^0, \mathbf{R})$.

The perturbation measure can be explicitly constructed when $h_\alpha(\mathbf{R}) = h_\alpha(R_\alpha)$ and $b_\alpha(\mathbf{R}) = b_\alpha(R_\alpha)$, which is the case in all examples considered here, giving:

$$d(\mathbf{R}^0, \mathbf{R}) = -\sum_\alpha \int_{R_\alpha^0}^{R_\alpha} \frac{h_\alpha(R'_\alpha)}{b_\alpha(R'_\alpha)}\, dR'_\alpha. \tag{7}$$

In online appendix A, we find b_α and construct the resulting perturbation measure for a range of specific models.

Data Analysis

To illustrate the consequences of MEPP, we simulated an invasion experiment based on a series of measurements by K. Rothhaupt of maximal clearance rates and growth curves for two species of zooplankton, *B. rubens* and *B. calyciflorus*, consuming two species of algae, *Monoraphidium minutum* and *Chlamydomonas sphaeroides* (Rothhaupt 1988).

The clearance rate is defined as the volume of water cleared of food per unit time per individual animal. This is equal to the per-capita consumption rate divided by the amount of food present, or $-q_{i\alpha}/R_\alpha$ in our notation. In this context $i = \{r, c\}$ indexes the zooplankton species, and $\alpha = \{m, c\}$ indexes the algae species. Rothhaupt measured the consumption rates $-q_{i\alpha}$ for all four pairwise combinations of zooplankton and algae, using short-term feeding experiments with radiolabeled food at a range of concentrations. The clearance rate is maximal at small food concentrations, where $q_{i\alpha}$ is approximately proportional to R_α, and decreases with increasing concentration as the consumption rate saturates. Rothhaupt divided the measured consumption rates by the food concentration for each measurement, and reported the average of $-q_{i\alpha}/R_\alpha$ over all measurements in the linear region. The results are reproduced in Table 1.

A saturating Monod model for consumption of multiple substitutable resources results in impact vectors of the form:

$$q_{i\alpha} = -\frac{c_{i\alpha} R_\alpha}{1 + \frac{\sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{K_i}}. \tag{8}$$

In this model, w_α represents the nutritional value (e.g., carbon content) of each unit of resource α, and consumption rates for species i begin to saturate when the available nutrient supply reaches

	Monoraphidium	*Chlamydomonas*
B. rubens	252 ± 7.4	49.0 ± 5.0
B. calyciflorus	211 ± 8.6	427 ± 32

Table 1: **Maximal clearance rates $c_{i\alpha}$ reported in Rothhaupt 1988**. Clearance rates for two species of zooplankton (rows) fed with two species of algae (columns) have been converted to units of μl/animal/day (from the originally reported units of μl/animal/hour). Each number is an average over between 11 and 38 animals, $\pm$ the standard error of measurement.

K_i. When the organism is supplied with only one resource type, and the food concentration is low enough that $w_\alpha R_\alpha \ll K_i$, then the impact vector can be approximated by the linear form

$$q_{i\alpha} \approx -c_{i\alpha} R_\alpha \tag{9}$$

so that $c_{i\alpha}$ are the maximal clearance rates.

Rothhaupt's measurements of per-capita growth rates as a function of resource concentration are reproduced in fig. 2. We model the growth rate as a weighted sum of the consumption rates, plus a constant offset:

$$g_i(\mathbf{R}) = \frac{e_i \sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{1 + \frac{\sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{K_i}} - e_i K_i + g_i^{\max}. \tag{10}$$

The new parameters here are e_i, which gives the efficiency with which species i converts consumed nutrients into biomass, and $g_i^{\max}$, which is the maximal growth rate that the species attains when the nutrient supply is large. It is straightforward to show that the relationship between g_i and $q_{i\alpha}$ in this model can be cast into the form of eq. (3), so that MEPP holds. The derivation is worked out in detail in online appendix A.

Since the resource abundances are reported in units of carbon mass per volume, we were able to set the unit nutritional value w_α equal to 1 for both algae species. We also required the maximal growth rate $g_i^{\max}$ to be the same for both species, leaving five free parameters for curve fitting. Rotthaupt points out that mechanical disturbance led to lower growth rates of *B. rubens* at high densities of *Chlamydomonas*, and so we included only the smallest three concentration values from

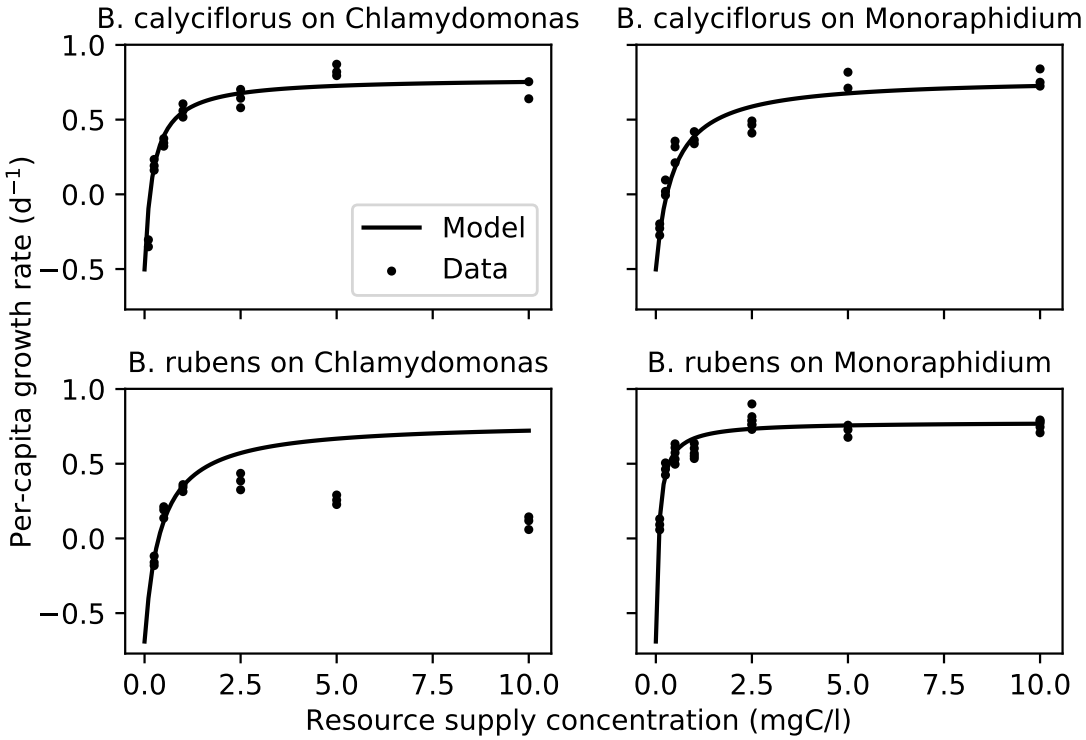

Figure 2: **Inferring parameters from for simulated experiments.** Data points were extracted from the printed figures in Rothhaupt 1988 using WebPlotDigitizer (Rohatgi 2010–). Solid lines represent the results of a simultaneous nonlinear least-squares fit of eq. (10) to all the data, with maximal clearance rates $c_{i\alpha}$ held fixed at their independently measured values listed in table 1. The best-fit values for the remaining five parameters were $g^{\max} = 0.78$ d^{-1}, $e_r = 0.073$ animal/ngC, $e_c = 0.014$ animal/ngC, $K_r = 20$ ngC/animal/day, $K_c = 3.86$ ngC/animal/day.

that growth curve in our analysis. The results of a simultaneous nonlinear least-squares fit to all four sets of growth rate measurements are plotted in fig. 2. Using the best-fit parameters and the full model described in online appendix A, we can predict equilibrium resource abundances under a variety of hypothetical experimental conditions.

Rothhaupt also reported the outcomes of competition between the two zooplankton species under twelve different experimental conditions, varying the dilution rate D and the externally provided resource supply $\mathbf{R}^0$. Dilution is incorporated into our model by simply subtracting D from the per-capita growth rates $g_i(\mathbf{R})$. The resource supply concentrations at $D = 0.45$/day

(*Chlamydomonas* ngC/μl : *Monoraphidium* ngC/μl) were 10:0, 8:2, 4:6, 6:4, 2:8, 0:10. The same concentrations were used in the series of experiments at $D = 0.2/\text{day}$, except that 8:2 was replaced by 6:1.5.

Results

The intrinsic environmental dynamics encoded in the supply vector $\mathbf{h}$ determine the form of the minimized perturbation measure $d(\mathbf{R}^0, \mathbf{R})$. Online appendix A contains detailed derivations of the perturbation measures arising from several commonly used models. In MacArthur's original Consumer Resource Model (MacArthur 1970), the resources are themselves self-replicating entities, such as plants or insects. This justifies a logistic form for $\mathbf{h}$, with exponential growth at low densities and saturation at a given carrying capacity. In this case, $d(\mathbf{R}^0, \mathbf{R})$ is a weighted Euclidean distance, as we have shown elsewhere (Mehta et al. 2018). In a chemostat culture, on the other hand, with a constant influx of resources and a fixed dilution rate, $\mathbf{h}$ is a linear function of $\mathbf{R}$, causing the resource abundances to relax exponentially to the chemostat supply point in the absence of consumers. The resulting perturbation measure turns out to be a weighted Kullback-Leibler divergence. This supply vector also be used as an approximate model of serial dilution experiments, where dilution and resource resupply occur at discrete time intervals. In both of these models, the contribution of each resource to the dissimilarity measure is weighted by its nutritional value, as discussed in appendix A and summarized in table 2.

Fig. 3 illustrates the application of MEPP to both of these paradigmatic cases, with contour plots of the two perturbation measures. In these two-dimensional examples, the equilibrium state can be visually identified as the point lying on the contour line closest to $\mathbf{R}^0$. Numerical simulations of the ecological dynamics with different initial conditions always end up at this point, even if the transient behavior is complicated.

When the uninvadable region is convex, the environmental perturbation has a unique local minimum along the boundary, which is therefore also the global minimum. In consumer resource

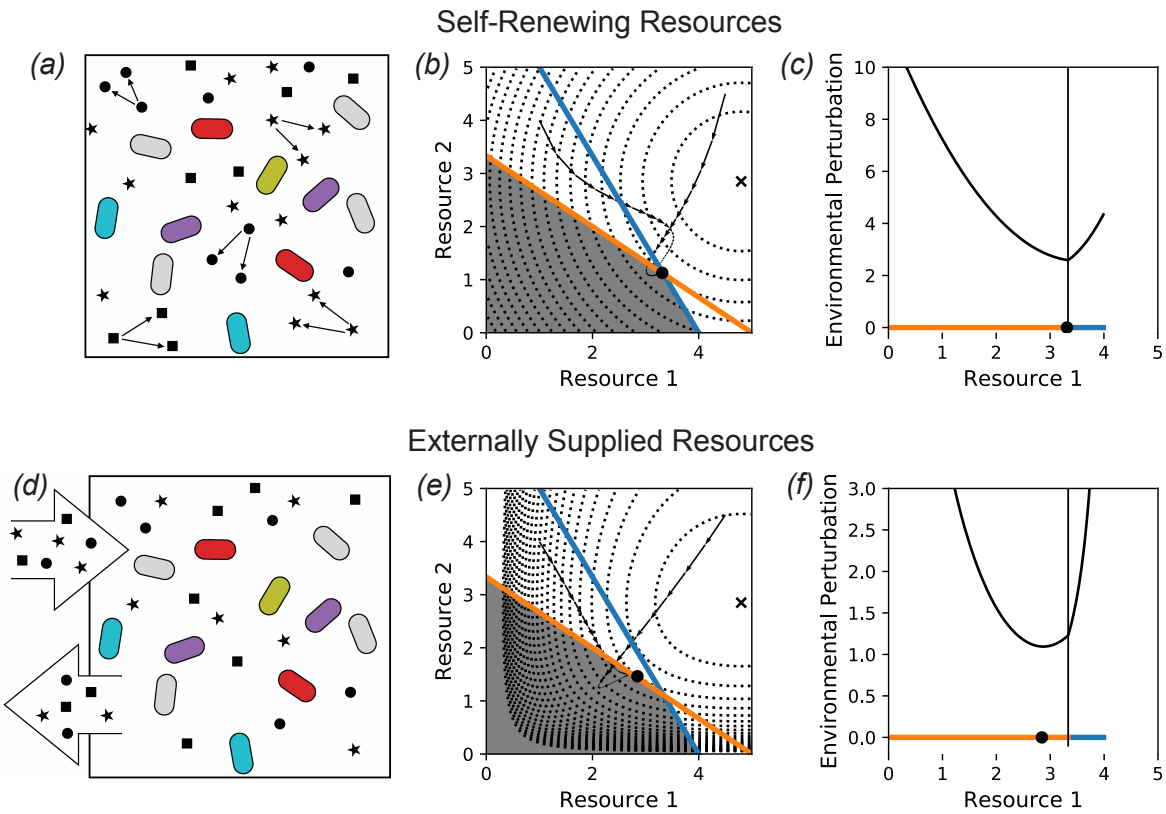

Figure 3: **Steady states of consumer resource models minimize environmental perturbation.** *(a)* Schematic of community with self-renewing resources (e.g. plants, insects). Colored shapes are consumers, with each color representing a different species, and small black shapes represent three different resource types. *(b)* Dynamical trajectories (arrows) of MacArthur's original model with self-renewing resource dynamics, starting from arbitrarily chosen initial resource concentrations, with equal populations of the two consumer species. Also shown are the contours of the relevant perturbation measure $d(\mathbf{R}^0, \mathbf{R})$ (dotted lines), the unperturbed environmental state $\mathbf{R}^0$ (black 'x'), and the ZNGI's (colored lines), with the uninvadable region shaded. *(c)* Environmental perturbation $d(\mathbf{R}, \mathbf{R}^0)$ along the boundary of the uninvadable region. Black dot/line indicates the location of the final steady state obtained in the simulations. *(d),(e),(f)* Same as above, but for externally supplied resources. See online appendices A and B for mathematical details and simulation parameters.

models with perfectly substitutable resources, where the ZNGI's are linear, one can show that this is always the case. This global optimization has directly testable consequences for ecological invasions. Before the invasion, the steady state of the ecosystem minimizes $d(\mathbf{R}^0, \mathbf{R})$ subject to the constraints imposed by all the currently existing consumer species. Invasion by a new species introduces a new constraint, without affecting any of the original constraints. Since $d(\mathbf{R}^0, \mathbf{R})$ was already minimized under the original set of constraints, the introduction of an additional constraint can only increase its value, or leave it unchanged. In other words, the steady-state value of $d(\mathbf{R}^0, \mathbf{R})$ can never decrease in any series of invasions.

Figure 4 shows how one might test this prediction in a microcosm experiment with two species of zooplankton, *B. rubens* and *B. calyciflorus*, that can consume two species of algae, *Monoraphidium minutum* and *Chlamydomonas sphaeroides* . This system was studied by K. Rothhaupt, who measured growth curves (which determine the ZNGI's) and maximal clearance rates (which determine the impact vectors) for each zooplankton-algae pair (Rothhaupt 1988), as described in the Methods. In fig. 2, we showed that these measurements are consistent with our key assumption, that the direction of each impact vector can be inferred from the corresponding ZNGI.

We consider an experiment in which *B. calyciflorus* is initially maintained in monoculture, fed with both species of algae in a fixed ratio. Once *B. calyciflorus* has reached an equilibrium density under constant supply and dilution rates, *B. rubens* is added to the system, and the species abundances eventually reach a new equilibrium state. The pre-invasion algae abundances can be predicted from Rothhaupt's measurements, using the model presented above in the Methods. We assume that the result of the invasion is identical to the result of a direct competition experiment starting with equal populations of the two species. Rothhaupt performed such experiments for 12 combinations of supply levels $\mathbf{R}^0$ and dilution rates, and recorded whether *B. calyciflorus*, *B. rubens* or both survived at non-zero density when the population dynamics reached equilibrium (Rothhaupt 1988). Using this information on the identity of the surviving species after the invasion, we can then infer the post-invasion algae abundances in the same way as the pre-invasion

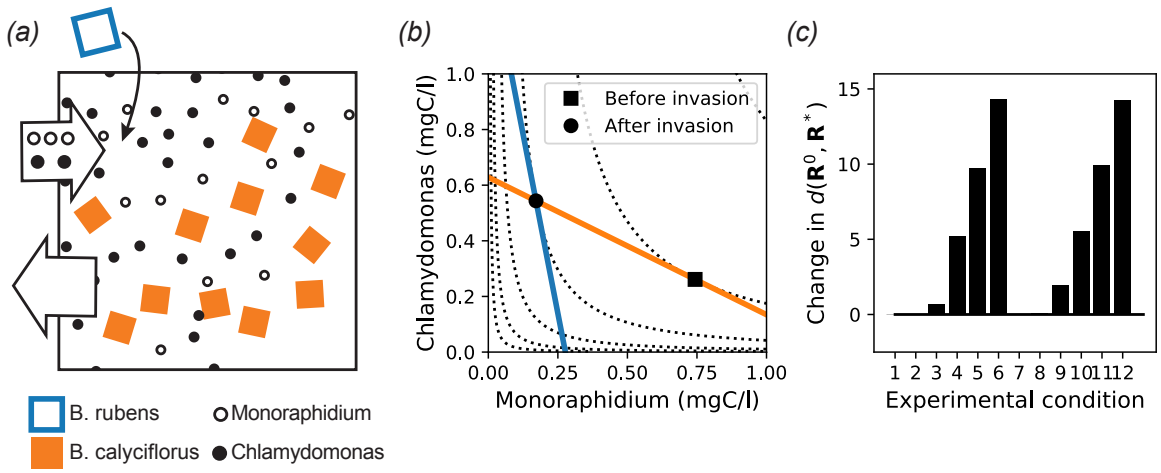

Figure 4: **Invasions monotonically increase the environmental perturbation.** *(a)* K. Rothhaupt tested the predictions of niche theory in an experimental system of two zooplankton, *B. rubens* and *B. calyciflorus*, competing for two algal species *Monoraphidium* and *Chlamydomonas* (Rothhaupt 1988). Using his measurements of growth and clearance rates, we simulated a simple invasion experiment: *B. calyciflorus* is initially at its steady-state population for a given supply of *Mono.* and *Chlamy.*, then *B. rubens* is introduced, and the system relaxes to a new steady state. The abundance of each of the algae is computed before and after the invasion, and these values determine the change in the environmental perturbation $d(\mathbf{R}^0, \mathbf{R}^*)$. *(b)* ZNGI's and contour lines of $d(\mathbf{R}^0, \mathbf{R})$. The pre-invasion steady state minimizes d along the *B. calyciflorus* ZNGI, while the post-invasion steady state lies on the boundary of the uninvadable region constructed from both ZNGI's. *(c)* The predicted change in $d(\mathbf{R}^0, \mathbf{R}^*)$ after invasion under the 12 different experimental conditions tested by Rothhaupt. See Methods for details.

abundances. Using our inferred values, we computed the environmental perturbation $d(\mathbf{R}^0, \mathbf{R}^*)$ before and after the invasion for each of the 12 conditions. These results are shown in Figure 4*(c)*. As expected, all conditions result in a change that is positive or zero.

As noted above, computing $d(\mathbf{R}^0, \mathbf{R}^*)$ requires knowledge of the nutritional unit value of each resources, in addition to the abundances. Since a *Chlamydomonas* cell occupies approximately forty times the volume of *Monoraphidium* (Rothhaupt 1988), one expects that the former will contribute much more to the growth rate per individual consumed than the latter. In Rotthaupt's experiments, however, algae abundances are reported in carbon mass per volume, rather than raw population sizes, and so the difference in cell size is already accounted for. The fits shown in fig. 2 confirm that in these units, equal value can be assigned to both resource types consistently with the data. This allows $d(\mathbf{R}^0, \mathbf{R}^*)$ to be computed directly from the reported resource abundances, with no additional parameters.

We now turn to the case where the ZNGI's are nonlinear, and multiple local mimima of $d(\mathbf{R}^0, \mathbf{R})$ can exist on the boundary of the uninvadable region. MEPP implies that each of these minima is a possible equilibrium state. Figure 5*(a)* shows the uninvadable region of an ecosystem with two alternative stable states, consisting of two species competing for two interactively essential resources (Ricklefs and Miller 1999). Panel *(c)* shows how $d(\mathbf{R}^0, \mathbf{R})$ varies along the boundary of this region. The latter plot reveals two local minima, one at the intersection of the two ZNGI's, where the two species coexist, and one at another point, where only one of the species survives. Numerical simulations in panel *(b)* with two different sets of initial population sizes confirm that each of the two minima can be reached under a suitable choice of initial conditions.

The presence of alternative stable states makes it possible for $d(\mathbf{R}^0, \mathbf{R}^*)$ to decrease under invasions. If the system begins in a local minimum of $d(\mathbf{R}^0, \mathbf{R})$ that is not the global minimum (e.g., the stable coexistence state in Figure 5*(c)*), invasion by a new species could destabilize this state, and send the system to the global minimum, decreasing the value of $d(\mathbf{R}^0, \mathbf{R}^*)$. This makes the environmental perturbation an interesting heuristic for investigating the existence of alternative stables states in microcosm or mesocosm experiments.

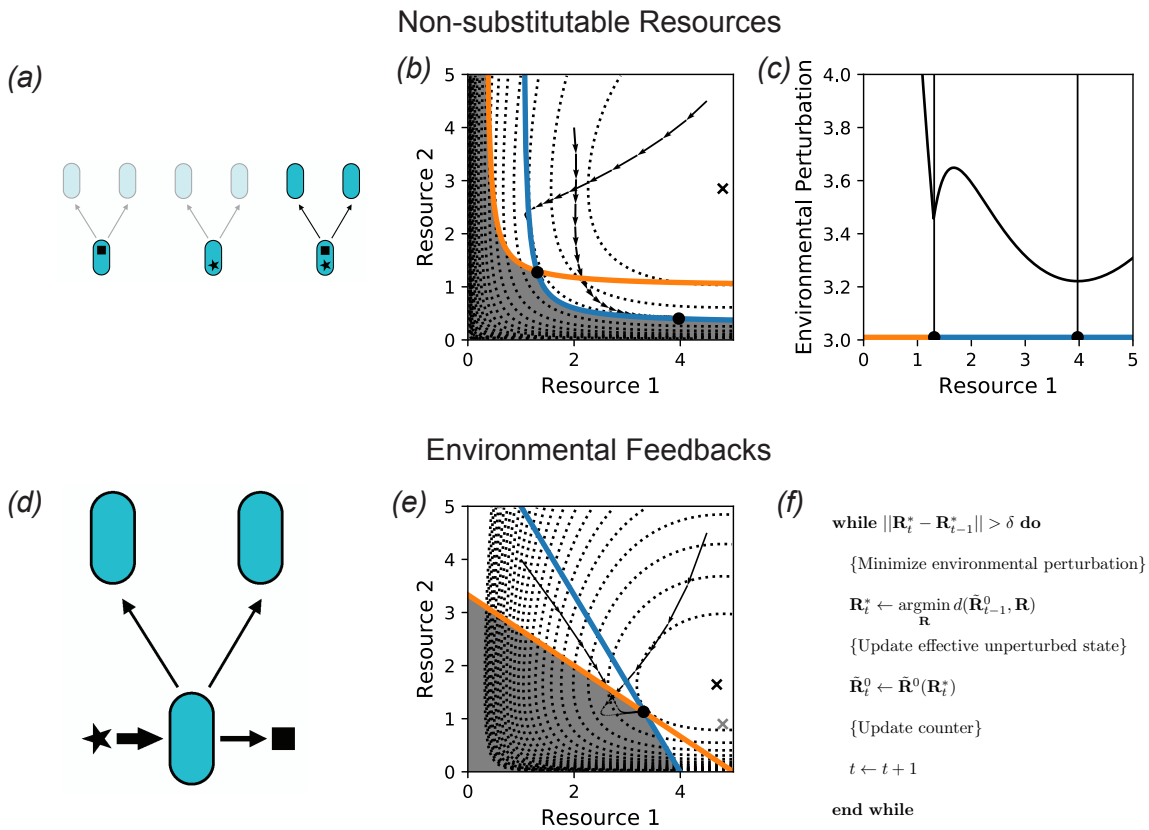

Figure 5: **Nonsubstitutable resources and environmental feedbacks.** *(a)* Growth law for a pair of essential resources (square and star). If either resource is absent, the organism cannot reproduce. *(b)* Dynamical trajectories of two interactively essential resources (arrows), starting two arbitrary sets of resource concentrations. The trajectory that leads to coexistence begins with a much smaller population of the orange species. The ZNGI's corresponding to this growth law are no longer straight lines, and they give rise to a non-convex uninvadable region. *(c)* Environmental perturbation along the boundary of the uninvadable region. The end points of the trajectories (black dots/lines) lie at two local minima which correspond to alternative stable states. *(d)* Schematic of an environmental feedback scenario, where consumers impact their environment not only by depleting resources but also by releasing metabolic byproducts. *(e)* Trajectories of two substitutable resources (arrows) in the presence of environmental feedbacks. The steady state no longer minimizes the dissimilarity with the true unperturbed state $\mathbf{R}^0$ (gray 'x'). Contours show the environmental perturbation with respect to an effective unperturbed state $\tilde{\mathbf{R}}^0$ (black 'x'), which incorporates the effect of the consumer mediated resource transformations. *(f)* Pseudocode of iterative algorithm for self-consistently finding $\tilde{\mathbf{R}}^0$ and the equilibrium resource concentrations $\mathbf{R}^*$. See Methods for mathematical details and parameters.

So far we have only considered cases of competition between pure consumers. But real ecosystems often exhibit more complex environmental feedbacks, sometimes referred to as 'bio-engineering' interactions (Chase and Leibold 2003): large mammals can destroy plant species by trampling or wallowing, beavers can create entirely new habitats for aquatic species, fish excrement can influence algal growth. In general, these effects break the required relationship between growth rates and impact vectors (Chase and Leibold 2003). In some important cases, however, an extended version of the principle still applies. One concrete example is cross-feeding among microbial species. Microbes generically release metabolic byproducts into their environment, serving as both consumers and producers of resources (Goldford et al. 2018; Pacheco et al. 2019). Since core carbon metabolism is largely conserved across many microbial species, these metabolic conversions can plausibly be approximately encoded into a universal stoichiometry matrix (Marsland III et al. 2019; Smith and Morowitz 2004).

Under this "universal chemistry" assumption, as illustrated in Figure 5*(b)* and *(c)*, byproduct secretion shifts the unperturbed environmental state from $\mathbf{R}^0$ (where the supply vector vanishes) to an "effective" location $\tilde{\mathbf{R}}^0$ (see online appendix A for details). Computing $\tilde{\mathbf{R}}^0$ requires prior knowledge of the steady-state resource concentrations $\mathbf{R}^*$, and so the equilibrium point cannot be found by straightforward optimization in this case. However, one can construct an iterative algorithm that starts with an arbitrary guess $\mathbf{R}_0^*$ for the equilibrium environmental state, as sketched in Figure 5*(f)*. The loop terminates when the self-consistency condition for the steady state is achieved, that is, when $\mathbf{R}_t^*$ minimizes $d(\tilde{\mathbf{R}}_t^0, \mathbf{R})$. This local version of the MEPP principle is intimately related to Expectation-Maximization algorithms used in statistical inference and machine learning (Mehta et al. 2019; Neal and Hinton 1998).

Discussion

The Minimum Environmental Perturbation Principle provides a new perspective on community ecology, by recasting the coexistence conditions of niche theory as the solution to an optimization

Description	Growth g_i	Impact $q_{i\alpha}$	Supply h_α	Perturbation d
Self-renewing CRM	$\sum_\alpha w_\alpha c_{i\alpha} R_\alpha - m_i$	$-c_{i\alpha} R_\alpha$	$rR_\alpha(R_\alpha^0 - R_\alpha)$	$\|\|\mathbf{R}^0 - \mathbf{R}\|\|_{\mathbf{w}}^2$
Externally supplied CRM	$\sum_\alpha w_\alpha c_{i\alpha} R_\alpha - m_i$	$-c_{i\alpha} R_\alpha$	$\tau^{-1}(R_\alpha^0 - R_\alpha)$	$D_{KL}^{\mathbf{w}}(\mathbf{R}^0 \|\| \mathbf{R})$
Saturating response	$\frac{e_i \sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{1+\frac{\sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{K_i}} - K_i + g_i^{\max}$	$-\frac{\sum_\alpha c_{i\alpha} R_\alpha}{1+\frac{\sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{K_i}}$	$\tau^{-1}(R_\alpha^0 - R_\alpha)$	$D_{KL}^{\mathbf{w}}(\mathbf{R}^0 \|\| \mathbf{R})$
Interactively essential resources (marginal-benefit foraging)	$\gamma \left(\sum_\alpha \frac{c_{i\alpha}}{R_\alpha}\right)^{-1} - m_i$	$-\gamma \frac{c_{i\alpha}}{R_\alpha} \left(\sum_\alpha \frac{c_{i\alpha}}{R_\alpha}\right)^{-2}$	$\tau^{-1}(R_\alpha^0 - R_\alpha)$	$D_{KL}(\mathbf{R}^0 \|\| \mathbf{R})$
Interactively essential resources (mass-conserving)	$\gamma \left(\sum_\alpha \frac{c_{i\alpha}}{R_\alpha}\right)^{-1} - m_i$	$-\gamma c_{i\alpha} \left(\sum_\alpha \frac{c_{i\alpha}}{R_\alpha}\right)^{-1}$	$\tau^{-1}(R_\alpha^0 - R_\alpha)$	$D_{KL}^{1/\mathbf{R}}(\mathbf{R} \|\| \mathbf{R}^0)$
Saturating transporters	$\sum_\alpha w_\alpha c_{i\alpha} \frac{R_\alpha}{K_\alpha + R_\alpha} - m_i$	$-c_{i\alpha} \frac{R_\alpha}{K_\alpha + R_\alpha}$	$\tau^{-1}(R_\alpha^0 - R_\alpha)$	$D_{KL}^{\mathbf{w}}(\mathbf{K} + \mathbf{R}^0 \|\| \mathbf{K} + \mathbf{R})$
Predation-limited	$K_i - \sum_\alpha d_{i\alpha} P_\alpha$	$d_{i\alpha} P_\alpha$	$-m_\alpha P_\alpha$	$\mathbf{m} \cdot \mathbf{P}$
Predation- and resource-limited	$\sum_\alpha c_{i\alpha} R_\alpha - \sum_\beta d_{i\beta} P_\beta$	$-c_{i\alpha} R_\alpha$, $d_{i\beta} P_\beta$	$rR_\alpha(R_\alpha^0 - R_\alpha)$, $-m_\alpha P_\alpha$	$\|\|\mathbf{R}^0 - \mathbf{R}\|\|^2 + 2\mathbf{m} \cdot \mathbf{P}$
Universal bioengineers	$g_i(\mathbf{R})$	$-Q_{\alpha\beta} R_\beta \partial_\beta g_i(\mathbf{R})$	$\tau^{-1}(R_\alpha^0 - R_\alpha)$	$D_{KL}^{1/Q}(\tilde{\mathbf{R}}^0 \|\| \mathbf{R})$
Species-dependent biomass value	$\sum_\alpha y_{i\alpha} c_{i\alpha} R_\alpha - m_i$	$-c_{i\alpha} R_\alpha$	$rR_\alpha(R_\alpha^0 - R_\alpha)$	n/a
Species-dependent saturation	$\sum_\alpha w_\alpha c_{i\alpha} \frac{R_\alpha}{K_{i\alpha} + R_\alpha} - m_i$	$-c_{i\alpha} \frac{R_\alpha}{K_{i\alpha} + R_\alpha}$	$\tau^{-1}(R_\alpha^0 - R_\alpha)$	n/a

Table 2: **Range of validity of MEPP.** This selective list of examples illustrates the range of systems that can be analyzed with the MEPP. It tabulates the perturbation measure that emerges from the general procedure given in eqs. 3 through 7 of the Methods for different choices of growth rate, impact vector, and supply vector. The perturbation is minimized over the vector of resource concentrations $\mathbf{R}$ and/or the vector of predator densities $\mathbf{P}$. $w_\alpha, y_{i\alpha}, e_i, m_i, g_i^{\max}, c_{i\alpha}, r, \tau, d_{i\alpha}, \gamma, K_\alpha$ and $Q_{\alpha\beta}$ are constants. We have defined the weighted KL divergence $D_{KL}^{\mathbf{w}}(\mathbf{x}||\mathbf{y}) \equiv \sum_\alpha w_\alpha \left[x_\alpha \ln \frac{x_\alpha}{y_\alpha} - (x_\alpha - y_\alpha)\right]$, and the weighted Euclidean distance $||\mathbf{x} - \mathbf{y}||_{\mathbf{w}}^2 \equiv \sum_\alpha w_\alpha (x_\alpha - y_\alpha)^2$. For $D_{KL}^{1/\mathbf{R}}$ and $D_{KL}^{1/Q}$, the superscripts are shorthand for vectors with elements $1/R_\alpha$ and $(Q^{-1})_{\alpha\alpha}$, respectively. For the "universal bioengineer" example, only a local form of MEPP applies, involving an effective unperturbed state $\tilde{\mathbf{R}}^0$ that depends on the steady-state resource vector $\mathbf{R}^*$ as discussed in the Environmental Feedbacks section of online appendix A. MEPP is not applicable in general to the last two examples, since the extra subscript on $w_{i\alpha}$ and $K_{i\alpha}$ means that the relationship between the ZNGI and the impact vector can be different for each species.

problem. The minimized quantity depends on the intrinsic environmental equilibrium state $\mathbf{R}^0$, and measures the perturbation away from this state induced by the organisms. Different models for the intrinsic environmental dynamics give rise to different perturbation measures: self-renewing resources such as insects or plants minimize the Euclidean distance, while externally supplied resources like sugars in a microbial culture minimize the Kullback-Leibler divergence.

Qualitatively, MEPP stems from a longstanding ecological intuition, though fleshed out in a new way. In the situations considered by niche theory, where all interactions are mediated by the environment, population growth eventually drives the environment to the edge of the uninvadable region (cf. Leibold 1995), unless the ecosystem exhibits chaos or limit cycles (Huisman and Weissing 2001*a*). Rates of population growth vanish only on this boundary, and so the population sizes will keep changing until the boundary is reached. This basic observation opens up a way to abstract from the complexities of the population dynamics, and focus on the environment. To do so, we further assume that the intrinsic dynamics of the environment exhibit a unique stable state $\mathbf{R}^0$, and are always acting to restore the environment to this state. If we now regard the population dynamics as simply a mechanism for constraining the environment to the boundary of the uninvadable region, these intrinsic dynamics will keep acting until the state cannot move any closer to $\mathbf{R}^0$ without leaving the boundary. Hence, the equilibrium state must be the one that is locally closest to $\mathbf{R}^0$.

MEPP thus provides a powerful new way of extending the graphical methods of ecosystem analysis popularized by Tilman to more complex situations. The 'R-star' criterion remains a necessary one for species coexistence (Chase and Leibold 2003; León and Tumpson 1975; Tilman 1982). Independently of any other assumptions, coexistence requires the ZNGI's to intersect, which means each species can survive at a lower concentration of a least one resource than any of its competitors. But the criterion for resource supply in the traditional graphical coexistence analysis requires that the supply vector field is linear (so that the supply vectors point towards the supply point), and is easiest to implement if the impact vectors are also fixed. These two assumptions are valid for self-renewing resources, but cannot simultaneously hold in even the

simplest models of externally supplied resources, as illustrated in Supplementary Figure S2. The basic geometrical intuition behind MEPP extends to a much broader class of systems, as illustrated in Table 1, with the supply vector always 'pulling' the equilibrium state to the closest point in the uninvadable region. Specific assumptions about the environmental dynamics are contained in the form of the perturbation measure, which determines the meaning of "closeness" in each case. Furthermore, as Chase and Leibold point out, the relationships among the impact vectors are increasingly difficult to visualize in higher dimensions (Chase and Leibold 2003). But high-dimensional constrained optimization is a long-established tool in many engineering disciplines, which have built up a deep pool of theorems, algorithms and intuition that can now be appropriated for ecological reasoning (Bertsekas 2014).

The validity of MEPP requires that impacts and requirements be related in a species-independent way. The correlation between these two niche components has long been known as a condition for ecosystem stability (Chase and Leibold 2003), and recent theoretical work has demonstrated that perfect correlation is indeed sufficient for stability in a broad class of models with externally supplied resources (Butler and O'Dwyer 2018, 2019). The fact that this assumption also leads to an optimization principle is further evidence of its importance for understanding ecological dynamics. Non-trophic interactions mediated by natural 'bioengineers' can easily break this postulate, underlining the ecological significance of environmental feedbacks from leaf litter to excrement decomposition (Chase and Leibold 2003; Jones et al. 1997; Molofsky and Augspurger 1992; Vanni et al. 2002).

We have identified one concrete example, involving zooplankton feeding on algae, where impacts and requirements have been independently measured, and this crucial assumption appears to hold. We have also showed that MEPP makes a definite prediction about the results of invasion experiments when growth is limited by substitutable resources: a successful invasion always increases the environmental perturbation $d(\mathbf{R}^0, \mathbf{R}^*)$, and can never cause it to decrease. With modern flow cytometry techniques, the resource concentrations in the zooplankton-algae system can be directly measured much more efficiently than when the experiments were orig-

inally carried out. This would facilitate the relevant measurements in the invasion experiment illustrated in Figure 4, where this prediction could be tested.

Acknowledgments

This work was supported by NIH NIGMS grant 1R35GM119461, Simons Investigator in the Mathematical Modeling of Living Systems (MMLS) to PM. We would also like to thank Ching-Hao Wang, Jacob Ferguson and William Ludington for useful discussions.

Literature Cited

Advani, M., G. Bunin, and P. Mehta. 2018. Statistical physics of community ecology: a cavity solution to MacArthur's consumer resource model. Journal of Statistical Mechanics 2018:033406.

Bertsekas, D. P. 1999. Nonlinear programming. Athena Scientific, Belmont, MA.

———. 2014. Constrained optimization and Lagrange multiplier methods. Academic Press, New York, NY.

Bishop, C. M. 2006. Pattern Recognition and Machine Learning. Springer, New York, NY.

Boyd, S., and L. Vandenberghe. 2004. Convex optimization. Cambridge University Press, Cambridge, UK.

Butler, S., and J. P. O'Dwyer. 2018. Stability criteria for complex microbial communities. Nature Communications 9:2970.

———. 2019. Cooperation and stability for complex systems in resource limited environments. bioRxiv:514018 .

Chase, J. M., and M. A. Leibold. 2003. Ecological niches: linking classical and contemporary approaches. University of Chicago Press, Chicago, IL.

Chesson, P. 1990. Macarthur's consumer-resource model. Theoretical Population Biology 37:26.

Elton, C. 1927. Animal Ecology. Sidgwick and Jackson, London, UK.

Gause, G. F., and A. A. Witt. 1935. Behavior of mixed populations and the problem of natural selection. The American Naturalist 69:596.

Goldford, J. E., N. Lu, D. Bajić, S. Estrela, M. Tikhonov, A. Sanchez-Gorostiaga, D. Segrè, P. Mehta, and A. Sanchez. 2018. Emergent simplicity in microbial community assembly. Science 361:469.

Huisman, J., and F. J. Weissing. 1999. Biodiversity of plankton by species oscillations and chaos. Nature 402:407.

———. 2001*a*. Biological conditions for oscillations and chaos generated by multispecies competition. Ecology 82:2682.

———. 2001*b*. Fundamental unpredictability in multispecies competition. The American Naturalist 157:488.

Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology .

Jones, C. G., J. H. Lawton, and M. Shachak. 1997. Positive and negative effects of organisms as physical ecosystem engineers. Ecology 78:1946–1957.

Jones, E., T. Oliphant, P. Peterson, et al. 2001–. SciPy: Open source scientific tools for Python.

Leibold, M. A. 1995. The niche concept revisited: mechanistic models and community context. Ecology 76:1371–1382.

León, J. A., and D. B. Tumpson. 1975. Competition between two species for two complementary or substitutable resources. Journal of theoretical Biology 50:185.

Levin, S. A. 1970. Community equilibria and stability, and an extension of the competitive exclusion principle. The American Naturalist 104:413–423.

MacArthur, R. 1969. Species packing, and what competition minimizes. Proceedings of the National Academy of Sciences 64:1369.

———. 1970. Species packing and competitive equilibrium for many species. Theoretical population biology 1:1–11.

MacArthur, R., and R. Levins. 1967. The Limiting Similarity, Convergence, and Divergence of Coexisting Species. The American Naturalist 101:377.

MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous forests. Ecology 39:599–619.

Marsland III, R., W. Cui, J. Goldford, A. Sanchez, K. Korolev, and P. Mehta. 2019. Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities. PLOS Computational Biology 15:e1006793.

Mehta, P., M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab. 2019. A high-bias, low-variance introduction to machine learning for physicists. Physics Reports (in press).

Mehta, P., W. Cui, C.-H. Wang, and R. Marsland III. 2018. Constrained optimization as ecological dynamics with applications to random quadratic programming in high dimensions. Physical Review E (in press).

Molofsky, J., and C. K. Augspurger. 1992. The effect of leaf litter on early seedling establishment in a tropical forest. Ecology 73:68–77.

Neal, R. M., and G. E. Hinton. 1998. A view of the EM algorithm that justifies incremental, sparse, and other variants. Pages 355–368 *in* Learning in Graphical Models. Springer, New York.

Pacheco, A. R., M. Moel, and D. Segrè. 2019. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. Nature Communications 10:103.

Posfai, A., T. Taillefumier, and N. S. Wingreen. 2017. Metabolic trade-offs promote diversity in a model ecosystem. Physical Review Letters 118:028103.

Rao, R., and M. Esposito. 2016. Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics. Physical Review X 6:041064.

Ricklefs, R. E., and G. L. Miller. 1999. Ecology. 4th ed. W.H. Freeman and Co., New York.

Rohatgi, A. 2010–. WebPlotDigitizer, https://automeris.io/WebPlotDigitizer.

Rothhaupt, K. O. 1988. Mechanistic resource competition theory applied to laboratory experiments with zooplankton. Nature 333:660.

Shear, D. 1967. An analog of the Boltzmann H-theorem (a Liapunov function) for systems of coupled chemical reactions. Journal of Theoretical Biology 16:212.

Smith, E., and H. J. Morowitz. 2004. Universality in intermediary metabolism. Proceedings of the National Academy of Sciences 101:13168–13173.

Taillefumier, T., A. Posfai, Y. Meir, and N. S. Wingreen. 2017. Microbial consortia at steady supply. eLife 6:e22644.

Tikhonov, M., and R. Monasson. 2017. Collective phase in resource competition in a highly diverse ecosystem. Physical Review Letters 118:048103.

———. 2018. Innovation rather than improvement: A solvable high-dimensional model highlights the limitations of scalar fitness. Journal of Statistical Physics 172:74.

Tilman, D. 1982. Resource Competition and Community Structure. Princeton University Press, Princeton, NJ.

Vanni, M. J., A. S. Flecker, J. M. Hood, and J. L. Headworth. 2002. Stoichiometry of nutrient recycling by vertebrates in a tropical stream: linking species identity and ecosystem processes. Ecology Letters 5:285–293.

Appendix A: Models

In the Methods, we noted that a sufficient condition for the applicability of MEPP is that the relationship between requirements and impacts can be written in the form

$$q_{i\alpha} = -a_i(\mathbf{R})b_\alpha(\mathbf{R})\partial_\alpha g_i(\mathbf{R}). \tag{A1}$$

where $a_i(\mathbf{R})$ and $b_\alpha(\mathbf{R})$ are positive-valued functions. The minimized perturbation measure can then be found by integrating h_α/b_α:

$$d(\mathbf{R}^0, \mathbf{R}) = -\sum_\alpha \int_{R^0_\alpha}^{R_\alpha} \frac{h_\alpha(R'_\alpha)}{b_\alpha(R'_\alpha)}\, dR'_\alpha. \tag{A2}$$

In this appendix, we find b_α and compute d for all five models discussed in the Results section.

Figures A1 and A2 graphically illustrate the effect of dividing by b_α on the impact and supply vectors for the first two models.

Self-renewing resources

In MacArthur's Consumer Resource Model (Chesson 1990; MacArthur 1970; MacArthur and Levins 1967), the environmental state is fully described by the abundances R_α of M substitutable resources, which are consumed by organisms of species i at a rate $c_{i\alpha}R_\alpha$. The per-capita growth rate g_i of each species is proportional to a weighted sum of per-capita consumption rates, minus a "maintenance cost" m_i that determines the consumption required to maintain a constant population size. The weights w_α in this sum measure the "quality" of each resource (e.g., energy density or carbon content). The resources are "self-renewing," with intrinsic dynamics that produce exponential growth up to a finite carrying capacity R^0_α. These definitions and assumptions lead to the following niche model:

$$g_i(\mathbf{R}) = \sum_{\alpha=1}^{M} w_\alpha c_{i\alpha} R_\alpha - m_i \tag{A3}$$

$$q_{i\alpha}(\mathbf{R}) = -c_{i\alpha}R_\alpha \tag{A4}$$

$$h_\alpha(\mathbf{R}) = r_\alpha R_\alpha (R^0_\alpha - R_\alpha). \tag{A5}$$

Since $\partial_\alpha g_i = w_\alpha c_{i\alpha}$, application of eq. (A1) yields

$$b_\alpha(R_\alpha) = \frac{R_\alpha}{w_\alpha}. \tag{A6}$$

Plugging in to eq. (A2), we obtain

$$d(\mathbf{R}^0, \mathbf{R}) = \frac{1}{2}\sum_\alpha w_\alpha r_\alpha (R_\alpha^0 - R_\alpha)^2. \tag{A7}$$

Dividing by R_α requires some care, however, since $h_\alpha = 0$ when $R_\alpha = 0$, and so it is possible for some R_α to vanish in the steady state. Examining the dynamics, we see that the steady state is only stable against invasion by extinct resources if

$$\partial_\alpha \left(d + \sum_i N_i g_i \right) = -w_\alpha \left[r_\alpha (R_\alpha^0 - R_\alpha) - \sum_{i=1}^{S} N_i c_{i\alpha} \right] > 0 \tag{A8}$$

whenever $R_\alpha = 0$. Since the gradient of the objective function points in the positive direction along this axis, all positive values of R_α are further away from the optimum. We can therefore accommodate this possibility within the constrained optimization framework, by simply adding a new set of constraints:

$$R_\alpha \geq 0. \tag{A9}$$

Externally supplied resources

A straightforward extension to MacArthur's Consumer Resource Model replaces the logistic dynamics of the resources with linear dynamics, which better capture the behavior of externally supplied nutrients that do not self-replicate:

$$g_i(\mathbf{R}) = \sum_{\alpha=1}^{M} w_\alpha c_{i\alpha} R_\alpha - m_i \tag{A10}$$

$$q_{i\alpha}(\mathbf{R}) = -c_{i\alpha} R_\alpha \tag{A11}$$

$$h_\alpha(\mathbf{R}) = \tau^{-1}(R_\alpha^0 - R_\alpha). \tag{A12}$$

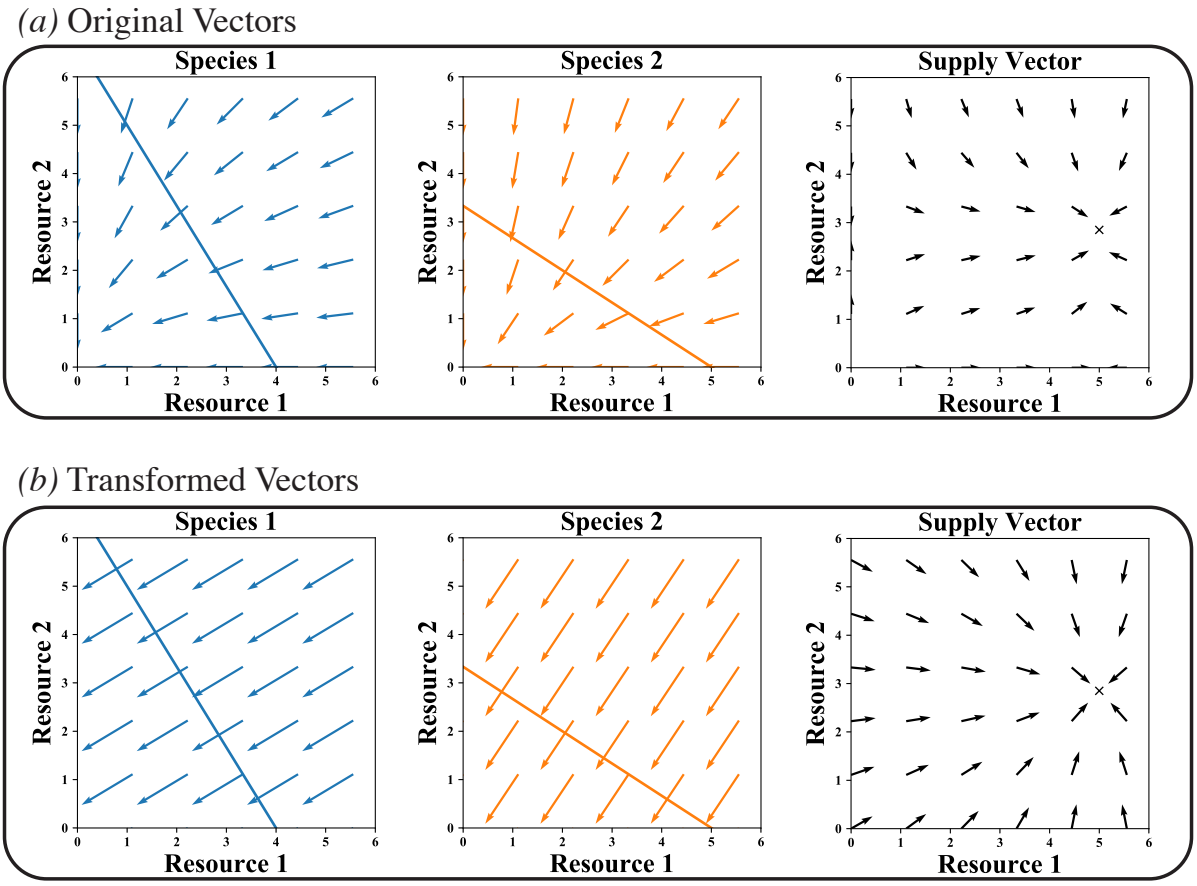

Figure A1: **Transforming the supply and impact vectors – MacArthur CRM.** *(a)* Impact vectors and ZNGI's for two species described by MacArthur's Consumer Resource Model, along with supply vectors corresponding to self-renewing resource dynamics. *(b)* Transformed impact and supply vectors resulting from division by the scaling factors b_α in eq. (A1), with the impact vectors now given by the gradient of the per-capita growth rate, and therefore orthogonal to the ZNGI's.

Since g_i and $q_{i\alpha}$ are unchanged, we obtain the same b_α as in the original model. Applying eq. (A2), we conclude:

$$d(\mathbf{R}^0, \mathbf{R}) = \sum_\alpha w_\alpha \tau^{-1} \left[R_\alpha^0 \ln \frac{R_\alpha^0}{R_\alpha} - (R_\alpha^0 - R_\alpha) \right]. \tag{A13}$$

Note that because R_α is not a normalized probability distribution, this formula differs from the usual definition of the Kullback-Leibler divergence employed in information theory, even when all the w_α's are equal, because of the new term $-(R_\alpha^0 - R_\alpha)$. This generalization to non-normalized vectors is widely used in chemical thermodynamics, where it is also known as Shear's Lyapunov Function (Rao and Esposito 2016; Shear 1967).

Model for Rothhaupt data analysis

In the Rothhaupt data, the empirically determined growth curves saturate at a fixed maximum growth rate $g_i^{\max}$ (Rothhaupt 1988). As noted in the Methods, they can be fit by a Monod curve with constant offset:

$$g_i(\mathbf{R}) = \frac{e_i \sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{1 + \frac{\sum_\alpha w_\alpha c_{i\alpha} R_\alpha}{K_i}} - K_i + g_i^{\max} \tag{A14}$$

where K_i and e_i are new parameters that control the saturation point and the efficiency of conversion from incoming carbon flux to biomass, respectively. The corresponding impact vectors are

$$q_{i\alpha} = -\frac{c_{i\alpha} R_\alpha}{1 + \frac{e_i \sum_\beta w_\beta c_{i\beta} R_\beta}{K_i}} \tag{A15}$$

where $c_{i\alpha}$ are the maximal clearance rates. Finally, the serial dilution protocol used in the experiments can be approximately treated as a chemostat, with supply vector

$$h_\alpha(\mathbf{R}) = \tau^{-1}(\mathbf{R}^0 - \mathbf{R}). \tag{A16}$$

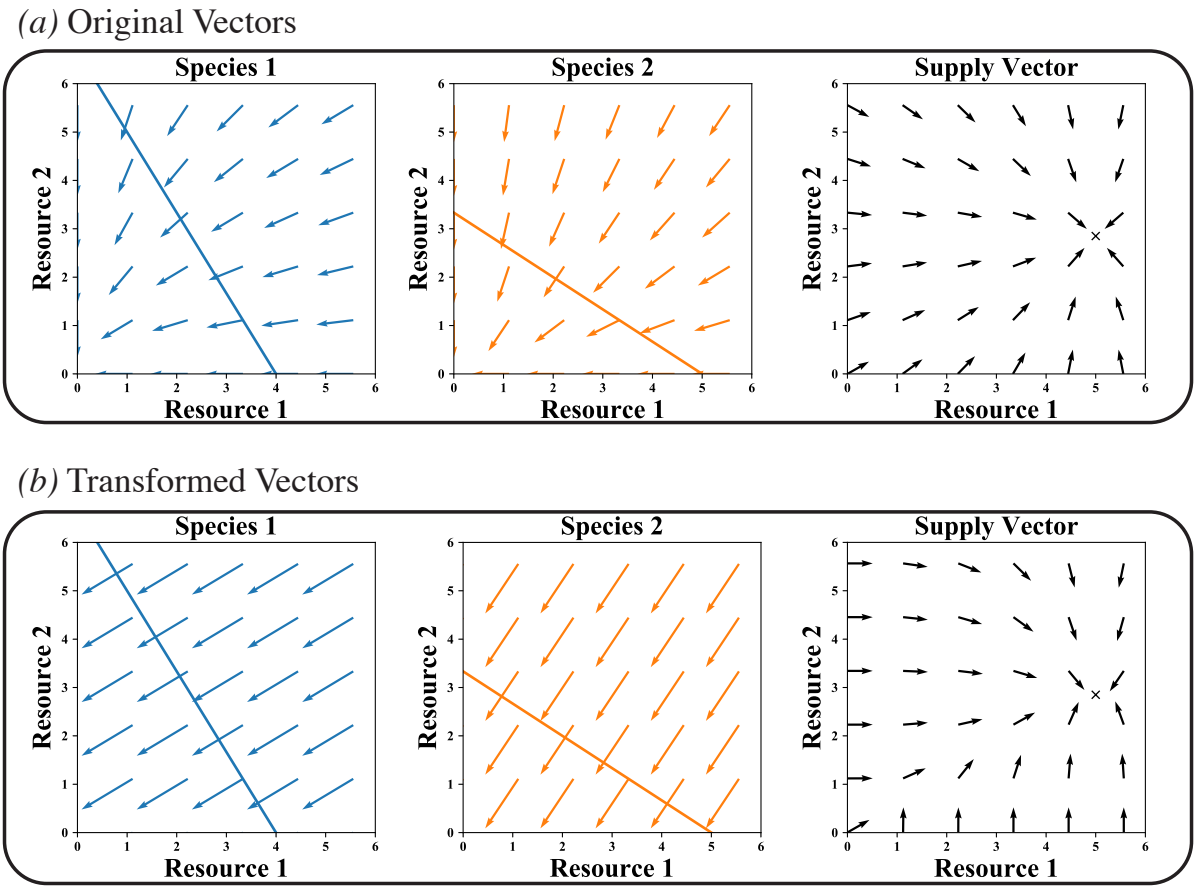

Figure A2: **Transforming the supply and impact vectors – External Supply.** *(a)* Impact vectors and ZNGI's for two species, along with supply vectors for externally supplied resource dynamics. *(b)* Transformed impact and supply vectors resulting from division by the scaling factors b_α in eq. (A1), with the impact vectors now given by the gradient of the per-capita growth rate, and therefore orthogonal to the ZNGI's. The simplest version of graphical ecosystem analysis does not apply here, because either the impact vectors are not constant (original vectors) or the supply vector field is nonlinear (transformed vectors).

The partial derivatives of g_i are:

$$\partial_\alpha g_i = \frac{e_i w_\alpha c_{i\alpha}}{1 + \frac{\sum_\beta w_\beta c_{i\beta} R_\beta}{K_i}} - \frac{e_i \sum_\beta w_\beta c_{i\beta} R_\beta \frac{w_\alpha c_{i\alpha}}{K_i}}{\left(1 + \frac{\sum_\beta w_\beta c_{i\beta} R_\beta}{K_i}\right)^2} \tag{A17}$$

$$= \frac{e_i w_\alpha c_{i\alpha}}{\left(1 + \frac{\sum_\beta w_\beta c_{i\beta} R_\beta}{K_i}\right)^2}. \tag{A18}$$

Applying eq. (A1), we find a complicated a_i, but the same b_α as in the other models:

$$a_i(\mathbf{R}) = \frac{1 + \frac{e_i \sum_\beta w_\beta c_{i\beta} R_\beta}{K_i}}{e_i} \tag{A19}$$

$$b_\alpha(R_\alpha) = \frac{R_\alpha}{w_\alpha}. \tag{A20}$$

Using eq. (A2), we find the same perturbation measure as in the previous section, with $\tau^{-1} = D$:

$$d(\mathbf{R}^0, \mathbf{R}) = \sum_\alpha w_\alpha D \left[R_\alpha^0 \ln \frac{R_\alpha^0}{R_\alpha} - (R_\alpha^0 - R_\alpha) \right]. \tag{A21}$$

Essential resources

In fig. 5, we consider the model of interactively essential resources described in (Taillefumier et al. 2017):

$$g_i(\mathbf{R}) = \gamma \left(\sum_\alpha \frac{c_{i\alpha}}{R_\alpha} \right)^{-1} - m_i \tag{A22}$$

where γ is a constant with units of inverse time, and the $c_{i\alpha}$ are constants with the same units as R_α, based on the ratios of resources required for building biomass. For the example presented in Figure 5*(a)*, we set the consumption rate per unit resource equal to the marginal benefit accrued from a slight increase of resource α:

$$-\frac{q_{i\alpha}}{R_\alpha} = \partial_\alpha g_i(\mathbf{R}) = \gamma \frac{c_{i\alpha}}{R_\alpha^2} \left(\sum_\alpha \frac{c_{i\alpha}}{R_\alpha} \right)^{-2} \tag{A23}$$

which becomes small if any resource is in short supply, because all resource types are required for growth. This results in the same b_α as the previous three examples, resulting in the same

forms for $d(\mathbf{R}^0, \mathbf{R})$. In the figure, we used externally supplied resources, so the perturbation measure was the weighted KL divergence.

Another reasonable choice for the impacts is to make them equal the growth rate times the biomass stoichiometry:

$$q_{i\alpha} = -c_{i\alpha}\gamma \left(\sum_{\beta} \frac{c_{i\alpha}}{R_\alpha} \right)^{-1}. \tag{A24}$$

Comparing this to the expression for the gradient of g_i, we find that

$$b_\alpha(R_\alpha) = R_\alpha^2. \tag{A25}$$

This gives

$$\frac{h_\alpha}{b_\alpha} = \frac{R_\alpha^0 - R_\alpha}{\tau} \frac{1}{R_\alpha^2}, \tag{A26}$$

which results in

$$d(\mathbf{R}^0, \mathbf{R}) = \sum_{\alpha} \frac{1}{\tau R_\alpha} \left[R_\alpha \ln \frac{R_\alpha}{R_\alpha^0} - (R_\alpha - R_\alpha^0) \right]. \tag{A27}$$

Environmental Feedbacks

The environmental feedbacks presented in Figure 5*(b)* are modeled by making each species return a fraction l_α of the "energy" acquired from each resource back to the environment, after transforming this energy into other resource types as specified by a universal chemical matrix $D_{\alpha\beta}$ (Goldford et al. 2018; Marsland III et al. 2019):

$$g_i(\mathbf{R}) = \sum_{\alpha=1}^{M} w_\alpha (1 - l_\alpha) c_{i\alpha} R_\alpha - m_i \tag{A28}$$

$$q_{i\alpha}(\mathbf{R}) = -c_{i\alpha} R_\alpha + \sum_{\beta=1}^{M} l_\beta c_{i\beta} D_{\alpha\beta} \frac{w_\beta}{w_\alpha} R_\beta \tag{A29}$$

$$h_\alpha(\mathbf{R}) = \tau^{-1}(R_\alpha^0 - R_\alpha) \tag{A30}$$

The impact vector can be rewritten in matrix form as:

$$q_{i\alpha}(\mathbf{R}) = -\sum_{\beta=1}^{M} Q_{\alpha\beta} R_\beta \partial_\beta g_i(\mathbf{R}) \tag{A31}$$

where

$$Q_{\alpha\beta} = \frac{1}{w_\beta(1-l_\beta)}\left[\delta_{\alpha\beta} - l_\beta D_{\alpha\beta}\frac{w_\beta}{w_\alpha}\right]. \tag{A32}$$

The role of b_α in eq. (A1) is now played by the matrix $Q_{\alpha\beta}R_\beta$. Instead of dividing through by b_α to obtain $q_{i\alpha} = -\partial_\alpha g_i$, we now must operate with the matrix inverse Q^{-1} of Q, whose elements satisfy $\sum_\beta (Q^{-1})_{\alpha\beta}Q_{\beta\gamma} = \delta_{\alpha\gamma}$, and then divide through by R_α.

When we perform this operation on the supply vector, we find that the perturbation measure must satisfy

$$\sum_\beta (Q^{-1})_{\alpha\beta}\frac{h_\beta(R_\beta)}{R_\alpha} = -\partial_\alpha d, \tag{A33}$$

which is the generalization of eq. (5) in the Methods. But the left-hand side now has a nonvanishing curl in general, as shown in fig. A3, and cannot be written as the gradient of an objective function.

We can locally recover MEPP in the vicinity of a given point in resource space by rewriting the left-hand side in the same form that arises in the absence of environmental feedbacks:

$$\sum_\beta (Q^{-1})_{\alpha\beta}\frac{h_\beta(R_\beta)}{R_\alpha} = \tilde{w}_\alpha \tau^{-1}\frac{\tilde{R}^0_\alpha - R_\alpha}{R_\alpha} \tag{A34}$$

where the weighting factors are

$$\tilde{w}_\alpha = (Q^{-1})_{\alpha\alpha} \tag{A35}$$

and the effective unperturbed environmental state is given by

$$\tilde{R}^0_\alpha(\mathbf{R}) = R^0_\alpha + \sum_{\beta\neq\alpha}\frac{(Q^{-1})_{\alpha\beta}\tau_\beta^{-1}}{(Q^{-1})_{\alpha\alpha}\tau^{-1}}(R^0_\beta - R_\beta). \tag{A36}$$

This state is a function of $\mathbf{R}$, but it needs to be a constant for $\mathbf{h}(\mathbf{R})$ to reduce to the desired form. We can make it constant by evaluating it at a fixed value of $\mathbf{R}$, but the point that minimizes the resulting perturbation measure $d(\tilde{\mathbf{R}}^0, \mathbf{R})$ is not necessarily the steady state. If we now re-evaluate $\tilde{\mathbf{R}}^0$ at the new value of $\mathbf{R}$, we will in general find a new optimum. When $\tilde{\mathbf{R}}^0(\mathbf{R})$ is evaluated at the true steady state location $\mathbf{R} = \mathbf{R}^*$, however, then the approximate resource equation with

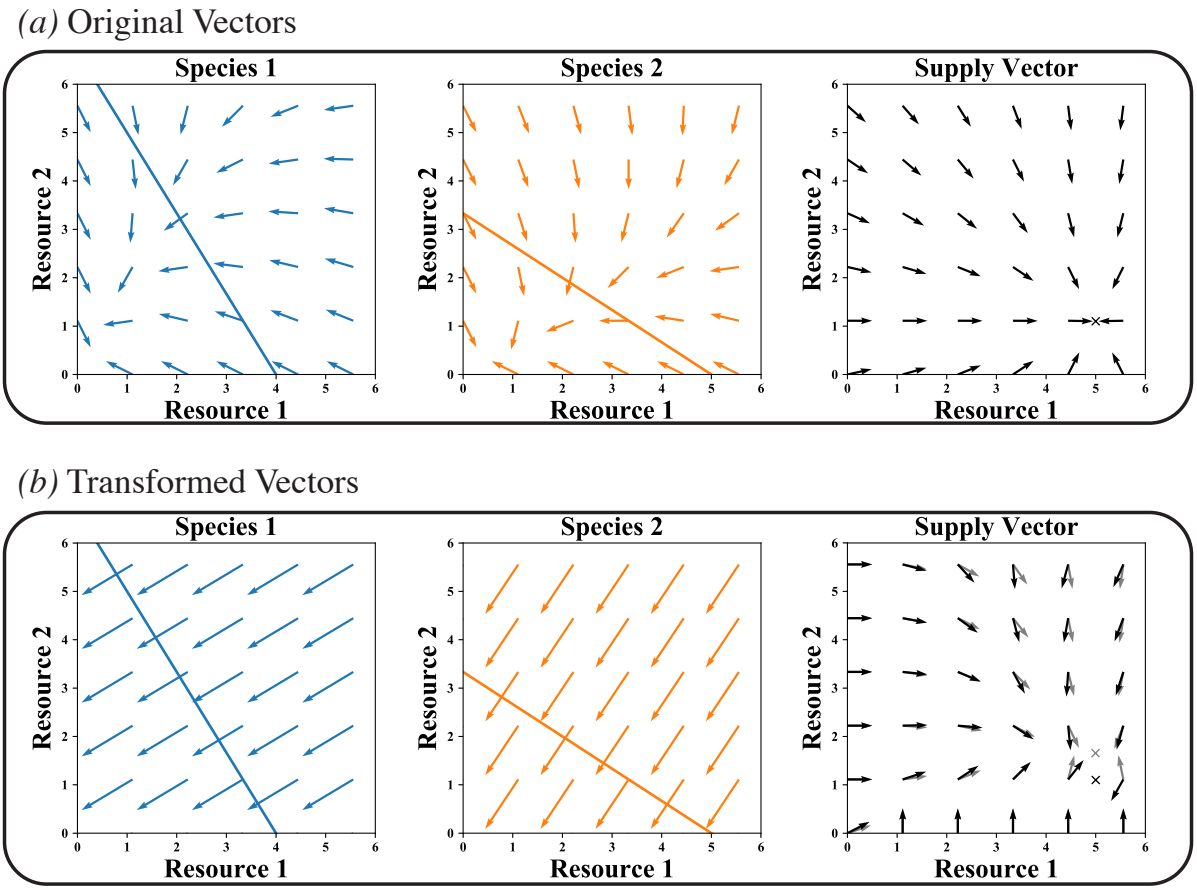

Figure A3: **Transforming the supply and impact vectors – Environmental Feedbacks.** *(a)* Impact vectors and ZNGI's for two species with environmental feedbacks, such that consumption of one resource is associated with production of the other, along with supply vectors corresponding to externally supplied resource dynamics. *(b)* Transformed impact and supply vectors, with the impact vectors now given by the gradient of the per-capita growth rate, and therefore orthogonal to the ZNGI's. The transformed supply vector field has a nonvanishing curl, and can therefore not be described as the gradient of any perturbation measure. It can be approximated by a curl-free field shown in gray, which becomes exact at the location of the steady state. The unperturbed environmental state corresponding to this approximate supply field is indicated with the gray 'x.'

constant $\tilde{\mathbf{R}}^0(\mathbf{R})$ also balances at $\mathbf{R}^*$. This means that the steady state does minimize $d(\tilde{\mathbf{R}}^0, \mathbf{R})$ for this choice of $\tilde{\mathbf{R}}^0$, and an algorithm that iteratively optimizes $d(\tilde{\mathbf{R}}^0, \mathbf{R})$ and updates $\tilde{\mathbf{R}}^0$ will halt when it finds this point. We note that this algorithm is intimately related to Expectation Maximization algorithms in statistical inference and machine learning (Mehta et al. 2019; Neal and Hinton 1998). The pseudocode for this algorithm is given in Figure 5*(f)* of the main text.

Appendix B: Simulation Parameters

In all simulations with substitutable resources (Figure 3 and bottom half of Figure 5), species 1 (blue) had preferences $c_{11} = 0.5, c_{12} = 0.3$, while species 2 (orange) had preferences $c_{21} = 0.4, c_{22} = 0.6$, and both species had the same maintenance cost $m_1 = m_2 = 2(1 - l)$. The leakage fraction l, which controls the magnitude of environmental feedbacks, was set to 0 for Figure 3, and to 0.5 for the bottom half of Figure 5. The metabolic matrix for the latter example $D_{12} = D_{21} = 1$, so that all secreted fluxes from consumption of one resource type were converted to the other resoruce type. The unperturbed resource concentrations for Figure 3 were $R_1^0 = 4.8, R_2^0 = 2.85$, while for the bottom half of Figure 5 they were $R_1^0 = 4.8, R_2^0 = 0.9$.

In the essential resource simulation from the top half of Figure 5, species 1 (blue) had $c_{11} = 1, c_{12} = 0.3$, species 2 (orange) had $c_{21} = 1, c_{22} = 0.3$, and both species had the same maintenance cost $m_1 = m_2 = 1$, with $\gamma = 1$. The unperturbed resource concentrations were $R_1^0 = 4.8, R_2^0 = 2.85$.

All simulations were initialized with $N_1 = 1, N_2 = 1$, except for the simulation with stable coexistence on non-substitutable resources, which had $N_1 = 1, N_2 = 0.01$.