🔓 OPEN ACCESS

Chaotic turnover of rare and abundant species in a strongly interacting model community

Emil Mallmin[a,1] (ID), Arne Traulsen[a] (ID), and Silvia De Monte[a,b] (ID)

The composition of ecological communities varies not only between different locations but also in time. Understanding the fundamental processes that drive species toward rarity or abundance is crucial to assessing ecosystem resilience and adaptation to changing environmental conditions. In plankton communities in particular, large temporal fluctuations in species abundances have been associated with chaotic dynamics. On the other hand, microbial diversity is overwhelmingly sustained by a "rare biosphere" of species with very low abundances. We consider here the possibility that interactions within a species-rich community can relate both phenomena. We use a Lotka–Volterra model with weak immigration and strong, disordered, and mostly competitive interactions between hundreds of species to bridge single-species temporal fluctuations and abundance distribution patterns. We highlight a generic chaotic regime where a few species at a time achieve dominance but are continuously overturned by the invasion of formerly rare species. We derive a focal-species model that captures the intermittent boom-and-bust dynamics that every species undergoes. Although species cannot be treated as effectively uncorrelated in their abundances, the community's effect on a focal species can nonetheless be described by a time-correlated noise characterized by a few effective parameters that can be estimated from time series. The model predicts a nonunitary exponent of the power-law abundance decay, which varies weakly with ecological parameters, consistent with observation in marine protist communities. The chaotic turnover regime is thus poised to capture relevant ecological features of species-rich microbial communities.

ecological chaos | rare biosphere | plankton ecology | disordered Lotka–Volterra model

The dynamic nature of ecological communities of species has long been recognized (1). Fluctuations in species' abundances can have fundamentally different origins depending on the spatial and temporal scales considered and the particular community of interest (1, 2). For example, if environmental filtering shapes community composition, abundance fluctuations may reflect changing external conditions. Another possibility is that stochastic effects of demography, dispersal, and diversification dominate over the ecological differences between species in driving turnover (3). As communities are increasingly monitored in the wake of climate change and biodiversity decline, there is growing opportunity and need to understand why abundances fluctuate and how fluctuations relate to patterns of biodiversity and biogeography.

An alternative to environmental and stochastic effects as the main determinants of abundance fluctuations is the hypothesis that they reflect intrinsically chaotic dynamics arising from the complexity of ecological interactions. Mathematically, chaos is the phenomenon whereby a deterministic, nonlinear dynamical system (e.g. describing the populations of interacting species) generates bounded but aperiodic trajectories that depend sensitively on initial conditions (4). While chaos can be readily identified in simple mathematical models, its presence in empirical time series is challenging to ascertain, and the relevance of chaos for natural communities has been controversial (5, 6). However, recent methodological advances and systematic assessment of a large ecological time series database using validated, nonparametric methods showed that ecological chaos is generally not rare (7, 8) and is particularly prevalent in planktonic communities, where it was found in $\sim 80\%$ of time series.

The biodiversity of microbial communities such as plankton is overwhelmingly sustained by the "rare biosphere" revealed by recent methods of high-throughput genomic sequencing (9–11)—an extreme instance of a near-universal observation that ecological communities harbor a few highly abundant, dominant species, and a much larger number of low-abundance, rare species. In plankton protist communities sampled from multiple distant locations in the world oceans, the number of rare species increases as a power-law as lower abundances are considered, a pattern that is quantitatively uniform

Significance

A prominent feature of ecological communities is that a few species are abundant while most are rare. Using a standard community model, in which species interactions are assigned fixed random values, we show that chaos is a generic outcome if interactions are strong and immigration prevents extinction. Each species then alternates, in an effectively stochastic way, between long periods of rarity and shorter periods of high abundance; yet the overall distribution of species abundances remains conserved and qualitatively consistent with observations in marine plankton protists. Our model results contribute to a rekindled debate about the role of chaos in ecological communities.

across samples despite their strong compositional differences (12). In addition to spatial variations, strong temporal turnover has been observed for plankton, where species abundances can change dramatically on a short time scale, even when abiotic conditions do not vary substantially (13–15). A role for intrinsic ecological dynamics in driving such complex oscillations is supported by mesocosm experiments, where sustained abundance fluctuations have been observed even under stable external conditions, both for plankton (16, 17) and other microbes (18, 19).

The conditions enabling ecological chaos can be investigated with mathematical models. Traditionally, models of population dynamics have considered only a handful of taxa. There, chaos tends to occur only within a narrow parameter range (20, 21). In contrast, high-dimensional dynamical systems (involving dozens or hundreds of interacting degrees of freedom) seem to display chaos more generically (22). Robust fluctuating states (variably identified as chaos) were found in models of species-rich communities with competitive (23–25), predator–prey (26, 27), or consumer–resource (28) interactions. Some studies reported that chaotic regimes tended to sustain a higher diversity than equilibria due to the availability of more (spatio-)temporal niches (20, 24, 27, 29). Nonetheless, stabilizing mechanisms are required to prevent large abundance fluctuations from causing diversity-limiting extinctions. A metacommunity structure (a network of patches connected through dispersal) offers one plausible solution. Under conditions where patches' abundance dynamics do not synchronize, local extinction can be compensated by migration from another patch, and the fluctuations persist on time scales much longer than the local dynamics (24–26, 30).

Here, we are instead interested in characterizing the within-patch chaotic dynamics in order to relate two complementary perspectives: that of the fluctuating abundance time series of individual species and that of local community-level statistics—such as the instantaneous distribution of abundances across species and the overall strength of interactions. We consider a general model for large communities where strong ecological interactions encompass, as in microbial ones, vigorous competition between the composing species, but also facilitation. We simplify spatially structured models by considering a local, well-mixed community with a constant, small immigration. Like a "seed bank" (31) or the effect of metacommunity dispersal, this prevents irreversible loss of species. Following a well-established "disordered" approach to complex communities (32–35), we consider pairwise interactions drawn from a random distribution but focus on the little-studied regime where only a handful of species can dominate the community at any time, while most other species are rare, consistent with the rare biosphere pattern.

We show that in this general setting, a broad range of model parameters allows species to alternate chaotically between rarity and abundance on a characteristic timescale such that the community composition moves through a succession of low-diversity states. The distribution of abundances attained by any given species over a long time series largely overlaps with the distribution of abundances found in the whole community at any given time, which is a power-law across many orders of magnitude in abundance values. This correspondence suggests an equivalence among different species despite their clear ecological differences and short-term competitive exclusion dynamics. We derive a stochastic focal-species model that captures, in a statistical sense, the dynamical features common to all species, and also identify the origin of species-specific deviations in the propensity to dominate the community.

1. Model

We describe a community of S species by their time-dependent absolute abundances $x_i(t)$, with $i = 1, 2, \ldots, S$ the index of a species. Microbial communities have been described by deterministic equations where changes in abundance relate to competition within species and pairwise species interactions (19, 36, 37). According to the Lotka–Volterra equations (38), the abundance of any species in isolation grows logistically: If initially the species is rare, its abundance grows exponentially at a maximum rate r, doubling every $(\ln 2)/r$ time units. Eventually, it saturates to a carrying capacity K set by resources, predation, and abiotic conditions, assumed constant and not modeled explicitly. For simplicity, we set r and K to unity for all species but discuss heterogeneity in these parameters in *SI Appendix*, Note S3. The interaction coefficients α_{ij} (real numbers) quantify the effect of species j on the growth rate of species i, by convention, detrimental when $\alpha_{ij} > 0$, and facilitative when $\alpha_{ij} < 0$. We include a small rate of immigration $\lambda \ll 1$ into the community, constant and equal for each species, to set a lowest level of rarity and prevent extinctions. Abundances thus change in time as:

$$\dot{x}_i(t) = x_i(t)\left(1 - x_i(t) - \sum_{j=1(\neq i)}^{S} \alpha_{ij}x_j(t)\right) + \lambda. \quad [1]$$

In species-rich communities, the number of potential interactions—$S \times S$—is very large, and their values hard to estimate in natural settings. A classic approach is therefore to model the set of interaction coefficients as a realization of a random interaction matrix A (32–35). When S is large, patterns of ecological interest are expected to depend on the summary statistics of A rather than its particular realization. We consider for simplicity Gaussian statistics $A_{ij} \sim \mathcal{N}(\mu, \sigma^2)$ ($i \neq j$). A correlation γ between diagonally opposed elements can be introduced, biasing interactions toward predator–prey ($\gamma = -1$) or symmetric competition ($\gamma = 1$); here, we focus on independent interaction coefficients ($\gamma = 0$) and discuss other cases in *SI Appendix*, Fig. S5.

The interaction coefficients for distinct species i, j can be represented in terms of the mean μ and SD σ of the interaction matrix, as:

$$\alpha_{ij} = \mu + \sigma z_{ij}, \quad [2]$$

where the z_{ij} are realizations of random variables with zero mean and unit variance. We note that, by convention, we have separated the self-interaction term from the intra-specific interaction terms in Eq. 1. The diagonal element α_{ii} therefore does not appear in the sum and is not defined.

Eq. 1 with randomly sampled interactions defines the disordered Lotka–Volterra (dLV) model. By tuning the ecological parameters S, μ, σ, λ, it exhibits a number of distinct dynamical behaviors which have been thoroughly explored in weak-interaction regime, where the interaction between any particular pair of species is negligible, but a species' net competition term from all other species is comparable to its (unitary) self-interaction. If species are near their carrying capacities, the net competition is approximately:

$$\sum_{j(\neq i)} \alpha_{ij} = S\mu + \sqrt{S}\sigma\, z_i, \quad [3]$$

where the net interaction bias:

$$z_i := \frac{1}{\sqrt{S}}\sum_{j(\neq i)} z_{ij} \quad [4]$$

is a realization of a random variable $\sim \mathcal{N}(0, 1)$. A finite net competition in the limit of a large species pool requires:

$$\mu = \frac{\tilde{\mu}}{S}, \quad \sigma^2 = \frac{\tilde{\sigma}^2}{S}, \qquad [5]$$

where $\tilde{\mu}, \tilde{\sigma}$ do not grow with S. Under this scaling, methods from statistical physics [dynamical mean-field theory (34, 39–41), random matrix theory (32, 42), and replica theory (43, 44)] allow exact analytical results in the limit of $S \to \infty$, although in practice $S \sim 100$ is sufficient for good agreement between theory and simulations. Sharp boundaries were shown to separate a region where species coexist at a unique equilibrium and one with multiple attractors, including chaotic steady states (34, 39–41).

Since we are here interested in the scenario of large differences in species abundance (rare biosphere pattern) and rapid turnover dynamics, we instead consider the strong-interaction regime where the statistics of the interaction matrix do not scale with species richness S according to Eq. 5. For $S\mu \gg 1$, the overall competitive pressure makes it impossible for all species to simultaneously attain abundances close to their carrying capacities. Abundant species tend to exclude one another, resulting in instability and complex community dynamics. Arguably, strong interactions are more plausible than weak ones for microbial communities, where metabolic cross-feeding, toxin release, phagotrophy, and competition over limited nutrients lead species to depend substantially on one another's presence (45, 46).

2. Results

In the strong-interaction regime, numerical simulations of the disordered Lotka–Volterra model show that the community can display several different classes of dynamics, from equilibrium coexistence of a small subset of species, to different kinds of oscillations, including chaos. In Sections 2.1–2.4, we focus on the reference value of the interaction statistics ($\mu = 0.5, \sigma = 0.3$) representative of chaotic dynamics, and describe its salient features. In Subsections 2.5–2.6, we describe how the dynamics depends qualitatively on the statistical parameters μ and σ. Unless otherwise stated, simulations use $S = 500$ and $\lambda = 10^{-8}$. Further details on the numerical implementation are presented in section 4.1.

2.1. A Chaotic Turnover of Rare and Abundant Species. For a broad range of parameters in the strong-interaction regime, the community undergoes a chaotic turnover of dominant species. As illustrated by the time series of stacked abundances in Fig. 1A, the overwhelming share of the total abundance at any given time is due to just a few species. Which species are abundant and which are rare changes on a characteristic timescale, $\tau_{\text{dom}} \approx 30$ time units, comparable to the time it would take an isolated species to attain an abundance on the order of its carrying capacity starting from the lowest abundance set by immigration. While the total abundance fluctuates moderately around a well-defined time average, individual species follow a "boom-bust" dynamics. If this simulation represented a natural microbial community, only the most abundant species—that we call the dominant component of the community—would be detectable by morphological inspection or shallow sequencing.

We wish to characterize the dominant component and understand how it relates to the pool of rarer species. In order to
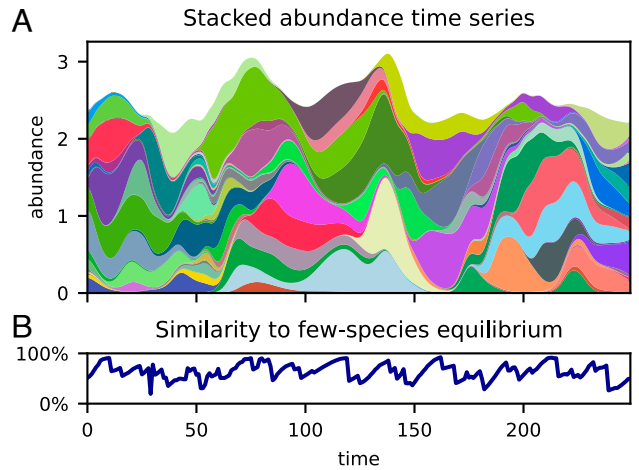

Fig. 1. Turnover of the dominant component. (*A*) The stacked abundances of all species under steady-state conditions: There is a turnover of species such that only the dominant component is visible at any given time (each species has a distinct random color). (*B*) Bray–Curtis index of community composition similarity between the dominant component of the community at time *t*, and the composition if it were isolated from the rare species and allowed to reach equilibrium: The community appears to approach the composition of few-species equilibria before being destabilized by invasion from the pool of rare species.

quantify the notion of dominance, we define the effective size of the community as Simpson's (reciprocal) diversity index (47),

$$S_{\text{eff}}(t) := \frac{1}{\sum_i p_i^2(t)}, \qquad [6]$$

where $p_i = x_i / \sum_j x_j$ denote relative abundances. S_{eff} approaches its lowest possible value of 1 when a single species is responsible for most of the total abundance, and its maximum S when all species have similar abundances. Its integer approximation provides the richness, i.e. , number of distinct species, of the dominant component.

The effective size S_{eff} of the community in our reference simulation fluctuates around an average of nine dominant species, which make up 90% of the total abundance. The relative abundance threshold for a species to be in the dominant component fluctuates around 3%, which is comparable to the arbitrary 1%-threshold used in empirical studies (48). In *SI Appendix*, Fig. S4, we show that the number of dominant species grows slowly (but super-logarithmically) with S, up to about 15 for $S = 10^4$. Thus, strong interactions limit the size of the dominant component, and the vast majority of species are rare at any point in time.

The turnover of dominant species is not periodic; indeed, even over a large time-window, where every species is found on multiple occasions to be part of the dominant component, its composition never closely repeats (*SI Appendix*, Fig. S3). This aperiodicity suggests the presence of chaotic dynamics. We give numerical evidence for sensitive dependence on initial condition and positive maximal Lyapunov exponent in *SI Appendix*, Figs. S1 and S2. The turnover dynamics has the character of moving, chaotically, between different quasi-equilibria corresponding to different compositions of the dominant community [cf. "chaotic itinerancy" (49)]. To reveal this pattern, we measure a "closeness-to-equilibrium," defined as the similarity in composition between the observed dominant component at a given time, and the equilibrium that this dominant component would converge to if it were isolated from the rare component and allowed to

equilibrate. As a similarity metric, we use the classical Bray–Curtis index (section 4.2), which has also been used to measure variations in community composition in plankton time series (13). In Fig. 1*B*, we see that the similarity at times slowly approaches 100%, followed by faster drops, toward about 50%, indicating the subversion of a coherent dominant community by previously rare invaders.

The fact that the community composition is not observed to closely repeat is arguably due to the vast number of possible quasi-equilibria that the chaotic dynamics can explore. In the weak-interaction regime, a number of unstable equilibria exponential in S has been confirmed (50, 51). It is therefore conceivable that the number of quasi-equilibria in our case is also exponentially large. The LV equations for $\lambda = 0$ admit up to one coexistence fixed point (not necessarily stable) for every chosen subset of species (38). Hence, we expect on the order of $\sim S^{S_{\mathrm{eff}}}$ quasi-equilibria, which for $S = 500$ and $S_{\mathrm{eff}} \approx 9$ evaluates to 10^{24}. If the dynamics explores the astronomical diversity of such equilibria on trajectories which depend sensitively on the initial conditions, the dominant component may look as if having been assembled "by chance" at different points in time.

The composition of the dominant community is not entirely arbitrary, though. While the abundance time series of most pairs of species have negligible correlations, every species tends to have a few other species with a moderate degree of correlation. In particular, if $(\alpha_{ij} + \alpha_{ij})/2$ is significantly smaller than the expectation μ, and hence species i and j are close to a commensal or mutualistic relationship, these species tend to "boom" one after the other (*SI Appendix*, Fig. S6).

2.2. Species' Abundance Fluctuations Follow a Power-Law.
In a common representation of empirical observations, where relative abundances are ranked in descending order a rank–abundance plot (52), microbial communities display an overwhelming majority of low-abundance species (10). Our simulated community reproduces this feature; Fig. 2*A*. The exact shape of the plot changes in time, as does the rank of any particular species, but the overall statistical structure of the community is highly conserved. An alternative way to display the same data is to bin abundances and count the frequency of species occurring within each bin, producing a species abundance distribution (SAD) (52). The histogram in Fig. 2*B* illustrates the "snapshot" SAD for the rank-abundance plot in Fig. 2*A* of abundances sampled at a single time point. Whenever observations are available for multiple time points, it is also possible to plot, for a given species, the histogram of its abundance in time. As time gets large (practically, we considered 100'000 time units after the transient), the histogram converges to a smooth distribution, that we call the abundance fluctuation distribution (AFD) (53). Its average shape across all species is also displayed in Fig. 2*B*.

Several conclusions can be drawn by comparing SADs and AFDs. First, a snapshot SAD appears to be a subsampling of the average AFD. Therefore, SADs maintain the same statistical structure despite the continuous displacement of single species from one bin to another. Second, every species fluctuates in time between extreme rarity ($x \approx \lambda = 10^{-8}$) and high abundance ($x \gtrsim 10^{-1}$). This variation is comparable to that observed, at any given time, between the most abundant and the rarest species. Third, species are largely equivalent with respect to the spectrum of fluctuations in time, as indicated by the small variation in AFDs across species. We evaluate the regularities and differences of single-species dynamics more thoroughly in Subsection 2.4.
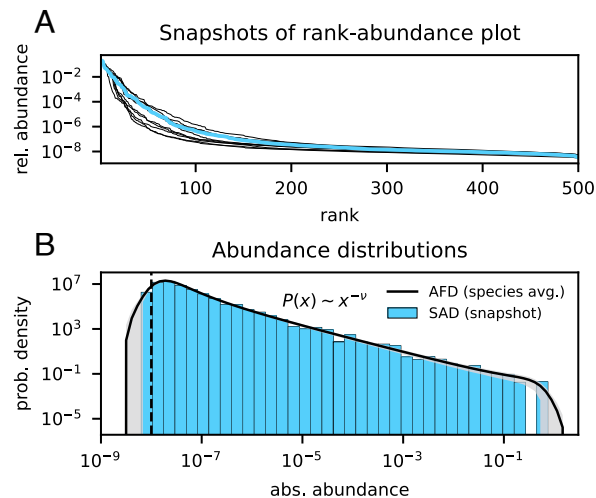
Fig. 2. Statistical features of abundance variations across species and in time. (*A*) Snapshot rank-abundance plot for the relative abundances in the reference simulation: Most species have orders of magnitude smaller abundances than the top ranks. Different lines represent observations at well-separated time points. (*B*) Species abundance distribution (SAD, blue histogram) corresponding to the blue rank-abundance plot; overlaid, abundance fluctuation distribution (AFD), averaged over all species (black line) with ± one SD across species shaded in gray: The snapshot SAD appears to be a subsampling of the average AFD, indicating an equivalence, but de-synchronization, of species in their abundance fluctuations. The one bar missing from the SAD is the effect of finite species richness, as high-abundance bins only ever contain a couple of species for $S = 500$. The vertical dashed line indicates the immigration level which determines a lower limit to abundances.

The most striking feature of these distributions, however, is the power-law $x^{-\nu}$ traced for intermediate abundances. This range is bounded at low abundances by the immigration rate and at high abundances by the single-species carrying capacity. The power-law exponent is $\nu \approx 1.18$ for the reference simulation, but it varies in general with the ecological parameters, as we discuss further in the following sections.

The regularity of the abundance distributions across species suggests that it may be possible to describe the dynamics of a "typical" species in a compact way—this is the goal of the next section.

2.3. A Stochastic Focal-Species Model Reproduces Boom-Bust Dynamics.
Fluctuating abundance time series are often fitted by one-dimensional stochastic models (7); for example, stochastic logistic growth has been found to capture the statistics of fluctuations in a variety of datasets on microbial abundances (53, 54). The noise term encapsulates variations in a species' growth rate whose origin may not be known explicitly. In our virtual Lotka–Volterra community, once the interaction matrix and initial abundances have been fixed, there is no uncertainty; nonetheless, the chaotic, high-dimensional dynamics results in species' growth rates fluctuating in a seemingly random fashion. We are therefore led to formulate a model for a single, focal species, for which explicit interactions are replaced by stochastic noise. Because we have found species to be statistically similar, its parameters do not depend on any particular species, but reflect the effective dynamics of any species in the community.

Following dynamical mean-field-like arguments and approximations informed by our simulations (section 4.5), we derive the focal-species model:

$$\dot{x}(t) = x(t)\left(g(t) - x(t)\right) + \lambda, \qquad [\textbf{7a}]$$
$$g(t) = -k + u\,\eta(t), \qquad [\textbf{7b}]$$

where $g(t)$ is a stochastic growth rate with mean $-k$, and fluctuations of magnitude u and correlation time τ. The process $\eta(t)$ is a colored Gaussian noise with zero mean and an autocorrelation that decays exponentially;

$$\langle \eta \rangle = 0, \quad \langle \eta(t)\,\eta(t') \rangle = e^{-|t-t'|/\tau}, \quad \text{[8]}$$

where brackets denote averages over noise realizations. The connection between the ecological parameters S, μ, σ, λ and the resulting dynamics of the disordered Lotka–Volterra model in the chaotic phase is then broken down into two steps: how the effective parameters k, u, τ relate to the ecological parameters and how the behavior of the focal-species model depends on the effective parameters.

For the first step, we find

$$k = \mu \overline{X} - 1 \quad \text{and} \quad u = \sigma \frac{\overline{X}}{\sqrt{\overline{S}_{\text{eff}}}}, \quad \text{[9]}$$

where X is the total community abundance of the original dynamics Eq. 1, the effective community size S_{eff} is as in Eq. 6, and an overline denotes a long-time average. Eq. 9 relates the focal species' growth rate to the time-averaged net competition ($\approx \mu \overline{X}$) from all other species. We find in simulations of Eq. 1 in the chaotic phase that competition is strong enough to make $k > 0$. The second relation captures the variation in the net competition that a species experiences because of turnover of the dominant community component. Due to sampling statistics, this variation is larger when the dominant component tends to have fewer species; hence, the dependence on $(\overline{S}_{\text{eff}})^{-1/2}$. The third effective parameter, the timescale τ, controls how long the focal species stays dominant, once a fluctuation has brought it to high abundance. This timescale is essentially equal to the turnover timescale τ_{dom} of the dominant component (defined more precisely by autocorrelation functions in section 4.5). In the weak-interaction regime, where any pair of species can be treated as effectively independent at all times, self-consistency relations such as $S\langle x \rangle = \overline{X}$ allow to implicitly express the focal-species model in terms of the ecological parameters. For strong interactions, however, the disproportionate effect of the few dominant species on the whole community invalidates this approach; we therefore relate the effective parameters to the community-level observables $\overline{X}, \overline{S}_{\text{eff}}, \tau_{\text{dom}}$ which are obtained from simulation of Eq. 1 at given values of the ecological parameters.

For the second step, we would like to solve Eq. 7 for general values of the effective parameters. While this is intractable due to the finite correlation time of the noise, the equations can be simulated and treated by approximate analytical techniques. In Fig. 3A, we compare the time series of an arbitrary species in the dLV model with a simulation of the focal-species model. By eye, the time series appear statistically similar. The typical abundance of a species can be estimated by replacing the fluctuating growth rate in Eq. 7 with its typical value (i.e. $\eta = 0$), yielding the equilibrium λ/k if $k > 0$, as indeed confirmed by the simulation. Thus the typical abundance value is on the order of the immigration threshold. Fig. 3B shows that the average AFD of the dLV agrees remarkably well with the stationary distribution of the focal-species model, in particular for the power-law section. Using the unified colored noise approximation (55) (section 4.6), one predicts that the stationary distribution, for $\lambda \ll x \ll 1$, takes the power-law form $x^{-\nu}$, where the exponent

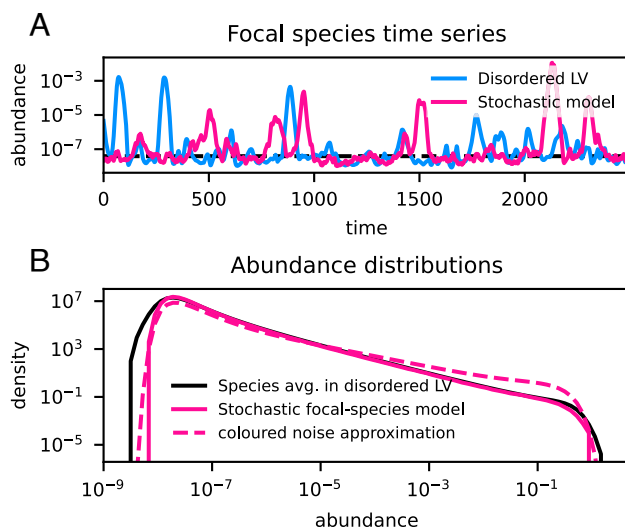$$\nu = 1 + \frac{k}{u^2 \tau} \quad \text{[10]}$$

Fig. 3. Comparison of the stochastic focal-species model and the chaotic dLV model. (A) Time series of one arbitrary species in the disordered Lotka–Volterra (dLV) model (blue), and one realization of the stochastic focal-species model Eq. 7 with parameters as in Eq. 9: The time series are statistically similar. (B) Comparison of the average abundance fluctuation distribution (AFD) from Fig. 2 (black), and the AFD of the focal-species model (pink): Excellent agreement is found for the power-law section. The "unified colored noise approximation" solution for the focal-species model's AFD (dashed, pink line) predicts the correct overall shape of the distribution, but not a quantitatively accurate value for the power-law exponent.

is strictly larger than one—the value predicted for weak interactions (41) and for neutral models (56). Even if Eq. 10 is not quantitatively precise (Fig. 3B), this formula suggests a scaling with the effective parameters that we will discuss later on.

2.4. Species with Lower Net Competition Are More Often Dominant. The similarity of all species' abundance fluctuation distributions in Fig. 2 is reflected in the focal-species model's dependence on collective properties like the total abundance. However, the logarithmic scale downplays the variance between species' AFDs, particularly at higher abundances. Indeed, while all abundances fluctuate over orders of magnitude, some species are observed to be more often dominant (or rare). Such differences are reminiscent of the distinction between "frequent" and "occasional" species observed in empirical time series (57, 58).

In order to assess the nature of species differences in simulations of chaotic dLV, we rank species by the fraction of time spent as part of the dominant component. Observing the community dynamics on a very long timescale of tens of thousands of generations (400 times longer than in Fig. 1), the first-ranked species appears to boom much more often than the last (Fig. 4A). The frequency of a species is chiefly determined by the number of booms rather than their duration, which is comparable for all species. The median dominance time decreases with the total species richness (Fig. 4B): A doubling of S leads to each species halving its dominance time fraction. As the community gets crowded—while its effective size hardly increases, as remarked in Subsection 2.1—all species become temporally more constrained in their capacity to boom. Yet some significant fraction of species is biased toward booming much more often or rarely than the median, regardless of community richness. We quantify this trend by plotting in Fig. 4C the dominance bias—the dominance time fraction normalized by the median across all species—against the relative rank (i.e., rank divided by S). For high richness ($S \sim 10^3$), the distribution of bias converges toward
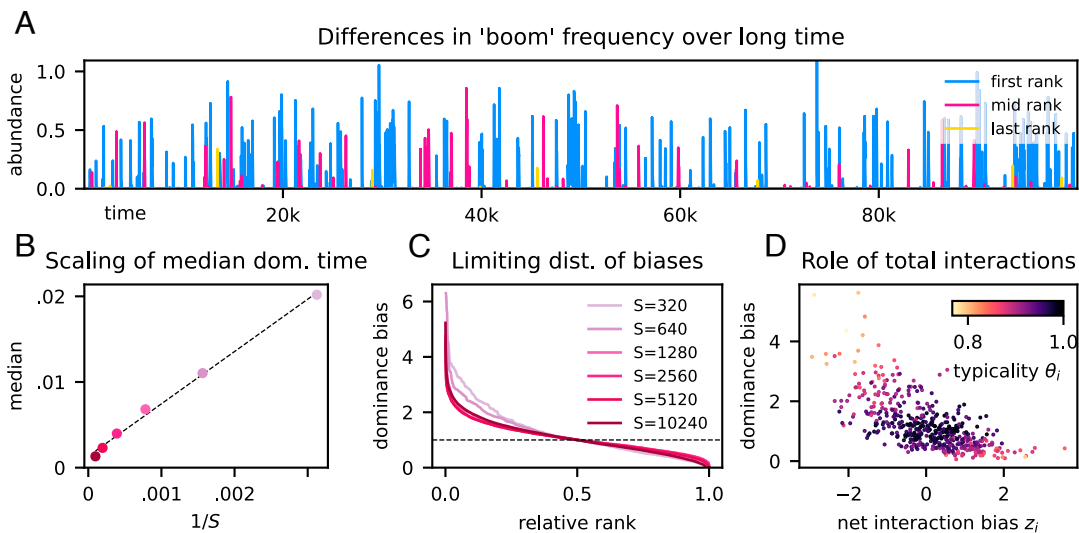
Fig. 4. Species differences in dominance. (*A*) Example of a long abundance time series for the three species who are ranked first, median, and last, with respect to the "dominance bias" (fraction of time spent in the dominant component relative to the species median). Some species "boom" more often than others. (*B*) The scaling of median fraction of time spent in the dominant component against reciprocal species pool size: Increasing S results in a proportional decrease in median dominance time. (*C*) Distribution of dominance biases against relative dominance rank for a range of S: there appears to be convergence toward a nonconstant limiting distribution, implying that net species differences are not due to small-S effects. Note that, by definition, the dominance bias is 1 for the middle rank, indicated by the dashed line separating positively from negatively biased species. (*D*) Scatter of dominance bias against the net interaction bias, z_i Eq. **4**: Lower net competition correlates with higher dominance bias. Species in the tails of the z_i distribution are also less "typical," with typicality quantified by the index θ_i, Eq. **16**, representing the similarity of a species AFD to the species-averaged AFD. Panel *A* and *D* are both for $S = 500$.

a characteristic, nonlinearly decreasing shape, where the most frequent species occur more than four times as often as the median, and the last-ranked species almost zero.

The persistence of inter-species differences with large S may seem to contradict the central limit theorem, as species' sets of interaction coefficients converge toward statistics that are identical for every species. In the chaotic regime, however, even the smallest differences in growth rates get amplified during a boom. As we show in section 4.4, if Eq. **1** is rewritten in terms of the proportions p_i, the relative advantage of species i is quantified by a selection coefficient whose time average scales as $-S^{-1/2}\sigma z_i$. Correspondingly, the relative, time-averaged growth rate is proportional to the net interaction bias z_i defined in Eq. **4**, resulting in species with larger z_i to have positive dominance bias (Fig. 4*D*). Outliers of the scatter plot, i.e. species that have particularly high or low dominance ranks, are also the species whose AFD is furthest from the average AFD of the community, as quantified by the typicality index $\theta_i \in [0, 1]$, defined in section 4.2.

In conclusion, the relative species-to-species variation in the total interaction strength drives the long-term differences in the dynamics of single species in the community. While the focal-species model emphasizes the similarity of species, species differences can also be taken into account by employing species-specific effective parameters. In particular, replacing k with a distribution of k_i's would create a dominance bias, and is in fact motivated upon closer examination of our focal-species model derivation (Fig. 7*D* in section 4.5).

2.5. Interaction Statistics Control Different Dynamical Phases. Hitherto, we have focused on reference values of the interaction statistics μ and σ that produce chaotic turnover of species abundances. We now broaden our investigation to determine the extent of validity of our previous analysis when the interaction statistics are varied. For every pair of (μ, σ) values, we run 30 independent simulations, each with a different sampling of

the interaction matrix and uniformly sampled abundance initial condition. After a transient has elapsed, we classify the trajectory as belonging to one of four different classes: equilibrium, cycle, chaos, or divergence. Fig. 5 displays the probability of observing chaos, demonstrating that it does not require fine tuning of parameters, but rather occurs across a broad parameter range.

The parameter region where chaos is prevalent, the "chaotic phase," borders on regions of qualitatively different community dynamics. For small variation in interaction strengths (below the line connecting $(0, \sqrt{2/S})$ to $(1, 0)$), the community has a unique, global equilibrium that is fully characterized for weak interactions (cf. Fig. 2 of ref. 34). The transition from equilibrium to chaos has been investigated with dynamical mean-field theory (41). For low interaction variance, but with mean exceeding the unitary strength of intra-specific competition, a single species comes to dominate, as expected by the competitive exclusion principle (59). Adiabatic simulations, implemented by continuously rescaling a single realization of the interaction matrix (details in *SI Appendix*, Fig. S8), reveal that lines radiating from the point $(\mu, \sigma) = (1, 0)$ separate sectors where stable fixed points have different numbers of coexisting species. Traversing these sectors anti-clockwise, S_{eff} increases by near-integer steps from one (full exclusion) up to about 8. From thence, a sudden transition to chaos occurs at the dashed line in Fig. 5. We note, however, that the parameter region between chaos and competitive exclusion contains attractors of different types: cycles and chaos, coexisting with multiple fixed points, resulting in hysteresis (*SI Appendix*, Fig. S8*B*). This "multiple attractor phase" (34, 41) is a complicated and mostly uncharted territory whose detailed exploration goes beyond the scope of this study. Finally, for large variation in interactions, some abundances diverge due to the positive feedback loop induced by strongly mutualistic interactions, and the model is biologically unsound.

Across the phase diagram, community-level observables such as the average total abundance $\overline{X}$ and effective community size $\overline{S}_{\text{eff}}$ vary considerably (*SI Appendix*, Fig. S9). The weak-interaction
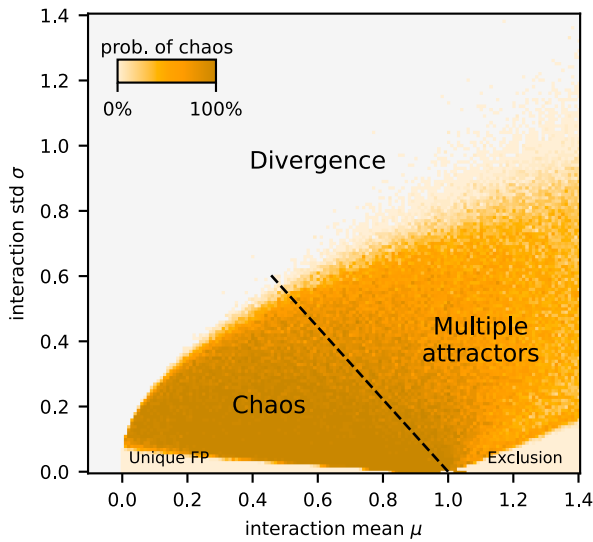
Fig. 5. Dynamical phases of the disordered Lotka–Volterra model as a function of the interaction mean and SD. Color indicates the probability of persistent chaos in long-time simulations: For each μ and σ (with 0.01 increment), 30 simulations were made, each with a different random initial condition $x_i \sim U(\lambda, 2/S)$ and realization of the interaction matrix. Parameters yielding divergence every time are marked with gray. The boundary separating the chaotic phase from the rest of the multiple-attractor phase (in which cycles and multi-stable fixed point are common in addition to chaos) is not sharp, unless probed adiabatically in the way explained in *SI Appendix*, Fig. S8. The unique fixed point phase has been studied analytically in the weak-interaction regime ($\mu \sim 1/S$). When inter-specific competition is in general stronger than intra-specific competition, a single species (identity depending on initial condition) dominates, in line with the classical competitive exclusion principle (59).

regime (whether in the equilibrium or chaotic phase) allows for high diversity, so $\overline{X}$ and $\overline{S}_{\mathrm{eff}}$ are of order S; strong interactions, on the other hand, imply low diversity, with $\overline{S}_{\mathrm{eff}}$ and $\overline{X}$ of order unity. An explicit expression for how these community-level observables depend on the ecological parameters (S, μ, σ, λ) is intractable although implicit formulas exist in the weak-interaction regime (34). Nonetheless, an approximate formula that we derive in section 4.3 allows to relate community-level observables to one another and to μ and σ:

$$\overline{X} \approx \left[\mu + \frac{1 - \mu}{\overline{S}_{\mathrm{eff}}} - \sigma\overline{\rho} \right]^{-1}, \qquad \textbf{[11]}$$

in which we introduce the collective correlation

$$\overline{\rho} := -\sum_{ij} z_{ij}\overline{p_i p_j}, \qquad \textbf{[12]}$$

involving the time-averaged product of relative abundances weighted by their normalized interaction coefficient Eq. **4**. By construction, the collective correlation is close to zero when all species abundances are uncorrelated over long times, as would follow from weak interactions. On the contrary, it is positive when pairs of species with interactions less competitive than the average tend to co-occur, and/or those with more competitive interactions tend to exclude one another.

Eq. **11** is particularly useful in understanding the role of correlations in the chaotic phase. As we observed in Subsection 2.3, the effective parameter $k = \mu\overline{X} - 1$ is positive in the chaotic phase, implying that the growth rate of a species is typically negative, and abundances are therefore typically on the order of

the small immigration rate rather than near carrying capacity. The existence of these two "poles" of abundance values is key to boom-bust dynamics. By combining $k > 0$ with Eq. **11**, we estimate a minimum, critical value of the collective correlation required for boom-bust dynamics:

$$\rho_c = \frac{1 - \mu}{\sigma} \frac{1}{\overline{S}_{\mathrm{eff}}}. \qquad \textbf{[13]}$$

Numerical simulations demonstrate that $\overline{\rho} \gtrsim \rho_c$ in the chaotic phase, where the critical value is approached at the boundary with the unique-equilibrium phase (*SI Appendix*, Fig. S11). With this result in hand, Eqs. **11** and **13** establish that $\overline{X} \gtrsim 1/\mu$ in the chaotic phase. For strong interactions, total abundances are predicted to be of order one, and for weak interactions $\overline{X} \approx S/\tilde{\mu}$ (recall Eq. **5**), which recovers the observed scalings of these observables. As one moves deeper into the chaotic phase, the collective correlation increases continuously, as the effective community size drops, suggesting a seamless transition from a weak-interaction, chaotic regime amenable to exact treatment (41, 60), to the strongly correlated regime that we have analyzed by simulations and the approximate focal-species model.

2.6. Self-Organization between Community-Level Observables Constrains Abundance Power-Law Variation. In Subsection 2.3, we established a focal-species model depending on the effective parameters $k, u,$ and τ, that were related to the ecological parameters S, μ, σ, λ indirectly via community-level observables $\overline{X}, \overline{S}_{\mathrm{eff}}, \tau_{\mathrm{dom}}$. Furthermore, in the previous section, we studied how the latter vary in the chaotic phase. Putting these results together, we here examine the corresponding variation of the effective parameters and of the focal-species model's predictions.

Because the trio k, u, τ ultimately derives from only two independent variables, μ, σ (considering fixed S, λ), they must be dependent. Fig. 6A demonstrates that, across the chaotic phase, an approximate linear relationship holds between k and u, as well as between u and τ. Because k and u are related to the mean and the variance of abundances via Eq. **9**, their proportionality is reminiscent of the empirical Taylor's law which posits a power-law relation between abundance mean and variance as they vary across samples (61). The slope of the relationship of u to k is close to one (and varying little with S and λ; *SI Appendix*, Fig. S10), which implies with Eq. **9** that:

$$\overline{X} \approx \left[\mu - \frac{\sigma}{\sqrt{\overline{S}_e}} \right]^{-1}. \qquad \textbf{[14]}$$

Comparison to Eq. **11** then yields that $\rho - \rho_c \approx \overline{S}_{\mathrm{eff}}^{-1/2}$. This empirical relationship thus supports the aforementioned convergence—in the limit where $\overline{S}_{\mathrm{eff}}$ is large, as for weak interactions—of the collective correlation to its critical value.

We find in Fig. 6 that the slope ν_{foc} of the power-law trend obtained from simulation of the focal-species model finds good agreement with the value ν from the full dLV model. There is a narrow overall variation of the exponent, a consequence of the interdependency of the effective parameters. As can be intuited by the approximate expression Eq. **10** for the focal-species model, the exponent is strictly larger than 1, a value it approaches if the turnover time scale diverges, as indeed it does on the boundary to the unique equilibrium phase. The exponent increases as interactions become more competitive, up to about 1.4 at $(\mu, \sigma) = (1, 0)$. However, the exponent also depends on S and λ, showing a constant slope against $\log S$ or $-1/\log \lambda$ (*SI Appendix*, Fig. S7).
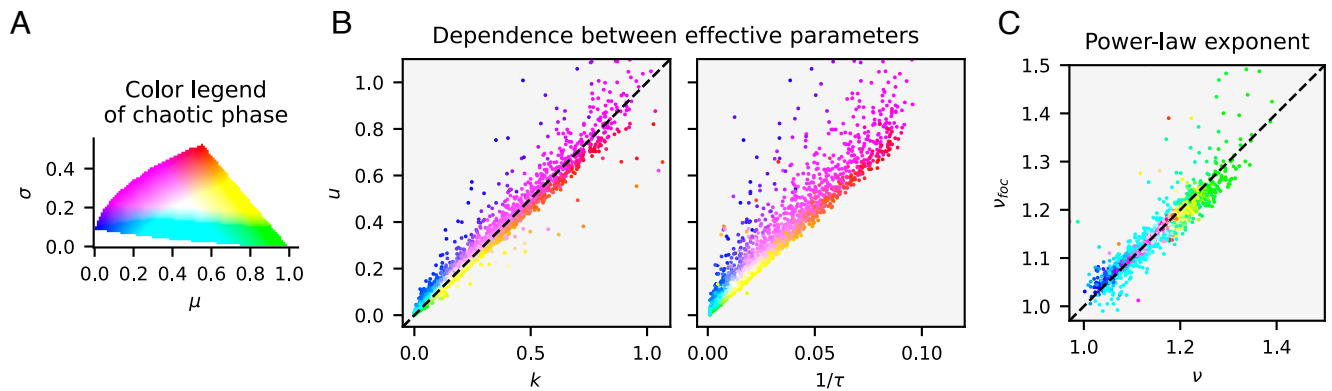
Fig. 6. Relations between effective parameters in the chaotic phase. (*A*) Color legend of the chaotic phase (boundaries from Fig. 5). Each pair of (μ, σ) has been mapped to a distinct color. (*B*) Co-dependence of the effective parameters u, k, τ: the amplitude u of growth-rate fluctuations approximately equals the absolute value k of the negative growth rate (only weakly depending on λ and S; *SI Appendix*, Fig. S10); u is roughly proportional to the inverse turnover time, but the slope of the relationship depends on λ and S. (*C*) The exponent ν of the power-law section of the AFD for the chaotic dLV model plotted against the analogue ν_{foc} obtained for the focal-species model: generally good agreement is found, with more outliers for parameters close to phase boundaries. A few outliers lie beyond the plotted range. Exponents have been estimated by fitting a power-law in the interval [100λ, 0.01] of the abundance distribution.

3. Discussion

Following growing empirical evidence for the presence of ecological chaos in natural and synthetic communities (7, 19), and increasing interest in the role of the rare biosphere (10, 62), we have sought a connection between the two through a minimal model of community dynamics: The disordered Lotka–Volterra (dLV) model with strong interactions and weak immigration. Our analysis of this model by extensive simulations, and through the derivation of an effective focal-species model, showed that first, persistent chaos arises generically and can drive fast and extensive turnover of rare and abundant species; second, a statistical equivalence between species emerges such that a single focal species' fluctuation statistics predict the largely invariant power-law abundance distributions; third, deviations from this equivalence are associated with species differences in frequency of occurrence. In the following, we discuss the generality of these results and their interpretation in the context of plankton ecology.

The chaotic turnover of rare and abundant species occurs because every subset of species that could stably coexist at high abundances is invadable by some rare species. This phenomenon should be robust to generalizations of the model as long as the dominant component remains exposed to a sufficient diversity of potential invaders and the niche space that underlies species interactions contains enough trade-offs that no species can be a superior competitor across many biotic contexts. Our simplifying assumptions such as uniform growth rates and carrying capacities, and uncorrelated interactions can be relaxed (see our limited explorations in *SI Appendix*, Fig. S5). Additional sources of modest noise should not cancel the deterministic contributions to fluctuations; indeed, the dynamical phases we have indicated are qualitatively similar to those arising in an individual-based version of the dLV model accounting for demographic stochasticity (23). On the other hand, if the connectivity of the interaction network were reduced, lowering the exposure to competitors, one might expect a loss of persistent chaos at some critical connectance value (63). Highly structured and hierarchical interactions would also undermine autonomous turnover on ecological timescales.

On a more technical note, the type of chaos we observe is likely "chaotic itinerancy" (49, 64). Lotka–Volterra systems without immigration admit heteroclinic networks (65–67), equilibria with stable and unstable directions (i.e. saddle points) connected by orbits. Without immigration, such saddles are found on the system boundary, corresponding to some subset of species

being extinct—in our case, these are the low-diversity equilibria reflected in the dominant component. The chaotic attractors appear when the saddles are "pushed off" the boundary by the immigration term. Consistent with chaotic itinerancy in the dLV, characteristics of heteroclinic orbits—dynamical slowdown and "aging"—appear in the limit of vanishing immigration (60, 68).

While the assumption of disordered interactions may appear ad hoc, predictions for the onset of instability by the dLV model qualitatively match experiments in synthetic bacterial communities (19). In a plankton context, we take the dLV to be a minimal yet relevant phenomenological representation of the relationships between species (or "operational taxonomic units" from sequencing) of marine protists of a similar size class: The protistan interactome is largely uncharted (69), the ubiquity of mixoplankton blurs consumer–resource distinctions (70), and the effects of a diversity of zooplankton and viruses can manifest as apparent competition between species.

For rare plankton protists, the empirical snapshot SADs show a clear power-law trend, with an exponent around 1.6, varying little between different locations in the world oceans, despite large composition differences across samples (12). The unified neutral theory of biodiversity, based on the interchangeability of individuals regardless of species identity, predicts a power-law tail of the SAD with exponent one (3, 56). To approach the empirical value, previous studies augmented neutral theory with nonlinear growth rates (12) or chaotic mixing (71) to find an exponent dependent on the model parameters. However, for large census sizes such as that of plankton communities, neutral theories predict astronomically large turnover timescales (72, 73), inconsistent with observation. As we have shown, the dLV exhibits fast turnover when interactions are strong and sufficiently varied. For this model, $\nu \to 1$ as immigration tends to zero (*SI Appendix*, Fig. S7, also shown in the weak-interaction limit (28, 60)), but, if interactions are not weak, ν is substantially larger than one for small but finite values of immigration. The approximate solution to the focal-species model, Eq. **10**, shows that the positive deviation from $\nu = 1$ depends on three inter-related effective parameters: the mean, amplitude, and timescale of fluctuations in each species' net competition. As these fluctuations drive the turnover pattern, boom-bust dynamics comes to be associated with a larger-than-one exponent. The relatively weak variation of ν across the space of ecological parameters moreover suggests a reason for

the limited geographical heterogeneity of the empirical value of the exponent.

A role for chaos in the plankton has long been advocated for (29, 74). Proposed mechanisms include coupling of population dynamics to chaotically fluctuating environmental variables (75, 76), nonlinearity of low-dimensional zooplankton–phytoplankton dynamics (77), resource competition between phytoplankton species (20), the effect of marine viruses on populations of cyanobacteria strains (26). These possibilities are not mutually exclusive, but relevant at different scales, from coarser to finer levels of taxonomic resolution. Adding a degree of structure to our species-level model to represent multiple functional groups would offer a way to investigate the connection between fluctuations at different scales. Empirical findings to replicate are the weakening signal for chaos as taxa are aggregated at higher orders (8) and more dynamical regularity and predictability in succession patterns at the level of functional groups (78). In fact, even our unstructured model captures the feature that fluctuations are less severe at the aggregated level (e.g. total biomass, the envelope in Fig. 1).

Besides explicit incorporation of structured interactions, an extension of our model with particular biological relevance would be to allow interactions to evolve, notably as they are reshaped by the appearance of novel species—a different scenario than our immigration term captures. Persistent turnover can then manifest on long timescales even if the ecological dynamics is—contrary to our case—at equilibrium. Such turnover has been shown in numerical models of evolving food webs structured by body size (79, 80) and when adaptive dynamics occurs in high-dimensional trait spaces (81). An open question is what evolutionary process may produce interactions that underpin chaotic turnover on ecological timescales. The observation that evolution sustains higher diversity under boom-bust ecology than under equilibrium ecology (82), together with the propensity of diversity to cause instability, suggests a possible role for eco-evolutionary feedbacks.

Our approximate derivation of an explicit focal-species model demonstrates how ecological chaos comes to resemble noise. Parallel work to ours shows that an exact but implicitly defined effective model can be derived in the combined limit of weak interactions and infinitesimal immigration, where compositional turnover is slow (60). In our model, the effective parameters could be used in fitting observational time series. Formally, Eq. 7 is similar to heuristic stochastic single-species logistic growth models that predict empirical distributions of microbial abundances (53, 54). A notable difference lies in the negative mean growth rate we find, which together with noise-correlation and immigration yields fluctuations over many orders of magnitude, from rare to abundant. An insight from our model is that a species may be rare for an exceedingly long time, without rarity being a permanent character. On the other hand, species differences in the propensity to become abundant could reflect small differences in effective parameters that depend on a multitude of factors, which—like interaction rates—might not be individually measurable with precision. Together, these findings suggest that the abundances of particular species may not be easily explained by their traits, should fluctuations be determined by community complexity rather than a more direct coupling to environmental variables. In closing, a comparison of time series data to focal-species models could provide a complement to nonparametric methods (6, 7) in establishing the plausibility of ecological chaos as a driver of abundance fluctuations.

4. Materials and Methods

4.1. Numerical Implementation. For Lotka–Volterra simulations, we used a fixed time-step Euler scheme with $\Delta t = 0.01$, applied to the logarithm of abundances. This guarantees the positivity of all abundances at all times, regardless of immigration rate. To automatically classify the long-time behavior of trajectories as fixed-points, cycles, or chaos, we used a heuristic method of counting abundance vector recurrences, validated against visual inspection of trajectories and calculated maximal Lyapunov exponent for a subset of trajectories. Further details are given in *SI Appendix*, Note S2.

4.2. Similarity Metrics. The Bray–Curtis similarity index (83) is defined as

$$BC(\boldsymbol{x}, \boldsymbol{y}) = \sum_i w_i \frac{\min(x_i, y_i)}{\text{mean}(x_i, y_i)}, \qquad [15]$$

where w_i is the relative abundance of species i with respect to the joined abundances $\boldsymbol{x} + \boldsymbol{y}$. By definition, $BC(\boldsymbol{x}, \boldsymbol{y}) = 1$ iff $\boldsymbol{x} = \boldsymbol{y}$, and $BC \approx 0$ when, for each i, either $x_i \gg y_i$ or $y_i \gg x_i$; this makes it suitable for communities where abundances span orders of magnitude.

For the similarity graph Fig. 1*B*, we have plotted $BC(\boldsymbol{x}^{\text{dom}}(t), \boldsymbol{y}^*(t))$, where $\boldsymbol{x}^{\text{dom}}(t)$ is the restriction of $\boldsymbol{x}(t)$ in the reference simulation to only the dominant species at time t, and $\boldsymbol{y}^*(t)$ is the fixed point reached from $\boldsymbol{x}^{\text{dom}}(t)$ as initial condition, with $\lambda = 0$.

To compare the similarity the AFD of species i, $P_i(x)$, to the species-averaged AFD $P = \sum_i P_i / S$, we define the index

$$\theta_i := 1 - \sup_x |F(x) - F_i(x)|, \qquad [16]$$

where F_i and F are the cumulative distribution functions of P_i and P, respectively; i.e., the index θ_i is based on the Kolmogorov–Smirnov distance (47) of the AFDs.

4.3. Derivation of Time-Averaged Total Abundance. Direct summation of Eq. 1 over i (assuming $r_i = 1$), and then division on both sides by $X(t) = \sum_i x_i(t)$, yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \ln X(t) = 1 - X(t) \left[\mu + (1 - \mu)/S_{\text{eff}} - \sigma \rho(t) \right] + \frac{S\lambda}{X(t)} \qquad [17]$$

with S_{eff} as in Eq. 6 and $\rho(t)$ as Eq. 12 but without the time average. Taking the long-time average of Eq. 17, the left-hand side becomes $\lim_{T \to \infty} (\ln X(T) - \ln X(0))/T$, which evaluates to zero on the assumption that no species diverges in abundance. The right-hand side contains terms such as $\overline{X/S_{\text{eff}}}$ and $\overline{X\rho}$. If the relative fluctuations in X, S_{eff}, ρ are small (see *SI Appendix*, Fig. S9), or these functions are at most weakly correlated to one another, then we obtain, approximately,

$$0 = 1 - \mu\overline{X} - (1 - \mu)\overline{X}/\overline{S}_{\text{eff}} + \sigma\overline{\rho} + O(S\lambda). \qquad [18]$$

We neglect the small immigration term; solving for $\overline{X}$ then finds Eq. 11. The relative error in $\overline{X}$ between Eq. 11 for simulated values of the community-level observables in the right-hand side, and the simulated value of $\overline{X}$, is typically less than a few percent (*SI Appendix*, Fig. S14).

4.4. Selective Advantage. The dynamics of the relative abundance $p_i = x_i/X$ is found by summing and differentiating Eq. 1 as:

$$\dot{p}_i = X p_i \left\{ \sum_j p_j^2 - p_i - \sum_j p_j \left(\alpha_{ij} - \sum_k \alpha_{jk} p_k \right) \right\} + \frac{\lambda S}{X} \left(\frac{1}{S} - p_i \right). \qquad [19]$$

Using Eqs. 3, 6, and 12 in defining:

$$\pi(t) := 1/S_{\text{eff}}(t) - \sigma\rho(t), \quad s_i(t) := -\sigma \sum_j z_{ij} p_j, \qquad [20]$$

we can write Eq. **19** as:

$$\dot{p}_i = Xp_i \{\pi + s_i - p_i\} + O(\lambda). \qquad [21]$$

The term s_i is responsible for the bias of species i against the reference proportion π. As a heuristic means of calculating the time-averaged bias, we suppose the p_i's can be treated independently of the z_{ij} and be replaced by $\overline{p_i} \approx 1/S$; then we obtain $\overline{s}_i \approx -\sigma z_i / \sqrt{S}$. On this basis, we expect z_i to be indicative of a species' dominance bias.

4.5. Derivation of the Stochastic Focal-Species Model from Dynamical Mean-Field Arguments. We write Eq. **1** as:

$$\dot{x}_i = x_i(g_i - x_i) + \lambda, \quad g_i = 1 - \sum_{j(\neq i)} \alpha_{ij} x_j. \qquad [22]$$

If we suppose that the abundances $\{x_j(t)\}$ (or, rather, their statistical properties) are independent of the particular realization $[\alpha_{ij}]$ of the interaction matrix, then, for a given realization of $\{x_j(t)\}$,

$$g_i(t) \sim \mathcal{N}\left(1 - \mu \sum_{j(\neq i)} x_j(t), \ \sigma^2 \sum_{j(\neq i)} x_j^2(t)\right), \qquad [23]$$

based on the properties of sums of Gaussian variables. The time-varying mean and variance of g_i means that, averaged over time, g_i does not necessarily follow a Gaussian distribution. We introduce

$$a(t) := 1 - \mu \sum_i x_i(t), \quad b(t) := \sigma \sqrt{\sum_i x_i^2(t)}, \qquad [24]$$

which are found to exhibit significant relative fluctuations, with skewed distributions (Fig. 7 *A* and *B*). However, once we shift and scale $g_i(t)$ into the "effective noise"

$$\eta_i(t) := \frac{g_i(t) - a(t)}{b(t)}, \qquad [25]$$

we recover (closely) a $\mathcal{N}(0, 1)$ distribution, for both the set $\{\eta_i(t)\}_{1,...,S}$ at any given time t, and for the stationary distribution of $\eta_i(t)$, at least for typical species (Fig. 7 *C* and *D*). The empirical distribution of the g_i across all species and times is closely approximated by the stationary distribution $\mathcal{N}(\overline{a}, \overline{b})$ (Fig. 7*E*). Therefore, we suppose that, despite their fluctuations, we can replace $a(t)$ and $b(t)$ with their time-averages and model g_i as a stochastic process $g(t) = \overline{a} + \overline{b}\eta(t)$, where $\eta(t)$ is a process with stationary distribution $\mathcal{N}(0, 1)$. The parameter correspondence in Eq. **9** follows by $k = -\overline{a}$, $u = \overline{b} \approx \overline{X}/\sqrt{S_{\text{eff}}}$, and $\tau = \tau_\eta$, the correlation time of η.

Note that, up to neglecting a diagonal term of the sum, the effective noise can be written:

$$\eta_i(t) = -\sum_{j(\neq i)} z_{ij} q_j(t), \qquad [26]$$

with $z_{ij} \sim \mathcal{N}(0, 1)$, and $\mathbf{q}(t) = \mathbf{x}(t)/\|\mathbf{x}(t)\|_2$. Given the chaotic turnover pattern, the latter is expected to perform something like a random walk on the S-sphere, with a de-correlation time corresponding to the turnover of dominant species. This timescale is inherited by the effective noise. More precisely, we compare autocorrelation functions (ACF). The ACF of a function f is defined as:

$$\phi_f(t_{\text{lag}}) := \text{mean}_t[\delta f(t) \cdot \delta f(t + t_{\text{lag}})]/\text{var}[f], \qquad [27]$$

with $\delta f = f - \overline{f}$. By definition $\phi_f(0) = 1$. For each species' effective noise we compute numerically $\phi_{\eta_i}(t_{\text{lag}})$, as shown in Fig. 7*F*. Due to the small number of "booms" per species, even over a large simulation time, ACFs are slightly irregular. In order to make estimations more accurate, we consider the averaged ACF $\phi_\eta := S^{-1} \sum_i \phi_{\eta_i}$ The decay of correlation is well-approximated by the exponential $\exp(-t_{\text{lag}}/\tau_\eta)$, where the parameter τ_η (fitted by least squares) represents the noise correlation timescale for a "typical" species.
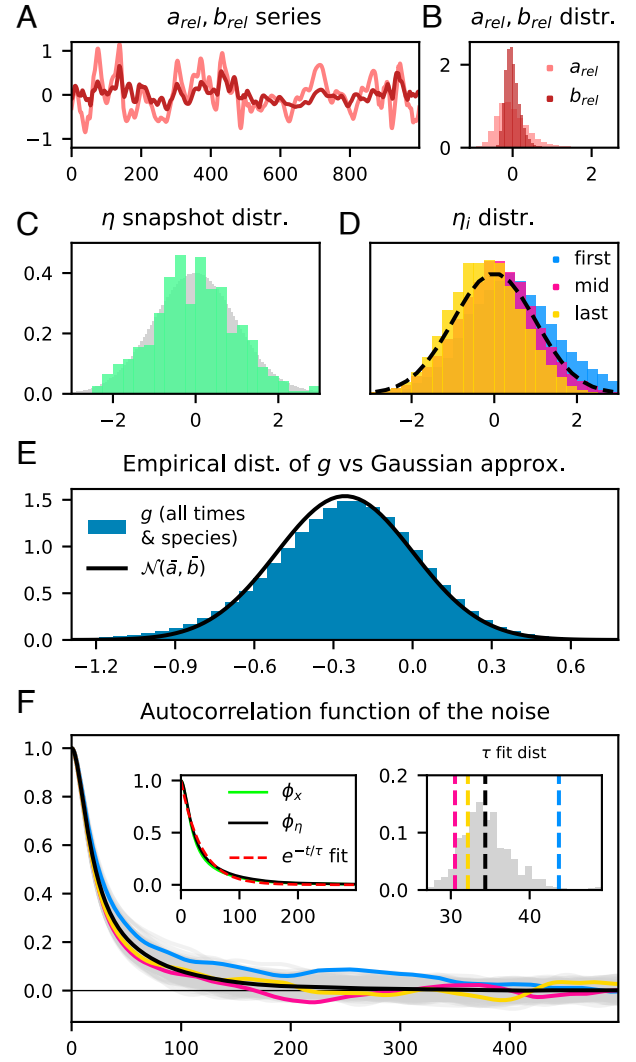

Fig. 7. Statistical properties of the effective noise. (*A* and *B*) Time series and distribution of $a_{\text{rel}} = a/\overline{a} - 1$, etc. (*C* and *D*) Histograms of $\eta_i(t)$ across all species and time (gray), over just species for one random time (green), over all time for the first/mid/last-ranked species with respect to average abundance (blue/pink/yellow), with $\mathcal{N}(0, 1)$ (black, dashed) for reference. (*E*) The empirical distribution of g in Eq. **22** over all species and times, compared to the distribution $\mathcal{N}(\overline{a}, \overline{b})$ assumed for g in the focal-species model. (*F*) Autocorrelation functions: for every species (gray), first/mid/last-rank species (blue/pink/yellow)), and the average over all species (black). The left inset compares the ACFs of $\mathbf{x}$ (green), η (black), and the exponential fit to the latter (red); the *Right Inset* shows the distribution of the τ parameter in exponential fits to each species ACF.

The approximately $\mathcal{N}(0, 1)$ distribution and exponential autocorrelation function of the effective noise η suggest that it can be modeled as an Ornstein–Uhlenbeck process, the only Markov process with these two properties;

$$\dot{\eta}(t) = -\frac{1}{\tau}\eta(t) + \sqrt{\frac{2}{\tau}}\xi(t), \qquad [28]$$

where $\xi(t)$ is a Gaussian white noise; $\langle\xi\rangle = 0$, $\langle\xi(t)\xi(t')\rangle = \delta(t - t')$. The timescale referred to as τ_{dom} in the main text can be defined as $\tau_{\mathbf{x}}$, the decay time of the exponential fit to the ACF of the abundance vector. For a vector-valued function, Eq. **27** gives

$$\phi_{\mathbf{x}} = \frac{1}{S}\sum_i w_i \phi_{x_i}, \quad w_i = \frac{\text{var}[x_i]}{\frac{1}{S}\sum_j \text{var}[x_j]}. \qquad [29]$$

ϕ_η and $\phi_{\mathbf{x}}$ match very well (*Inset* of Fig. 7*F*) for the reference simulation; as do the associated timescales $\tau_{\mathbf{x}}$ and τ_η for all (μ, σ) in the chaotic phase

(*SI Appendix*, Fig. S13). This observation motivates identifying τ_η of the focal-species model with the turnover timescale τ_{dom}. Thus, the focal species model and its parameters have been fully specified.

We point out a critical difference between our approach and dynamical mean-field theory applied in the weak-interaction regime (34, 41). Under strong interactions, $a(t)$ and $b(t)$ are determined by a small number (S_{eff}) of dominant species that, during the time of co-dominance, have strong effects on each other. Therefore, they can not be determined "self-consistently" from the focal-species equations by assuming that every species can described simultaneously as an independent realization of it. For example, the self-consistency relation for k in Eq. **7b** is $k = \mu S \langle x \rangle - 1$. In our reference simulation $k = 0.26$, whereas $\mu S \langle x \rangle - 1 = -0.31$ even has the wrong sign. This discrepancy is due to the neglected inter-species correlations needed for the collective correlation $\overline{\rho}$, Eq. **12**, to exceed the critical value Eq. **13** associated with $k > 0$ and boom-bust dynamics.

4.6. Steady-State Solution of the Focal-Species Model under the Unified Colored Noise Approximation.

The unified colored noise approximation (55) assumes overdamped dynamics to replace a process $\dot{x} = F(x) + G(x)\eta$, driven by Gaussian correlated noise η of correlation-time τ, with a process driven by white noise. The approximation is exact in the limits $\tau \to 0$ or $\tau \to \infty$. The stationary distribution of the corresponding white-noise process is

$$P^*(x) \propto \exp\left\{ \int^x v(x')\,dx' \right\}, \qquad [30]$$

with

$$v = \left(\tau^{-1/2} H_\tau\right)\frac{F}{G^2} + \left(\ln\frac{H_\tau}{G}\right)', \qquad [31]$$

and H_τ a function of F, G, and τ. For Eq. **7**

$$F(x) = -x(k + x) + \lambda, \quad G(x) = ux, \qquad [32]$$

$$H_\tau(x) = \tau^{-1/2} - \tau^{1/2}(x + \lambda x^{-1}). \qquad [33]$$

With these functions, the integral in Eq. **30** can be performed exactly, yielding

$$P^*(x) = \frac{1}{\mathcal{N}} e^{-[q_+(x)+q_-(\lambda/x)]} x^{-\nu}\left(\tau^{-1} + x + \frac{\lambda}{x}\right), \qquad [34]$$

where ν is given by Eq. **10**, and

$$q_\pm(y) := \frac{\left(y + (\tau^{-1} \pm k)\right)^2}{2u^2}. \qquad [35]$$

1. D. L. DeAngelis, J. C. Waterhouse, Equilibrium and nonequilibrium concepts in ecological models. *Ecol. Monogr.* **57**, 1–21 (1987).
2. M. Vellend, *The Theory of Ecological Communities* (Princeton University Press, 2016).
3. S. P. Hubbell, *The Unified Neutral Theory of Biogeography* (Princeton University Press, 2001).
4. H. G. Schuster, *Deterministic Chaos* (VCH, Weinheim, ed. 3, 1995).
5. A. A. Berryman, J. A. Millstein, Are ecological systems chaotic - and if not, why not? *Trends Ecol. Evol.* **4**, 26–28 (1989).
6. S. B. Munch, T. L. Rogers, B. J. Johnson, U. Bhat, C. H. Tsai, Rethinking the prevalence and pelevance of chaos in ecology. *Annu. Rev. Ecol. Evol. Syst.* **53**, 227–249 (2022).
7. T. L. Rogers, B. J. Johnson, S. B. Munch, Chaos is not rare in natural ecosystems. *Nat. Ecol. Evol.* **6**, 1105–1111 (2022).
8. T. L. Rogers, S. B. Munch, S. ichiro, S. Matsuzaki, C. C. Symons, Intermittent instability is widespread in plankton communities. *Ecol. Lett.* **26**, 470–481 (2023).
9. M. L. Sogin *et al.*, Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12115–12120 (2006).
10. M. D. J. Lynch, J. D. Neufeld, Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* **13**, 217–229 (2015).
11. F. Pascoal, R. Costa, C. Magalhães, The microbial rare biosphere: Current concepts, methods and ecological principles. *FEMS Microbiol. Ecol.* **97** (2021).
12. E. Ser-Giacomi *et al.*, Ubiquitous abundance distribution of non-dominant plankton across the global ocean. *Nat. Ecol. Evol.* **2**, 1243–1249 (2018).
13. J. A. Fuhrman, J. A. Cram, D. M. Needham, Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* **13**, 133–146 (2015).
14. A. M. Martin-Platero *et al.*, High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nat. Commun.* **9**, 266 (2018).
15. J. A. Gilbert *et al.*, Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).
16. E. Benincà *et al.*, Chaos in a long-term experiment with a plankton community. *Nature* **451**, 822–825 (2008).
17. I. V. Telesh *et al.*, Chaos theory discloses triggers and drivers of plankton dynamics in stable environment. *Sci. Rep.* **9**, 20351 (2019).
18. L. Becks, F. M. Hilker, H. Malchow, K. Jürgens, H. Arndt, Experimental demonstration of chaos in a microbial food web. *Nature* **435**, 1226–1229 (2005).
19. J. Hu, D. R. Amor, M. Barbier, G. Bunin, J. Gore, Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science* **378**, 85–89 (2022).
20. J. Huisman, F. J. Weissing, Biodiversity of plankton by species oscillations and chaos. *Nature* **402**, 407–410 (1999).
21. P. Schippers, A. M. Verschoor, M. Vos, W. M. Mooij, Does "supersaturated coexistence" resolve the "paradox of the plankton"? *Ecol. Lett.* **4**, 404–407 (2001).
22. I. Ispolatov, V. Madhok, S. Allende, M. Doebeli, Chaos in high-dimensional dissipative dynamical systems. *Sci. Rep.* **5**, 12506 (2015).
23. D. A. Kessler, N. M. Shnerb, Generalized model of island biodiversity. *Phys. Rev. E* **91**, 042705 (2015).
24. F. Roy, M. Barbier, G. Biroli, G. Bunin, Complex interactions can create persistent fluctuations in high-diversity ecosystems. *PLoS Comput. Biol.* **16**, e1007827 (2020).
25. J. D. O'Sullivan, J. C. D. Terry, A. G. Rossberg, Intrinsic ecological dynamics drive biodiversity turnover in model metacommunities. *Nat. Commun.* **12**, 3627 (2021).
26. M. T. Pearce, A. Agarwala, D. S. Fisher, Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14572–14583 (2020).
27. P. Rodríguez-Sánchez, E. H. van Nes, M. Scheffer, Neutral competition boosts cycles and chaos in simulated food webs. *R. Soc. Open Sci.* **7**, 191532 (2020).
28. I. Dalmedigos, G. Bunin, Dynamical persistence in high-diversity resource-consumer communities. *PLoS Comput. Biol.* **16**, e1008189 (2020).
29. M. Scheffer, S. Rinaldi, J. Huisman, F. J. Weissing, Why plankton communities have no equilibrium: Solutions to the paradox. *Hydrobiologia* **491**, 9–18 (2003).
30. J. Denk, S. Martis, O. Hallatschek, Chaos may lurk under a cloak of neutrality. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16104–16106 (2020).
31. J. T. Lennon, S. E. Jones, Microbial seed banks: The ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* **9**, 119–130 (2011).
32. R. May, Will a large complex system be stable? *Nature* **238**, 413–414 (1972).
33. S. Allesina, S. Tang, The stability-complexity relationship at age 40: A random matrix perspective. *Popul. Ecol.* **57**, 63–75 (2015).
34. G. Bunin, Ecological communities with Lotka-Volterra dynamics. *Phys. Rev. E* **95**, 042414 (2017).
35. M. Barbier, J. F. Arnoldi, G. Bunin, M. Loreau, Generic assembly patterns in complex ecological communities. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2156–2161 (2018).
36. A. F. Ansari, Y. B. S. Reddy, J. Raut, N. M. Dixit, An efficient and scalable top-down method for predicting structures of microbial communities. *Nat. Comput. Sci.* **1**, 619–628 (2021).
37. A. Skwara *et al.*, Statistically learning the functional landscape of microbial communities. *Nat. Ecol. Evol.* (2023).
38. J. Hofbauer, K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, 2002).
39. M. Barbier, J. F. Arnoldi, The cavity method for community ecology. bioRxiv (2017). 10.1101/147728 (Accessed 28 July 2023).
40. T. Galla, Dynamically evolved community size and stability of random Lotka-Volterra ecosystems. *Europhys. Lett.* **123**, 48004 (2018).
41. F. Roy, G. Biroli, G. Bunin, C. Cammarota, Numerical implementation of dynamical mean field theory for disordered systems: Application to the Lotka-Volterra model of ecosystems. *J. Phys. A: Math. Theor.* **52**, 484001 (2019).
42. J. W. Baron, T. J. Jewell, C. Ryder, T. Galla, Breakdown of random-matrix universality in persistent Lotka-Volterra communities. *Phys. Rev. Lett.* **130**, 137401 (2023).
43. G. Biroli, G. Bunin, C. Cammarota, Marginally stable equilibria in critical ecosystems. *New J. Phys.* **20**, 083051 (2018).
44. A. Altieri, F. Roy, C. Cammarota, G. Biroli, Properties of equilibria and glassy phases of the random Lotka-Volterra model with demographic noise. *Phys. Rev. Lett.* **126**, 258301 (2021).
45. O. S. Venturelli *et al.*, Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.* **14**, e8157 (2018).
46. D. Machado *et al.*, Polarization of microbial communities between competitive and cooperative metabolism. *Nat. Ecol. Evol.* **5**, 195–203 (2021).
47. S. Xu, L. Böttcher, T. Chou, Diversity in biology: Definitions, quantification and models. *Phys. Biol.* **17**, 031001 (2020).
48. R. Logares *et al.*, Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.* **24**, 813–821 (2014).
49. K. Kaneko, I. Tsuda, Chaotic itinerancy. *Chaos* **13**, 926–936 (2003).
50. G. B. Arous, Y. V. Fyodorov, B. A. Khoruzhenko, Counting equilibria of large complex systems by instability index. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2023719118 (2021).
51. V. Ros, F. Roy, G. Biroli, G. Bunin, A. M. Turner, Generalized Lotka-Volterra equations with random, nonreciprocal interactions: The typical number of equilibria. *Phys. Rev. Lett.* **130**, 257401 (2023).

52. T. J. Matthews, R. J. Whittaker, Fitting and comparing competing models of the species abundance distribution: Assessment and prospect. *Front. Biogeogr.* **6**, 67 (2014).

53. J. Grilli, Macroecological laws describe variation and diversity in microbial communities. *Nat. Commun.* **11**, 4743 (2020).

54. L. Descheemaeker, S. de Buyl, Stochastic logistic models reproduce experimental time series of microbial communities. *eLife* **9**, e55650 (2020).

55. P. Jung, P. Hänggi, Dynamical systems: A unified colored-noise approximation. *Phys. Rev. A* **35**, 4464–4466 (1987).

56. D. A. Kessler, N. M. Shnerb, Neutral selection. arXiv [Preprint] (2012). http://arxiv.org/abs/1211.3609 (Accessed 28 July 2023).

57. A. E. Magurran, P. A. Henderson, Explaining the excess of rare species in natural species abundance distributions. *Nature* **422**, 714–716 (2003).

58. W. Ulrich, M. Ollik, Frequent and occasional species and the shape of relative abundance distributions. *Div. Distrib.* **10**, 263–269 (2004).

59. G. Hardin, The competitive exclusion principle. *Science* **131**, 1292–1297 (1960).

60. T. A. de Pirey, G. Bunin, Many-species ecological fluctuations as a jump process from the brink of extinction. arXiv [Preprint] (2023). http://arxiv.org/abs/2306.13634 (Accessed 28 July 2023).

61. Z. Eisler, I. Bartos, J. Kertész, Fluctuation scaling in complex systems: Taylor's law and beyond1. *Adv. Phys.* **57**, 89–142 (2008).

62. X. Jia, F. Dini-Andreote, J. F. Salles, Community assembly processes of the microbial rare biosphere. *Trends Microbiol.* **26**, 738–747 (2018).

63. G. B. Stav Marcus, Ari. M. Turner, Local and extensive fluctuations in sparsely-interacting ecological communities. arXiv [Preprint] (2023). http://arxiv.org/abs/2308.01828v1 (Accessed 28 July 2023).

64. K. Hashimoto, T. Ikegami, Heteroclinic chaos, chaotic itinerancy and neutral attractors in symmetrical replicator equations with mutations. *J. Phys. Soc. Japan* **70**, 349–352 (2001).

65. J. Hofbauer, Heteroclinic cycles in ecological differential equations. *Tatra Mountains Math. Publ.* **4**, 105–116 (1994).

66. V. Afraimovich, I. Tristan, R. Huerta, M. I. Rabinovich, Winnerless competition principle and prediction of the transient dynamics in a Lotka-Volterra model. *Chaos* **18**, 043103 (2008).

67. C. Bick, M. I. Rabinovich, On the occurrence of stable heteroclinic channels in Lotka-Volterra models. *Dyn. Syst.* **25**, 97–110 (2009).

68. T. A. de Pirey, G. Bunin, Aging by near-extinctions in many-variable interacting populations. *Phys. Rev. Lett.* **130**, 098401 (2023).

69. M. F. M. Bjorbækmo, A. Evenstad, L. L. Røsæg, A. K. Krabberød, R. Logares, The planktonic protist interactome: Where do we stand after a century of research? *ISME J.* **14**, 544–559 (2019).

70. N. C. Millette *et al.*, Mixoplankton and mixotrophy: Future research priorities. *J. Plankton Res.* **45**, 576–596 (2023).

71. P. Villa Martín, A. Koldaeva, S. Pigolotti, Coalescent dynamics of planktonic communities. *Phys. Rev. E* **106**, 044408 (2022).

72. D. Kessler, S. Suweis, M. Formentin, N. M. Shnerb, Neutral dynamics with environmental noise: Age-size statistics and species lifetimes. *Phys. Rev. E* **92**, 022722 (2015).

73. M. J. Behrenfeld, R. O'Malley, E. Boss, L. Karp-Boss, C. Mundt, Phytoplankton biodiversity and the inverted paradox. *ISME Commun.* **1** (2021).

74. M. Scheffer, Should we expect strange attractors behind plankton dynamics – and if so, should we bother? *J. Plankton Res.* **13**, 1291–1305 (1991).

75. A. Bracco, A. Provenzale, I. Scheuring, Mesoscale vortices and the paradox of the plankton. *Proc. R. Soc. London. Ser. B: Biol. Sci.* **267**, 1795–1800 (2000).

76. F. d'Ovidio, S. De Monte, S. Alvain, Y. Dandonneau, M. Lévy, Fluid dynamical niches of phytoplankton types. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18366–18370 (2010).

77. A. B. Medvinsky *et al.*, Chaos far from the edge of chaos: A recurrence quantification analysis of plankton time series. *Ecol. Compl.* **23**, 61 (2015).

78. C. M. Mutshinda, Z. V. Finkel, C. E. Widdicombe, A. J. Irwin, Ecological equivalence of species within phytoplankton functional groups. *Funct. Ecol.* **30**, 1714–1722 (2016).

79. K. T. Allhoff, D. Ritterskamp, B. C. Rall, B. Drossel, C. Guill, Evolutionary food web model based on body masses gives realistic networks with permanent species turnover. *Sci. Rep.* **5** (2015).

80. M. Hamm, B. Drossel, The concerted emergence of well-known spatial and temporal ecological patterns in an evolutionary food web model in space. *Sci. Rep.* **11** (2021).

81. M. Doebeli, I. Ispolatov, Chaos and unpredictability in evolution. *Evolution* **68**, 1365–1373 (2014).

82. M. Doebeli, E. C. Jaque, Y. Ispolatov, Boom-bust population dynamics increase diversity in evolving competitive communities. *Commun. Biol.* **4**, 502 (2021).

83. J. R. Bray, J. T. Curtis, An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27**, 325–349 (1957).

84. E. Mallmin, Simulation code & data for "Chaotic turnover of rare and abundant species in a strongly interacting model community". Zenodo. https://doi.org/10.5281/zenodo.10646601. Deposited 21 February 2024.