Subpopulations and Stability in Microbial Communities

Pierre A. Haas,* Nuno M. Oliveira,† and Raymond E. Goldstein†
Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences,
University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom
(Dated: December 18, 2019)

In microbial communities, each species often has multiple, distinct phenotypes, but studies of ecological stability have largely ignored this subpopulation structure. Here, we show that such implicit averaging over phenotypes leads to incorrect linear stability results. We then analyze the effect of phenotypic switching in detail in an asymptotic limit and partly overturn classical stability paradigms: abundant phenotypic variation is linearly destabilizing but, surprisingly, a rare phenotype such as bacterial persisters has a stabilizing effect. Finally, we extend these results by showing how phenotypic variation modifies the stability of the system to large perturbations such as antibiotic treatments.

Over forty years ago, May suggested that equilibria of large ecological communities are overwhelmingly likely to be linearly unstable [1]. His approach did not specify the details of the dynamical system that describes the full population dynamics, but rather assumed that the linearized dynamics near the fixed point were represented by a random Jacobian matrix. Invoking results from random matrix theory, he concluded that unstable eigenvalues are more likely to arise as the number of interacting species increases. Actual large ecological communities certainly seem stable, and a major research theme in theoretical ecology has been to identify those features of the population dynamics that stabilize them [2–7].

Recent advances in our understanding of large natural microbial communities such as the human microbiome have emphasized the important link between stability and function: adult individuals typically carry the same microbiome composition for long periods of time and disturbances thereof are often associated with disease [8–10]. Moreover, while genetically identical organisms may exhibit different phenotypes [11–14] and despite the known ecological importance of phenotypic variation [15], studies of stability have largely ignored the existence of such subpopulations within species. Most models are therefore implicit averages over subpopulations.

We show here that this averaging yields incorrect stability results. With stochastic switching between phenotypes as an example of subpopulation structure, we show that while multiple abundant phenotypes are destabilizing, a rare phenotype can be stabilizing. This surprising result partly overturns May’s paradigm, and stresses the importance of phenotypic variation in ecological stability.

Our starting point is the Lotka–Volterra model [16], one of the most studied in population dynamics: \( N \) species with abundances \( \mathbf{A} \) compete as

\[
\dot{\mathbf{A}} = \mathbf{A} (\mathbf{\alpha} - \mathbf{D} \cdot \mathbf{A}), \tag{1}
\]

where \( \mathbf{\alpha} \) is a vector of birth rates and \( \mathbf{D} \) a matrix of non-negative competition strengths [17]; this is the competitive (as opposed to predator-prey) flavor of the model. If \( \det \mathbf{D} \neq 0 \), Eq. (1) has a unique equilibrium \( \mathbf{A}_* = \mathbf{D}^{-1} \cdot \mathbf{\alpha} \) of coexistence of all \( N \) species. This equilibrium is feasible (i.e. \( \mathbf{A}_* > 0 \)) if and only if \( \mathbf{\alpha} \) lies in the positive span of the columns of \( \mathbf{D} \).

If the \( N \) species each have two subpopulations with respective abundances \( \mathbf{B}, \mathbf{C} \), then

\[
\dot{\mathbf{B}} = \mathbf{B} (\beta - \mathbf{E} \cdot \mathbf{F} \cdot \mathbf{C}), \quad \dot{\mathbf{C}} = \mathbf{C} (\gamma - \mathbf{G} \cdot \mathbf{H} \cdot \mathbf{C}), \tag{2}
\]

where \( \beta, \gamma \) are birth rates, and \( \mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H} \) are competition strengths [18]. The dynamics of the sum \( \mathbf{B} + \mathbf{C} \) derive from the subpopulation-resolving “full” system (2) are not of the averaged form (1). However, to a coexistence equilibrium \( (\mathbf{B}_*, \mathbf{C}_*) \) of Eq. (2) we can associate an equilibrium \( \mathbf{A}_* \) of Eq. (1), determined by the requirement that, at equilibrium, population sizes, births, and competition be equal [Fig. 1(a)], i.e. that

\[
\mathbf{A}_* = \mathbf{B}_* + \mathbf{C}_*, \quad \mathbf{\alpha} \mathbf{A}_* = \beta \mathbf{B}_* + \gamma \mathbf{C}_*, \tag{3a}
\]

FIG. 1. Stability of communities with subpopulations. (a) The consistency conditions (3) relate the full model (2), in which each species has two subpopulations (dark and light individuals) that interact with all subpopulations of every other species, to the corresponding averaged model (1) without subpopulations. (b) Probability that a random stable equilibrium of the averaged model (1), \( \odot \), or the full model (2), \( \oplus \), is unstable in the other model, as a function of the number of species, \( N \). Probabilities were estimated from up to \( 10^7 \) random systems. Error bars indicate 95% confidence intervals larger than the plot markers. (c) Corresponding results for models with \( N = 2 \) species each having \( M \) subpopulations, as a function of \( M \).
and
\[ A_* DA_* = B_*(EB_* + FC_*) + C_*(GB_* + HC_*) . \] (3b)

These consistency conditions are a property of the model, based on the interpretation of its terms. Given Eqs. (2), they uniquely define an equilibrium \( A_* \) and an averaged model of the form (1), and this \( A_* \) is an equilibrium of this averaged model, and feasible if \( (B_*, C_*) \) is.

We select random averaged and subpopulation-resolving systems by sampling model parameters from a uniform distribution [19], and analyze the stability of their coexistence equilibria by computing the eigenvalues of their Jacobians. As the number of species \( N \) increases, stable equilibria of the averaged model (1) are increasingly likely to be unstable in the full model [Fig. 1(b)]. This is because the full model effectively has \( 2N \) species, and stable equilibria become increasingly rare as the number of species increases [1, 3]. It is therefore all the more striking that, as \( N \) increases, stable equilibria of the full model are also increasingly likely to be unstable in the averaged model [Fig. 1(b)]. The full model (2) can be extended to species with \( M \geq 2 \) subpopulations each, but increasing \( M \) at fixed \( N = 2 \) does not significantly affect the probability that a stable equilibrium of the full model destabilizes in the averaged model [Fig. 1(c)].

These toy models thus show that implicit averaging of subpopulations leads to incorrect stability results, and hence underlines their importance. Mathematically, this result is not fundamentally surprising: the determinant of the sum of two matrices is not the sum of their determinants, and so the linear relations between the Jacobians of the two systems resulting from Eqs. (3) cannot be expected to lead to simple relations between their stability.

We now specialize by taking phenotypic variation in microbial communities as an instance of subpopulation structure. It is useful to focus on one particular biological example: bacterial persisters [24, 26]. Bacteria such as Escherichia coli switch between a normal growth state and a persister state, in which they significantly suppress growth and competition of the persisters or the large separation of switching rates [29]. Adding a small parameter \( \varepsilon \ll 1 \), we therefore modify Eqs. (4) into
\[
\ddot{B} = B(\beta - E \cdot B - \varepsilon F \cdot C) - \kappa B + \lambda C , \quad \ddot{C} = C(\gamma - G \cdot B - H \cdot C) + \kappa B - \lambda C , \quad \varepsilon \ll 1 , \quad \varepsilon \approx 10^{-5} \text{ [29]} \cdot \text{ (5a)}
\]

where \( \kappa, \lambda \) are rates of (stochastic) switching. This form has previously been used to study phenotypic switching of a single species without competition [25, 29]. Since the rates are balanced, given an equilibrium of Eqs. (4), the consistency conditions (3) still define a corresponding "averaged" model (1) without phenotypic variation.

To analyze the effect of this switching, we again compare the stability of the full and averaged models. Steady states of Eqs. (4) cannot in general be found in closed form. To sample random systems, we must therefore sample parameters indirectly [19]. With increasing number of species, random stable coexistence states of either this full model or the corresponding averaged model again become increasingly likely to be unstable in the other model (Fig. 2).

Eqs. (4) do not however take into account the weak growth and competition of the persisters or the large separation of switching rates [29]. Adding a small parameter \( \varepsilon \ll 1 \), we therefore modify Eqs. (4) into
\[
\dot{B} = B(\beta - E \cdot B - \varepsilon F \cdot C) - \kappa B + \lambda C , \quad \dot{C} = C(\gamma - G \cdot B - H \cdot C) + \kappa B - \lambda C . \quad \varepsilon \ll 1 , \quad \varepsilon \approx 10^{-5} \text{ [29]} \cdot \text{ (5b)}
\]

Hence \( B \) are normal cells and \( C \) are persisters. For wild-type E. coli, \( \varepsilon \approx 10^{-5} \) [29], but here, we take \( \varepsilon = 0.01 \) for numerical convenience. We justify this below by confirming the numerical results by an asymptotic analysis of Eqs. (5). A more intricate asymptotic separation of parameters arises in the hipQ mutant of E. coli [29], but we do not pursue this further.

The separation of growth and competition terms and switching rates in Eqs. (5) allows for steady states to be found by expansion in \( \varepsilon \). Writing \( B_0 = B_0 + \varepsilon B_1 + O(\varepsilon^2) \) and \( C_0 = C_0 + \varepsilon C_1 + O(\varepsilon^2) \), we find \( B_0 = E^{-1} \cdot \beta \) and \( C_1 = \kappa B_0/\lambda \), but \( B_1 = C_0 = 0 \) [19]. This is the expected asymptotic separation of the two population sizes: few cells are persisters, at least under laboratory conditions [29]. This asymptotic solution enables direct
sampling of all model parameters [19].

To analyze the stability of equilibrium of Eqs. (5), we expand its Jacobian, $J_0 = J_0 + \varepsilon J_1 + O(\varepsilon^2)$, finding

$$
J_0 = \begin{pmatrix} -B_0E & \lambda I \cr 0 & -\lambda I \end{pmatrix}, \quad J_1 = \begin{pmatrix} -\kappa I & -B_0F \\ \kappa I & (\gamma - G \cdot B_0)I \end{pmatrix}, \tag{6}
$$

with $I$ the identity and $O$ the zero matrix [19]. The averaged model has Jacobian $K_0 = K_0 + \varepsilon K_1 + O(\varepsilon^2)$, with

$$
K_0 = -B_0E, \quad K_1 = B_0E K_0 \frac{\kappa}{\lambda}. \tag{7}
$$

Since $J_0$ is block-upper-triangular, its eigenvalues are those of $-\lambda I$, which are stable, and those of $-B_0E = K_0$. Hence any unstable eigenvalues of $J_0$ and $K_0$ are equal to lowest order in the expansion. Equivalently, at $\varepsilon = 0$, the full phenotype-resolving model (5) is stable if and only if the corresponding averaged model is stable. This result is not borne out however by numerics at finite $\varepsilon$: as $N$ increases, the probability that a random stable equilibrium of the full model is destabilized in the corresponding averaged model still increases (Fig. 2), although the probability is reduced compared to the previous case. Much more strikingly, the probability that a stable equilibrium of the averaged model is destabilized in the full model is vastly reduced (Fig. 2, inset). This is all the more surprising as we argued earlier that the opposite behavior was to be expected since larger systems are more likely to be unstable. We have also sampled exact equilibria of Eqs. (5), similarly to our analysis of Eqs. (4) above, yielding results in qualitative agreement with those based on the asymptotic equilibria (Fig. 2, inset). This justifies basing the detailed analysis of the destabilization mechanism on the asymptotic results.

To explain these surprising results, we analyze the spectra of the Jacobians in more detail. With the exception of a single outlier eigenvalue that is large and negative, the eigenvalues of $K_0$ lie approximately within a circle [Fig. 3(a)], as expected from the circular law of random matrices [30]. The spectral distribution of $J_0$ is the sum of this distribution and the (uniform) distribution of eigenvalues of $-\lambda I$ [Fig. 3(b)]. The outlier eigenvalue can be analyzed in great generality [31], but heuristics suffice here: denoting by $-\mu_0 < 0$ the mean of the distribution of entries of $K_0$ and neglecting correlations between entries, each row of $K_0$ has approximate sum $-N\mu_0$, and so $K_0$ has an approximate eigenvector $1 = (1,1,\ldots,1)$ with eigenvalue $-N\mu_0 < 0$, as argued in the Supplemental Material of Ref. [3]. The other eigenvalues of $K_0$ [Fig. 3(c)] are uniformly distributed on a disk of radius $r \sim \sqrt{N}$ for $N \gg 1$ by the circular law [30]. (Hence, by the Perron–Frobenius theorem [22], the outlier eigenvalue is indeed real.) An eigenvalue $\nu_0$ of $K_0$ has $|\text{Re}(\nu_0)| \lesssim \varepsilon$ with probability $\varpi \sim (c\varepsilon)/r^2 \sim \varepsilon/\sqrt{N}$, i.e. $\varpi = c\varepsilon/\sqrt{N}$, for some $c = O(1)$. The average distance $\delta$ between eigenvalues is determined by $N\delta^2 \sim r^2$, so $\delta = O(1)$ and the eigenvalues of $K_0$ are pairwise different at order $O(1)$ [Fig. 3(c)]. Hence, if $\nu_0$ is an eigenvalue of $K_0$, then $K_0$ has an eigenvalue $\nu_0 = \nu_0 + O(\varepsilon)$ [19]. Thus $\nu_*$ is stable if either (i) $\text{Re}(\nu_0) < 0$ and $|\text{Re}(\nu_0)| \lesssim \varepsilon$ or (ii) $|\text{Re}(\nu_0)| \lesssim \varepsilon$ and the small real part of $\nu_*$ is stabilized by $K_1$. By definition, $|\text{Re}(\nu_0)| \lesssim \varepsilon$ with probability $\varpi$, so (i) occurs with probability $(1 - \varpi)/2$. Let $q$ denote the probability of stabilization by $K_1$ in case (ii). Summing over the $N - 1$ non-outlier eigenvalues of $K_0$, the probability $p = P(K_0 \text{ stable})$ is

$$
p = \sum_{k=0}^{N-1} \binom{N-1}{k} \varpi^k q^{N-1-k} \sim \left(1 + \frac{c(2q-1)\varepsilon}{2\sqrt{N}}\right)^{N-1} \sim \exp\left[\frac{c(2q-1)\varepsilon}{2\sqrt{N}}\right], \tag{8}
$$

for $N \gg 1$, using the binomial theorem and $(1 + 1/x)^x \sim e^x$ for $x \gg 1$. A similar expression determines $p' = P(J_0 \text{ stable})$, with $q$ replaced by $q'$. Eq. (6) shows that $J_1$ acts on $K_0$ as the negative definite matrix $-\lambda I$, so $q' > q$ [19]. It follows that

$$
\frac{p'}{p} \sim \exp\left[2c(q' - q)\varepsilon\sqrt{N}\right] \rightarrow \infty \quad \text{as} \quad N \rightarrow \infty, \tag{9}
$$

confirming the trend in Fig. 2: the full model is much more likely to be stable than the averaged model.

Hence, switching to a rare phenotype such as persisters can enhance the stability of a community. The detailed analysis of Eqs. (5) above has emphasized that this
effect relies on both the spectral distribution (which allows terms beyond leading order to change the stability of the system) and the detailed structure of the system (which can suppress or enhance this mechanism). By contrast, switching to an abundant phenotype destabilizes the community: the introduction of such a phenotype essentially increases the effective number of species, which is destabilizing [1, 3]. Reference [6] recently introduced a family of models with explicit resource dynamics for which any feasible equilibrium is stable. Switching to an abundant phenotype also destabilizes a phenotype-resolving version of the model of Ref. [6] provided that the difference in switching rates is large enough [19], confirming that this effect is generic. These conclusions are therefore likely to be relevant for the stability of microbial communities such as the microbiome, for which competitive interactions are known to play an important, stabilizing role [10].

Linear stability analysis cannot elucidate the effect of large perturbations on coexistence. These might arise from antibiotic treatments, which bacterial communities can survive by forming persisters [24]. In the final part of this Letter, we therefore explore such perturbations numerically. Rather than modelling the dynamics of antibiotic treatment in detail [19], we suppose that it reduces the abundances of both normal cells and persisters, and we ask: does the community converge back to coexistence, or do some species disappear from the community? Do the answers from the averaged and full models differ? To answer these questions, we reconsider the exact stable equilibria of the full model (5) that are also stable in the averaged model, and evolve both systems from consistent random small initial conditions using the stiff solver ode15s of MATLAB (The MathWorks, Inc.).

Fig. 4(a) shows the distribution of possible outcomes: (1) convergence back to coexistence of all $N$ species, (2) convergence to a new coexistence state of $n < N$ species, and (3) convergence to a limit cycle. (The trivial equilibria $A = 0$ and $B = C = 0$ of the averaged and full models are clearly unstable, so $1 \leq n < N$.) The probability of outcome (2) increases with $N$ [Fig. 4(a)] and the distribution of $n$ in Fig. 4(b) shows that $n \gtrapprox N/2$ is somewhat more likely than $n \lesssim N/2$. Thus if the whole community does not survive, then at least half does. The averaged and full persister models give comparable outcome distributions. This does not contradict our earlier result that a rare phenotype stabilizes the community since here, we only consider states stable in both the averaged and full models so that results can be compared meaningfully. In fact, as $N$ increases, individual realizations of full and averaged models are increasingly likely to give different outcomes [Fig. 4(c)]. In most cases, outcomes differ because the system converges back to coexistence of all species in one model only; in a small number of cases, different species die out in the full and averaged models [Fig. 4(c)]. These observations thus extend our results for linear perturbations to large perturbations.

Here, we have revealed the strong effects of subpopulation structure on the stability of competing microbial communities and the surprising stabilizing effect of stochastic switching to a rare phenotype. Very recently, Ref. [32] similarly emphasized the stabilizing effect of phenotypic variation using a different model, without phenotypic switching. While the competitive interactions considered here are important in systems such as the human microbiome [10], future work will need to explore the phenotypic structure in more detail. The interaction structure of ecological communities without phenotypic variation is known to affect their stability [3, 4], but these and related studies, in the spirit of May’s seminal work [1], are based on the analysis of random Jacobians. By contrast, here, we could not avoid specifying explicit dynamical systems, since we had to establish a correspondence between full and averaged models (and indeed our analysis has shown that the details of the model structure can matter here). In particular, this precludes the question of feasibility to the question of stability. This question of feasibility can be treated in a statistical sense [33, 34] but cannot be eschewed in gen-

![FIG. 4. Stability of microbial communities to large perturbations. (a) Distribution of possible long-time behaviors after large perturbations of the full model (5) and the corresponding averaged model as a function of $N$. Possible behaviors are (1) convergence back to coexistence of all species, (2) convergence to coexistence of $n < N$ species, or (3) non-convergence to a steady state. In some cases (‘#’), numerical solution failed. Probabilities were estimated from up to $10^4$ random systems. (b) Distribution of the number $n$ of remaining species for outcome (2), for $N = 10$ and for both the full and averaged models. (c) Probability that the full and averaged models converge to different coexistence states, given that both models converge to some coexistence state. The contribution from systems for which one model converges to coexistence of all species is highlighted. Error bars indicate 95% confidence intervals.](image-url)
eral. Nonetheless, our analysis only relying on generic properties of the spectral distribution suggests that our conclusions apply not only to competitive, but also to more general interactions.

The authors gratefully acknowledge support from Engineering and Physical Sciences Research Council Established Career Fellowship EP/M017982/1 and Gordon and Betty Moore Foundation Grant 7523 (both to R.E.G.), a Herchel Smith Postdoctoral Research Fellowship (N.M.O.), and Magdalene College, Cambridge (Nevile Research Fellowship to P.A.H.).

* P.A.Haas@damtp.cam.ac.uk
† N.M.Oliveira@damtp.cam.ac.uk
‡ R.E.Goldstein@damtp.cam.ac.uk

[17] Throughout, we imply elementwise multiplication of vectors and (rows or columns of) matrices by writing symbols next to each other. Dots denote tensor contractions.
[18] For $N = 2$, systems of this form have been considered in analyses of evolutionary stability, see e.g. R. Cressman and J. Garay, Evolutionary stability in Lotka–Volterra systems, J. Theor. Biol. 222, 233 (2003).
[19] See Supplemental Material at [url to be inserted], which includes Refs. [5, 6, 20–25], for (i) a detailed description of the selection of random systems, (ii) details of the asymptotic calculations leading to Eqs. (6) and (7), (iii) details of the spectral analysis leading to Eq. (9), including a note on the stability of sums of stable matrices, (iv) a discussion of models with explicit resource dynamics, based on those of Ref. [6], and (v) a brief discussion of the dynamics of antibiotic treatments.