Journal of Theoretical Biology 269 (2011) 150–165

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

A mathematical synthesis of niche and neutral theories in community ecology

Bart Haegeman [a],*, Michel Loreau [b]

[a] INRIA Research Team MERE, UMR MISTEA, 2 Place Pierre Viala, 34060 Montpellier, France
[b] McGill University, Department of Biology, 1205 Avenue Docteur Penfield, Montreal, Québec, Canada H3A 1B1

ARTICLE INFO

Article history:
Received 7 July 2010
Received in revised form
22 September 2010
Accepted 6 October 2010
Available online 12 October 2010

Keywords:
Demographic stochasticity
Immigration
Lotka–Volterra model
Neutral community model
Species abundance distribution

ABSTRACT

The debate between niche-based and neutral community theories centers around the question of which forces shape predominantly ecological communities. Niche theory attributes a central role to niche differences between species, which generate a difference between the strength of intra- and interspecific interactions. Neutral theory attributes a central role to migration processes and demographic stochasticity. One possibility to bridge these two theories is to combine them in a common mathematical framework. Here we propose a mathematical model that integrates the two perspectives. From a niche-based perspective, our model can be interpreted as a Lotka–Volterra model with symmetric interactions in which we introduce immigration and demographic stochasticity. From a neutral perspective, it can be interpreted as Hubbell's local community model in which we introduce a difference between intra- and interspecific interactions. We investigate the stationary species abundance distribution and other community properties as functions of the interaction coefficient, the immigration rate and the strength of demographic stochasticity.

1. Introduction

Community ecology aims at describing the forces that structure ecological communities. Classical theories explain community dynamics in terms of species niches (Chase and Leibold, 2004; Loreau, 2010). Niche theory states that the long-term coexistence of species is possible only if their niches are sufficiently separated. Niche differences can be due to a range of mechanisms, such as different ways to use resources, different interactions with competitors or mutualists, and different spatial and temporal characteristics. The Lotka–Volterra competition model provides a convenient mathematical framework to deal with species niches (MacArthur and Levins, 1967; MacArthur, 1972). In this model, niche differences are effectively taken into account as differences between the strength of intra- and interspecific competition. The more species niches overlap, the larger the ratio of inter- and intraspecific competition strength (equal to the parameter α in our model). Hence, the Lotka–Volterra model can be considered as a minimal model of niche theory.

Neutral theory takes a quite different approach to community ecology (Caswell, 1976; Hubbell, 2001). It starts from the assumption that species are identical in all characteristics that may affect their population dynamics. Community structure is the result of stochastic birth–death processes. Species coexistence is guaranteed in a trivial way, by considering a constant species flow into the community (interpreted as immigration or, on a larger scale, as speciation). Hubbell's local community model, in which a demographically fluctuating community of fixed size receives immigrants from a large species pool, has become the reference neutral model (Hubbell, 2001; Alonso et al., 2006). This minimal combination of demographic stochasticity and immigration can generate a range of community patterns, matching some empirical data surprisingly well.

It is now generally accepted that niche-based and neutral community models should not be seen as radically opposed model paradigms (Gravel et al., 2006; Holyoak and Loreau, 2006; Leibold and McPeek, 2006; Adler et al., 2007). Rather, each of these model classes emphasizes a distinct set of ecological mechanisms. Neither the niche-based nor the neutral framework exclude the integration of additional processes in principle. Hence, it should be possible to construct more general community models that take into account the mechanisms involved in both niche theory and neutral theory. Such integrative models would allow us to bridge the conceptual gap between the two theories. In particular, they could reveal which mechanisms exactly underlie the simplifying approach of neutral community models, and its empirical successes.

In this paper we build an integrative community model that incorporates demographic stochasticity, immigration flow, and (competitive or mutualistic) species interactions. To avoid intractable constructions, our model combines a minimal niche model and a minimal neutral model. As a minimal niche model, we consider the Lotka–Volterra equations with symmetric species interactions. This means that (a) intraspecific interaction strength is the same for

---

* Corresponding author.
E-mail addresses: bart.haegeman@inria.fr (B. Haegeman), michel.loreau@mcgill.ca (M. Loreau).

0022-5193/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2010.10.006

all species, and that (b) interspecific interaction strength is the same for all species pairs (but intra- and interspecific interaction strengths can differ). As a minimal neutral model, we consider Hubbell's local community model. The resulting model has a limited number of parameters, can be handled analytically, and allows a systematic study of its stationary behaviour.

Recently, other proposals have been made to introduce niche features into the neutral community framework. Some papers have considered intraspecific interactions, or equivalently, species-level density dependence, meaning that growth rates depend on the density of conspecifics only (Volkov et al., 2005). Other papers have considered interspecific interactions, or equivalently, community-level density dependence, meaning that growth rates depend on the total density of all individuals in the community (Kadmon and Allouche, 2007; Haegeman and Etienne, 2008; Allouche and Kadmon, 2009). The community model of this paper includes both intra- and interspecific interactions, and therefore unifies previous, separate treatments of species-level and community-level density dependence. Related models have been studied by Loreau and de Mazancourt (2008), who used a linear approximation to study the synchronization of population fluctuations, and by Volkov et al. (2009), who used moment equations to infer species interactions from abundance data.

Niche processes can also be introduced more explicitly into neutral-like models. For example, species-specific habitat preferences can be defined in a spatially heterogeneous environment, and regulate the stochastic birth–death dynamics of the different species. This is close in spirit to metacommunity models (Mouquet and Loreau, 2003), and has been considered in a neutral setting by several authors (Tilman, 2004; Schwilk and Ackerly, 2005; Gravel et al., 2006; Zillio and Condit, 2007). Here we do not consider spatial heterogeneity, but restrict our attention to a local community. We consider immigration from a large species pool, without modeling the species pool dynamics. We compare the effects of the immigration process and the internal community dynamics (intra- and interspecific interactions, and demographic stochasticity) on community patterns.

2. Population model

Before tackling the multi-species community model, we study a dynamical model for a single species. The population model will be the basic building block for the community model of the next section. It can be obtained from the logistic growth model by adding demographic stochasticity and immigration. The resulting stochastic logistic model (with or without immigration) has been studied extensively (Pielou, 1977; Renshaw, 1991; Matis and Kiffe, 2000). Here we recall some basic model properties, and introduce a number of analytical tools; both properties and tools will be useful for the study of our community model.

As a baseline model, we consider the deterministic population model with logistic growth and immigration,

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN\left(1-\frac{N}{K}\right)+\mu, \tag{1}$$

with population size N, intrinsic growth rate r, carrying capacity K and immigration rate μ. The model (1) has a single equilibrium N^*,

$$N^* = \frac{K}{2}\left(1+\sqrt{1+4\frac{\mu}{rK}}\right), \tag{2}$$

which is globally stable. For weak immigration ($\mu \ll rK$), the equilibrium population size N^* is close to the carrying capacity K. For stronger immigration, the equilibrium population size N^* increases with the immigration rate μ. In the latter case, the population is externally forced to a larger size than its internal dynamics can sustain.

2.1. Construction of the stochastic model

To include demographic stochasticity into (1), we take into account the discrete nature of the population size N, i.e., the population size N can only take integer values $0,1,2,\ldots$, in contrast to the continuous variable N of model (1). The stochastic model dynamics consist of a series of events affecting the population size N: the population can increase by one individual due to a birth or immigration event, and it can decrease by one individual due to a death event. We have to specify the rate at which these events occur: denote by $q_+(N)$ the rate of population increase, and by $q_-(N)$ the rate of population decrease. This means that during a small time interval δt, the probability

that the population size increases by one	equals	$q_+(N)\delta t$,
that the population size decreases by one	equals	$q_-(N)\delta t$,
that the population size stays the same	equals	$1-(q_+(N)+q_-(N))\delta t$.

Note that by going from a continuous to a discrete variable N, we have simultaneously gone from a deterministic to a stochastic model (see Appendix A for the notation we use to describe stochastic models).

To construct a stochastic version of population model (1), we have to specify the transition rates $q_+(N)$ and $q_-(N)$. In Appendix B we show that the deterministic part of a stochastic population model with transition rates $q_+(N)$ and $q_-(N)$ is given by the difference $q_+(N)-q_-(N)$. Formally,

$$\frac{\mathbb{E}[\delta N]}{\delta t} = q_+(N)-q_-(N), \tag{3}$$

where $\mathbb{E}[\delta N]$ is the expected change in population size in a small time interval δt (see (B.1) for a more rigorous formulation). Hence, by requiring that

$$q_+(N)-q_-(N) = rN\left(1-\frac{N}{K}\right)+\mu \tag{4}$$

we guarantee that the expected, i.e., deterministic, behaviour of the stochastic population model is identical to that of the corresponding deterministic model (1).

Condition (4) does not fix uniquely the transition rates $q_+(N)$ and $q_-(N)$. After imposing the difference $q_+(N)-q_-(N)$, we still can choose the sum $q_+(N)+q_-(N)$ independently. In Appendix B we show that the sum of the transition rates $q_+(N)$ and $q_-(N)$ measures the intensity of the stochastic fluctuations superimposed on the deterministic model (3). Formally,

$$\frac{\mathrm{Var}[\delta N]}{\delta t} = q_+(N)+q_-(N), \tag{5}$$

where $\mathrm{Var}[\delta N]$ is the variance of the change in population size, see (B.2). As a consequence, there are different ways to incorporate demographic stochasiticity into model (1), leading to different stochastic population models.

To illustrate this, we consider two possible choices for the transition rates $q_+(N)$ and $q_-(N)$. As a first choice, take

$$q_+(N) = r_+N+\mu,$$
$$q_-(N) = r_-N+\frac{r}{K}N^2, \tag{6}$$

with $r = r_+ - r_-$. Transition rates (6) attributes density dependence entirely to death events: the per capita death rate $r_- + (r/K)N$ increases as population size increases, whereas the per capita birth rate r_+ is constant. We call this the density-dependent mortality version of the model. Rates (6) satisfy condition (4), so that the corresponding stochastic model has deterministic part given by (1).

A second choice attributes density dependence entirely to birth events,

$$q'_{+}(N) = r'_{+}N - \frac{r}{K}N^2 + \mu,$$

$$q'_{-}(N) = r'_{-}N, \tag{7}$$

with $r = r'_{+} - r'_{-}$. The per capita birth rate $r'_{+} - (r/K)N$ decreases as population size increases, whereas the per capita death rate r'_{-} is constant. We call this the density-dependent natality version. Again, transition rates (7) satisfy condition (4), and lead to a deterministic part given by (1). Note that to avoid negative transition rates, there should be a maximal population size for the density-dependent natality case (7).

The choice between (6) or (7) depends on the nature of the density dependence, whether it affects birth or death rate. One could also consider intermediary cases, in which density dependence is present in both rates $q_{+}(N)$ and $q_{-}(N)$, which would also be compatible with condition (4). Because the density-dependent mortality (6) and natality (7) versions can be considered as two limiting cases, we focus on these two versions to study the sensitivity of our model to $q_{+}(N)$ and $q_{-}(N)$.

2.2. *Stationary distribution*

We have constructed two stochastic versions of our population model given by (6) and (7), both adding demographic stochasticity to the deterministic model (1). We compute and compare the stationary distribution for the population size N for both versions.

The stationary distribution of population abundance can be computed explicitly, for any combination of transition rates $q_{+}(N)$ and $q_{-}(N)$, see Appendix C. Fig. 1A shows the stationary distribution for different parameter combinations, and for the density-dependent mortality (6) and natality (7) versions. Recall that, by construction, the two forms of density dependence have the same deterministic behaviour (3), but the intensity of their stochastic fluctuations (5) is different. Parameters were chosen such that at carrying capacity $K \approx N^*$ the variance of population fluctuations is the same (this condition is satisfied by taking $r_{+} = r'_{-}$). Note that the density-dependent mortality version (6) has larger fluctuations for population size $N > K$, while the density-dependent natality version (7) has larger fluctuations for population size $N < K$.

We compute the stationary distribution for different values of r_{-}, which can be considered as a proxy for the intensity of demographic stochasticity. The stationary distributions for the density-dependent mortality version (6) (red curve) and for the density-dependent natality version (7) (green curve) almost coincide for all values of r_{-}. For small values of r_{-} (left panel), the distribution is concentrated at the equilibrium population size N^* of the deterministic model. When increasing the value of r_{-}, the intensity of demographic stochasticity increases, and the stationary distribution gets wider. The main mode of the distribution is still located at $N = N^*$, but a second, smaller mode appears at $N = 0$ (middle panel). For even more intense demographic stochasticity, the mode at $N = N^*$ decreases and ultimately disappears, while the probability of a small population further increases (right panel).

Fig. 1A also shows the stationary distribution of a linear approximation that is often used in population ecology (Renshaw, 1991; Matis and Kiffe, 2000; Lande et al., 2003). The approximation consists in linearizing the non-linear population model (1), and replacing the discrete randomness of demographic stochasticity by continuous Gaussian random variables, see

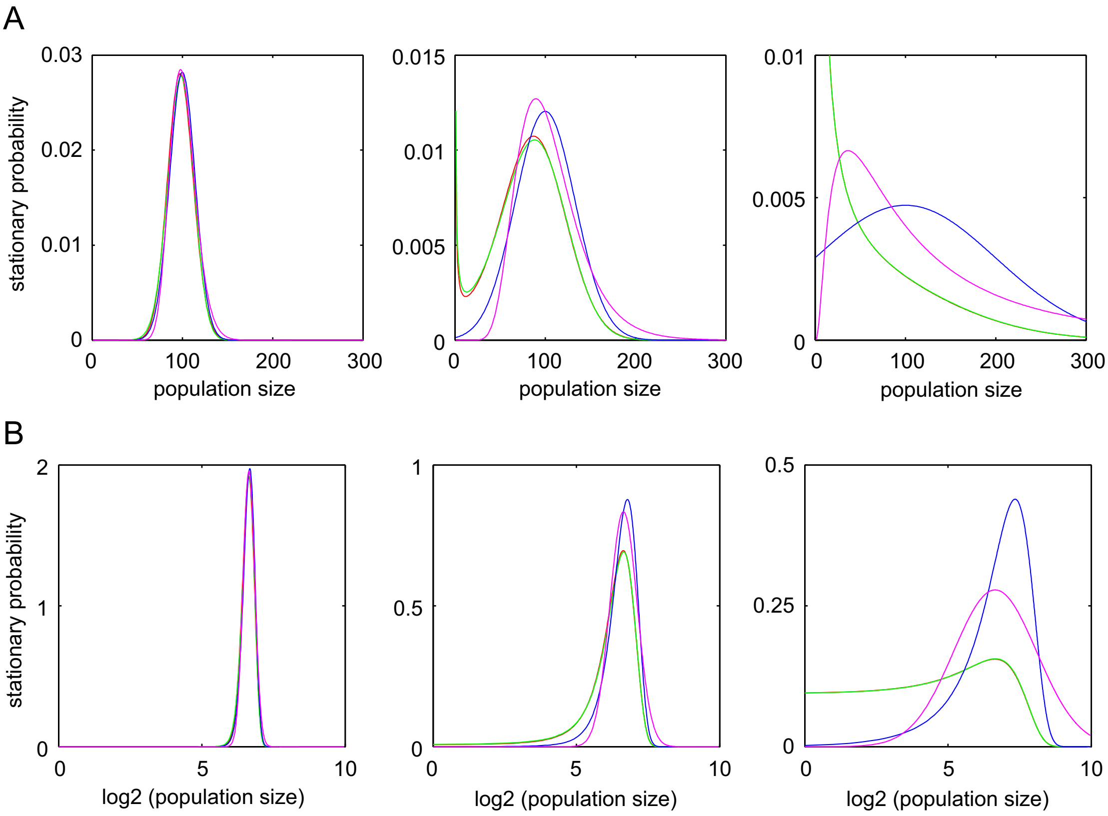

Fig. 1. Comparison of stationary population size distribution for different stochastic models and their approximations: (A) linear scale, (B) logarithmic scale. In red: density-dependent mortality (6); in green: density-dependent natality (7); in blue: linear approximation (B.4); in magenta: linear approximation (B.5). Left: $r_{-} = 1$; middle: $r_{-} = 10$; right: $r_{-} = 100$. Other parameters are $K = 100$, $\mu = 0.1$, $r = 1.0$, $r_{+} = r + r_{-}$, $r'_{-} = r_{+}$, $r'_{+} = r + r'_{-}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Appendix B. A first variant of the linear approximation (blue curve) is based on a linearization in terms of the population size N, see Eq. (B.4) in Appendix B; a second variant (magenta curve) is based on a linearization in terms of the logarithmic population size ln N, see Eq. (B.5). The two versions of our model have the same linear approximations because they correspond to the same deterministic population model (1), and they have the same fluctuation variance at carrying capacity. Both approximations are excellent for small demographic stochasticity, i.e., for small r_-, but deteriorate rapidly for larger values of r_-. In particular, the approximations describe poorly populations for which the probability of extinction is not negligible.

Population size can vary over orders of magnitude, and is therefore more conveniently represented on a logarithmic scale. Fig. 1B shows the same stationary distributions as Fig. 1A, but now as probability densities for the logarithmic population size $\log_2 N$. We use a base-two logarithm, as is customary in Preston plots. The transformation is given by

$$\mathbb{P}[\log_2 N] \approx \ln(2) N \mathbb{P}[N], \tag{8}$$

where we assumed that the population size can be considered as a continuous variable. Note that the population size $N=0$ is not representable on a logarithmic scale, so that the population size distribution is conditioned on $N>0$. For small demographic stochasticity, the distribution is log-normal, i.e., normal on a logarithmic scale. For larger demographic stochasticity, the logarithmic population size distribution is left-skewed. Note that, although the mode at $N=N^*$ disappears in the linear size distribution, it remains in the logarithmic size distribution.

2.3. Population properties

We have compared the stationary population size distributions of the density-dependent mortality and the density-dependent natality versions of our model. We have shown that the two versions lead to very similar distributions over a wide range of parameter values. We now perform a more systematic study for the density-dependent mortality case (6).

Rather than computing the entire stationary distribution for all parameter combinations, we consider here a limited number of population properties:

- the mean population size $\mathbb{E}[N]$;
- the variability of population size, measured by the coefficient of variation $\mathrm{CV}[N]$;
- the probability that the population is extinct, $\mathbb{P}[N=0]$.

To facilitate comparison with the community model in the next section, however, we will consider the mean number of species $\mathbb{E}[S] = 1-\mathbb{P}[N=0]$ instead of the extinction probability. Note that for the population model the statements $S=1$ and $N>0$ are equivalent. Note also that extinction is not fatal in our model, because immigration can initiate the population again.

Fig. 2 plots the three population properties as functions of the intensity of demographic stochasticity r_- and the immigration rate μ. For small demographic stochasticity, the population size has a sharp distribution ($\mathrm{CV}[N] \approx 0.1$) centered at the equilibrium N^*. In this parameter region, the stochastic model is close to its deterministic counterpart. The correspondence between the deterministic and stochastic models gets even better for larger immigration rates. For larger demographic stochasticity, mean population size decreases below the equilibrium N^*, and population size variability increases rapidly. For small immigration rate, there is a sharp transition to population extinction ($\mathbb{E}[S] \approx 0$).

It is interesting to compare the exact results of Fig. 2 with the commonly used linear approximations, see Appendix B. Fig. S1, in Appendix G shows the same population properties as in Fig. 2 but computed using approximation (B.4). The results for approximation (B.5) are very similar (not shown). In the linear approximation, mean population size and mean number of species are independent of demographic stochasticity, and identical to the exact values for small demographic stochasticity. The approximate variability $\mathrm{CV}[N]$ coincides with the exact values when demographic stochasticity is weak, but is too small for strong demographic stochasticity. Again, we find that the linear approximation is accurate for small values of r_-, but deteriorates rapidly for larger values of r_-.

3. Community model

In the previous section we added demographic stochasticity to a minimal population model. Here we generalize this approach to a multi-species community model. We start from the Lotka–Volterra model, which can be considered as a minimal community model with species interactions. Analogously with the population model, we propose an individual-based, stochastic community model, and compute its stationary distribution.

We build up from a deterministic community model, including (competitive or mutualistic) species interactions and immigration. The internal community dynamics are governed by the Lotka–Volterra equations:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = rN_i\left(1-\frac{N_i+\alpha\sum_{j\neq i}N_j}{K'}\right) = rN_i\left(1-\frac{(1-\alpha)N_i+\alpha\sum_j N_j}{K'}\right), \quad i=1,2,\ldots,S_{\mathrm{T}}, \tag{9}$$

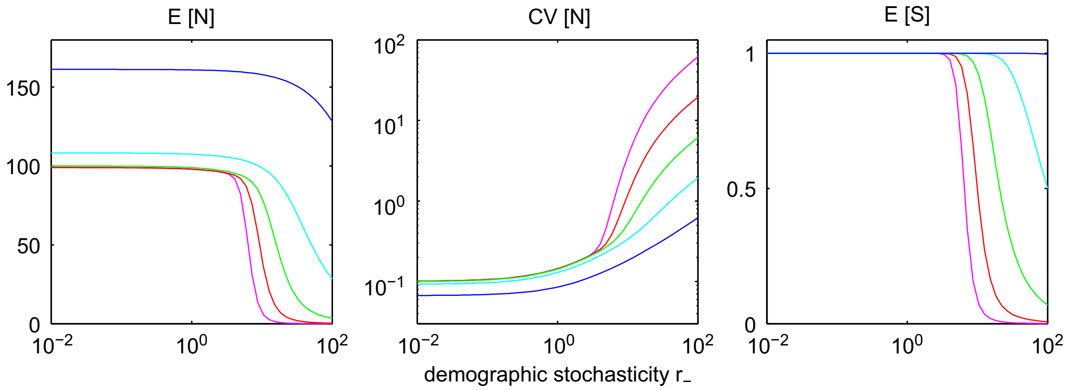

Fig. 2. Population properties as functions of model parameters. Left: mean population size $\mathbb{E}[N]$. Middle: coefficient of variation of population size $\mathrm{CV}[N]$. Right: expected number of species $\mathbb{E}[S]$. The demographic stochasticity coefficient r_- is plotted on the x-axis; the curves are parametrized by the immigration rate μ: 0.01 (magenta), 0.1 (red), 1 (green), 10 (cyan), 100 (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with N_i the abundance of species i, r the intrinsic per capita growth rate, K' the species-level carrying capacity, α the interaction coefficient, and S_T the total number of species.

Species interactions are competitive for $\alpha > 0$, and mutualistic for $\alpha < 0$. When species interactions are competitive, if the abundance of a species is increased by δN, then its own per capita growth rate is decreased by $(r/K')\delta N$, while the per capita growth rate of another species is decreased by $\alpha(r/K')\delta N$. Using the change in growth rate as a measure of interaction strength, the interaction coefficient α can be interpreted as the ratio between inter- and intraspecific interaction strength. When species interactions are mutualistic, a similar interpretation holds, but in this case an increase in the abundance of one species causes an increase in the growth rate of the other species. Note that the factor $1-\alpha$ is proportional to the difference between intra- and interspecific interaction strengths.

The parameter K' appearing in model (9) is called the species-level carrying capacity, because it is equal to the equilibrium population size in the absence of other species. It should be contrasted with the community-level carrying capacity K, defined as the equilibrium community size,

$$K = \sum_i N_i^* = \frac{S_T K'}{1+\alpha(S_T-1)}, \tag{10}$$

with N_i^* the equilibrium abundance of species i,

$$N_i^* = N^* = \frac{K'}{1+\alpha(S_T-1)}.$$

When $0 < \alpha < 1$, the species-level carrying capacity K' is smaller than the community-level carrying capacity K. Because the interaction between heterospecifics is weaker than between conspecifics, the carrying capacity K' as perceived by an isolated species is smaller than the carrying capacity K of the entire community. When $\alpha = 1$, all individuals interact with the same strength irrespective of their species identity, and both carrying capacities K and K' are equal. When $\alpha = 0$, the niches of the various species do not overlap, their dynamics are independent, and the community-level carrying capacity is the sum of the species-level carrying capacities, $K = S_T K'$. For competitive interactions, we have $\alpha > 0$ and $K < S_T K'$ due to niche overlap; for mutualistic interactions, we have $\alpha < 0$ and $K > S_T K'$. Note that for fixed K, the species-level carrying capacity K' goes to zero for $\alpha \to -1/(S_T-1)$; for fixed K', the community-level carrying capacity K diverges for $\alpha \to -1/(S_T-1)$.

The parameters r and K' are the same for all species, and the interaction coefficient α is the same for all species pairs. This implies that the community model (9) has a symmetry: permutating the species does not change the model equations.

Next, we add immigration to the internal dynamics (9). We assume that individuals can immigrate into the community from a much larger species pool. All S_T species are present in the species pool, and have the same abundance. Although different from Hubbell's model, this assumption is natural for neutral community models (Bell, 2000). As a result, the immigration rate μ from the species pool is the same for all species. This leads to the following equations:

$$\frac{dN_i}{dt} = rN_i\left(1-\frac{(1-\alpha)N_i+\alpha\sum_j N_j}{K'}\right)+\mu \quad i=1,2,\ldots,S_T. \tag{11}$$

Again, all species have the same parameters, so that this model has species permutation symmetry. Model (11) has a single equilibrium

$$N_i^* = N^* = \frac{K}{2S_T}\left(1+\sqrt{1+4\frac{S_T\mu}{Kr}}\right), \tag{12}$$

which is globally stable for $\mu > 0$ and $\alpha \le 1$. For weak immigration ($\mu S_T \ll rK$), the equilibrium species abundance N^* is close to the species carrying capacity K/S_T. For stronger immigration, the species populations and the community as a whole are externally pushed above their carrying capacities.

3.1. Construction of the stochastic model

To introduce demographic stochasticity into the deterministic model (11), we first replace the continuous species abundances N_i by discrete variables that can only take values 0,1,2,…. The community composition is described by the abundance vector $\vec{N} = (N_1,N_2,\ldots,N_{S_T})$, a vector of S_T integers.

The dynamics occur in the form of a series of stochastic events. During each event, one of the species increases its abundance by one individual due to a birth or immigration event, or decreases its abundance by one individual due to a death event. We have to specify the transition rates. Given community composition $\vec{N}$, we denote the rate of increase for species i by $q_{i+}(\vec{N})$, and the rate of decrease for species i by $q_{i-}(\vec{N})$.

We look for a stochastic model the expected, i.e., deterministic, behaviour of which is given by (11). In Appendix D we show that the deterministic part of the stochastic community model is

$$\frac{\mathbb{E}[\delta N_i]}{\delta t} = q_{i+}(\vec{N}) - q_{i-}(\vec{N}),$$

where $\mathbb{E}[\delta N_i]$ is the expected change of species abundance N_i in a small time interval δt, see (D.2). Hence, the transition rates $q_{i+}(\vec{N})$ and $q_{i-}(\vec{N})$ have to satisfy

$$q_{i+}(\vec{N}) - q_{i-}(\vec{N}) = rN_i\left(1-\frac{(1-\alpha)N_i+\alpha\sum_j N_j}{K'}\right)+\mu.$$

Analogously with the population model, different choices are possible for the rates $q_{i+}(\vec{N})$ and $q_{i-}(\vec{N})$. Here we use a generalization of the density-dependent mortality version of our population model (6),

$$q_{i+}(\vec{N}) = r_+ N_i + \mu,$$

$$q_{i-}(\vec{N}) = r_- N_i + rN_i\frac{(1-\alpha)N_i+\alpha\sum_j N_j}{K'}, \tag{13}$$

with $r = r_+ - r_-$. Both species-level density dependence (first term in nominator of $q_{i-}(\vec{N})$) and community-level density dependence (second term in nominator of $q_{i-}(\vec{N})$) are incorporated into the death rate. Transition rates (13) are only valid for $\alpha \ge 0$, because $q_{i-}(\vec{N})$ can become negative for $\alpha < 0$. For $\alpha \le 0$ we take

$$q_{i+}(\vec{N}) = r_+ N_i + rN_i\frac{(-\alpha)\sum_j N_j}{K'} + \mu,$$

$$q_{i-}(\vec{N}) = r_- N_i + rN_i\frac{(1-\alpha)N_i}{K'} \tag{14}$$

again with $r = r_+ - r_-$. Now species-level density dependence is part of the death rate, and community-level density dependence is part of the birth rate. Note that definitions (13) and (14) coincide for $\alpha = 0$.

Transition rates (13) and (14) define the stochastic community model that we study in this paper. Table 1 summarizes all model parameters. Putting $r = 1$ corresponds to fixing time units, which can be done without loss of generality. Once r is fixed, the parameter r_- can be interpreted as a measure of the intensity of demographic stochasticity. Indeed, increasing r_- does not affect the deterministic part of the community model, but augments the variance of stochastic fluctuations, as we show in Appendix D, see (D.3).

Table 1
Parameters of the stochastic community model.

Symbol	Meaning	Value
S_T	Total number of species (some species can be absent from the community)	$S_T=20$
r	Per capita intrinsic rate of population increase	$r=1$
K	Carrying capacity of community	$K=400$[a]
α	Interaction coefficient (competitive for $\alpha>0$, mutualistic for $\alpha<0$)	$\alpha\in[-0.05,1]$
r_-	Per capita intrinsic death rate (measure for intensity of demographic stochasticity)	$r_-\in[0.1,100]$
μ	Immigration rate of a species	$\mu\in[0.001,100]$
r_+	Per capita intrinsic birth rate	$r_+=r+r_-$
K'	Species-level carrying capacity	See (10)[a]

Three parameters (S_T,r,K) are unchanged in all figures (except in Fig. 5B); three parameters (α,r_-,μ) are varied over the ranges indicated; two parameters (r_+,K') are simple functions of the previous ones.

[a] In Fig. 5B the species-level carrying capacity K' is constant, $K'=100$, and the community-level carrying capacity K is computed from (10).

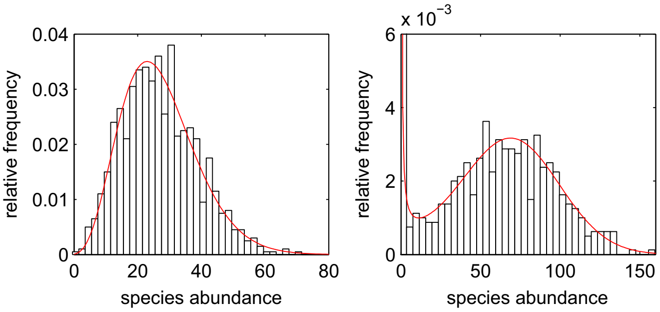

Fig. 3. Comparison between simulated and computed stationary species abundance distribution. The histograms give the species abundance distribution of a simulated trajectory. The model was simulated during 2.10^4 time units. The first half was used to eliminate the transient dynamics; the second half was used to sample the population size every 10 time units (1000 samples in total). The red curves correspond to the approximation (E6). Parameters are $r\_=1$, $\alpha=0.5$, $K=400$. Left panel: $\mu=10$. Right panel: $\mu=0.01$.

3.2. *Stationary distribution*

We study the stationary species abundance distribution of the stochastic community model defined by transition rates (13) and (14). A full derivation can be found in Appendix E; here we give an outline of the computations.

The stationary distribution can be solved exactly for two special cases. When $\alpha=0$, the interaction between individuals is purely intraspecific, and the S_T species have independent dynamics. The stationary distribution of the community model is the product of the stationary distributions of the various species, see (E.2). When $\alpha=1$, the interaction between individuals is completely symmetrical, i.e., individuals interact with each other independently of the species they belong to. In that case, the model belongs to a class of community models with community-level density dependence (Haegeman and Etienne, 2008), for which the stationary distribution is known explicitly, see (E.4).

For other values of the interaction coefficient α, $-1/(S_T-1)<\alpha<0$ and $0<\alpha<1$, we were unable to obtain an explicit expression for the stationary distribution. In Appendix E we present an approximation that matches closely the stationary distribution obtained from stochastic simulations, see Fig. 3. Moreover, by taking the limits $\alpha\to 0$ and $\alpha\to 1$ of our approximation, we recover (exactly, or with very good accuracy) the solutions for $\alpha=0$ and $\alpha=1$. We also compared the results obtained in Fig. 5A from simulating the stochastic process with our approximation: the results were almost identical. Therefore, we confidently use our approximation to investigate the stationary distribution of the community model.

We also derive a linear approximation for the community model and compute the corresponding stationary distribution in Appendix D. The linear approximation is useful to study the general behaviour of the community model. But the approximate stationary distribution can differ significantly from the exact solution, in particular when demographic stochasticity is important, as in the population model (Fig. 1).

3.2.1. *Stationary distribution as a function of α and μ*

Fig. 4 shows the stationary probability distribution for a population size N_i (left panel) and for the community size J (right panel),

$$J=\sum_{i=1}^{S_T} N_i.$$

Note that due to the species permutation symmetry of model (13) and (14), all species have the same population size distribution. We varied the immigration rate μ in each panel, and the interaction coefficient α between panels. We kept the community carrying capacity K constant, so that the species-level carrying capacity K' changes when varying α, see (10). The demographic stochasticity coefficient r_- was kept constant. Population and community sizes are represented on a logarithmic scale, using transformation (8).

For large immigration rates μ, the community and population size distributions are peaked; the center of these distributions coincides with the equilibrium value (12) of the deterministic model. The peak of community size is located at S_T times that of population size. Hence, the distributions exhibit almost no randomness, and are well described by the deterministic model. The community composition is a somewhat blurred image of the species pool. This deterministic behaviour is present for all values of α.

When the immigration rate μ decreases, the population size distribution gets wider, and its mode shifts towards smaller population size. In this case, some species become extinct due to demographic fluctuations. Note that their extinction is temporary, as immigration can reintroduce them from the species pool in the community. The dynamical balance between immigration and demographic stochasticity leads to a left-skewed population size distribution, which we also encountered in the population model (Fig. 1). For any interaction coefficient α, there is a range of parameter μ for which immigration and demographic stochasticity are balanced.

When the immigration rate μ decreases further, we have to distinguish the cases of positive and negative α. When $\alpha<0$, the community size distribution shifts to smaller values, becomes wider, and community disappearance becomes probable. When $\alpha>0$, the community size distribution keeps a constant peaked shape down to very small values of μ. In this case, community size regulation prevents community extinction. Simultaneously, the population size distribution shifts to larger values, becomes more peaked, and closely resembles the community size distribution. The community is then dominated by a few species that have stochastically excluded the other species.

3.3. *Community properties*

Here we perform a more systematic study of the stationary properties of model (13) and (14), in particular:

- the mean community size $\mathbb{E}[J]$;
- the variability of the community size, measured by its coefficient of variation $\mathrm{CV}[J]$;

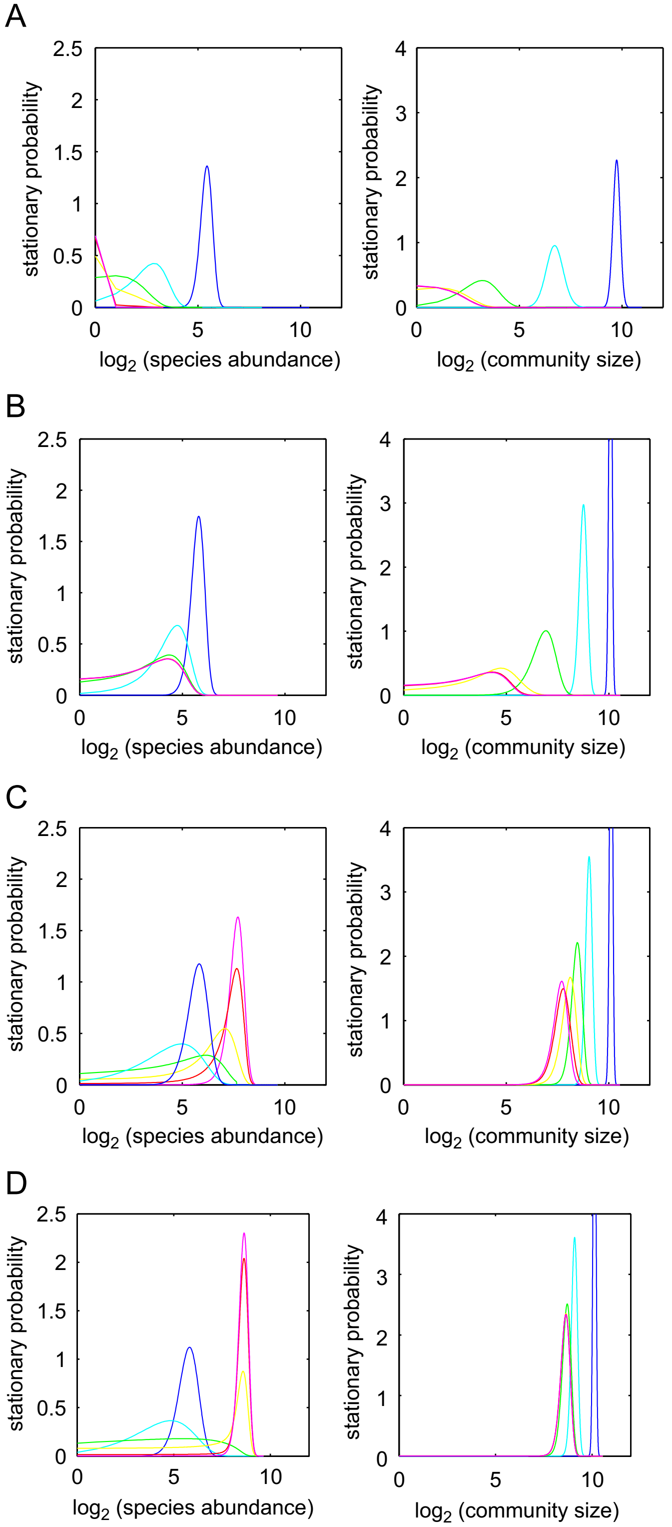

Fig. 4. Species abundance and community size distributions as functions of interaction coefficient α and immigration rate μ. The immigration rate μ takes values 0.001 (magenta), 0.01 (red), 0.1 (yellow), 1 (green), 10 (cyan), and 100 (blue). The demographic stochasticity coefficient r_- is the same for all distributions, $r_-=10$. The interaction coefficient α takes four different values: (A) $\alpha=-0.02$; (B) $\alpha=0$; (C) $\alpha=0.5$; (D) $\alpha=1$. The community carrying capacity K is constant, $K=400$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- the mean population size $\mathbb{E}[N_i]$; however, this property does not contain new information compared to $\mathbb{E}[J]$ because $\mathbb{E}[J]=S_T\mathbb{E}[N_i]$;
- the variability of the population size, measured by its coefficient of variation CV[N_i]; note that this property does contain new information compared to CV[J];
- the mean number of species $\mathbb{E}[S]$ in the community; note that $\mathbb{E}[S]=S_T(1-\mathbb{P}[N_i=0])$;
- the mean Simpson diversity index $\mathbb{E}[D]$ in the community, defined as the probability that two randomly sampled individuals from the community belong to different species.

Other properties can easily be derived from these properties. For example, community synchrony as defined by Loreau and de Mazancourt (2008) equals the ratio of CV[J] and CV[N_i]. Community evenness can be defined by comparing the Simpson diversity index $\mathbb{E}[D]$ and the mean species richness $\mathbb{E}[S]$. Most of these properties are readily obtained from the approximate stationary distribution of the community model. Those for which the computation is not straightforward are considered in Appendix F.

3.3.1. *Community properties as functions of α and μ*

Fig. 5 plots the above community properties as functions of the interaction coefficient α (on x-axis) and the immigration rate μ (different colors). Demographic stochasticity r_- is kept constant. Both negative and positive values of α are plotted; we use a finer scale for the mutualistic case (the scale for $\alpha<0$ is ten times finer than the scale for $\alpha>0$). To help interpret the results, we plot the same community properties obtained from the linear approximation in Fig. S2, in Appendix G.

In Fig. 5A we keep the community carrying capacity K constant when varying α, as in Fig. 4. In this case, there is no direct effect of the interaction coefficient α on the mean community size. Note that to keep K constant, the species-level carrying capacity K' has to decrease for decreasing α.

The mean community size $\mathbb{E}[J]$ increases with the immigration rate μ, as expected. More surprisingly, the mean community size also increases with the interaction coefficient α. As we keep the community carrying capacity K constant, the mean community size is independent of α in the linear approximation, see Fig. S2A in Appendix G. The dependence on α in the full model is due to demographic stochasticity, which can drive the community to extinction when μ and α are small. When $\alpha\approx1$, community-level density dependence prevents community extinction.

The variability of community size CV[J] decreases with μ, and decreases with α, reaching a minimum at $\alpha=1$. Community-level density dependence at $\alpha=1$ regulates the community size, decreasing its variability. For smaller and negative α, community extinction is increasingly probable, and variability increases steeply. The variability of population size CV[N_i] decreases with μ, decreases with α for $\alpha<0$, and increases with α for $\alpha>0$, reaching a minimum at $\alpha=0$. Species-level density dependence at $\alpha=0$ regulates the population sizes. For negative α, there is again a steep increase in variability due to demographic stochasticity.

The expected number of species $\mathbb{E}[S]$ increases with μ, and is maximal for $\alpha=0$. The latter result is due to density dependence regulating the size of each population, so that population extinction is less probable. The number of species decreases for negative α because the entire community can disappear; the number of species decreases for positive α because the community is increasingly dominated by a few species, and eventually (for small μ and $\alpha\approx1$) by a single species. The Simpson diversity index $\mathbb{E}[D]$ has a similar behaviour.

In Fig. 5B we keep the species-level carrying capacity K' constant when varying α. As a consequence, the interaction coefficient α directly affects the mean community size. Increasing competition decreases community size, and increasing mutualism increases community size.

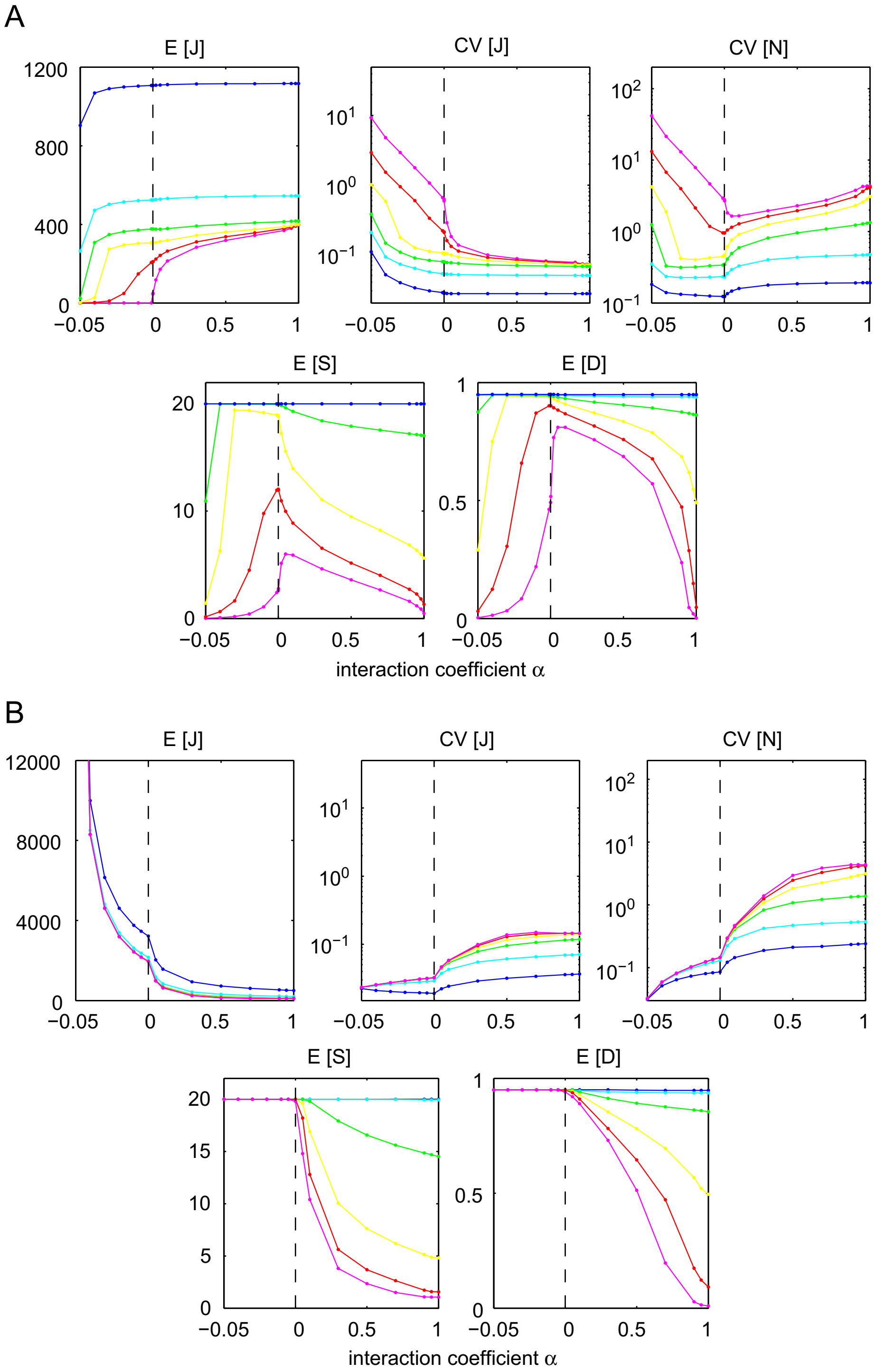

Fig. 5. Community properties as functions of interaction coefficient α for different immigration rates μ. The immigration rate varies from 0.001 to 100 with the same color code as in Fig. 4. The demographic stochasticity coefficient r_- is the same for all curves: $r_-=1$. The scale for mutualistic interactions ($\alpha<0$) is 10 times finer than the scale for competitive interactions ($\alpha>0$). In part (A) the community-level carrying capacity K is constant, $K=400$; in part (B) the species-level carrying capacity K' is constant, $K'=100$.

Note that to keep K' constant, the community carrying capacity K has to increase steeply for smaller α, especially when $\alpha<0$.

When α decreases, the mean community size $\mathbb{E}[J]$ increases, and the variabilities CV[J] and CV[N] decrease. Mutualistic (or less competitive) interactions lead to large population sizes, eliminating entirely the effect of demographic stochastic observed in Fig. 5A. When α is negative, the number of species $\mathbb{E}[S]$ and the Simpson diversity index $\mathbb{E}[D]$ reach their maximal value. Community extinction is extremely improbable and all populations have the same size because demographic stochasticity does not affect large populations.

3.3.2. Community properties as functions of r_- and μ

Fig. 6 plots the community variables as functions of parameters r_- (on x-axis) and μ (different colors) for either positive (Fig. 6A,

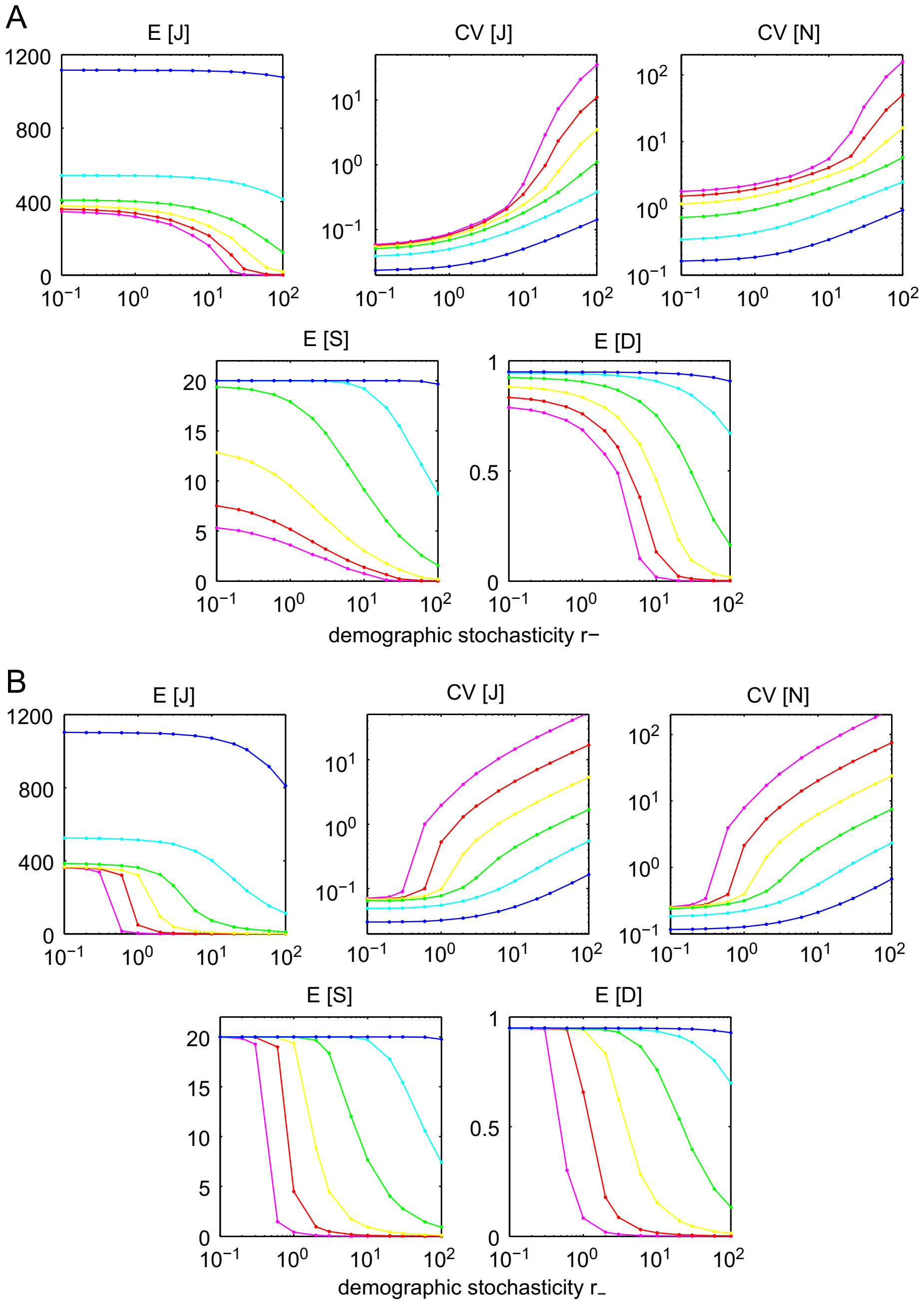

Fig. 6. Community properties as functions of demographic stochasticity coefficient r_- for different immigration rates μ. The immigration rate varies from 0.001 to 100 with the same color code as in Fig. 4. The interaction coefficient α takes two different values: (A) $\alpha = 0.5$, (B) $\alpha = -0.02$. The community carrying capacity K is constant, $K = 400$.

$\alpha = 0.5$) or negative (Fig. 6B, $\alpha = -0.02$) values of α. The same community properties obtained from the linear approximation are plotted in Fig. S3 in Appendix G for comparison.

The mean community size decreases with demographic stochasticity r_-, especially for small immigration rates μ, both for positive and negative α. This is due to the increased probability of species extinction. The variability of community and population size increases with r_-, both for positive and negative α. Increasing demographic stochasticity increases the probability that either individual populations ($\alpha > 0$) or the community as a whole ($\alpha < 0$) disappear, which further increases the variability of population and community sizes. Similarly, the mean number of species and the Simpson diversity index decrease with r_- due to demographic stochasticity. This decrease is gradual for $\alpha > 0$, as species disappear one by one, and more abrupt for $\alpha < 0$, as the entire community can disappear at once.

4. Discussion

We have proposed a minimal community model that combines the basic ingredients of niche-based and neutral community models. Our model can be interpreted as Hubbell's local community model in which we have replaced the condition of invariant community size with a dynamical regulation of population and community sizes by intra- and interspecific interactions. Alternatively, it can be interpreted as a classical Lotka–Volterra model to which we have added demographic stochasticity and immigration from an external species pool. The model is minimal in the sense that we exploited species symmetry as much as possible: all species have the same birth rate, death rate, carrying capacity and immigration rate, and all species pairs have the same interaction coefficient.

We have presented a detailed analysis of the properties of the stationary state as functions of three model parameters: the immigration rate μ, the demographic stochasticity intensity r_- and the interaction coefficient α, defined as the ratio of inter- and intraspecific interaction strength. The general conclusions of this analysis can be summarized as follows:

- For strong immigration and weak demographic stochasticity, community structure is deterministic, i.e., predictable. The local community is a faithful representation of the species pool. The noise in this representation increases by lowering the immigration rate μ or raising the strength of demographic stochasticity r_-.
- Increasing noise when species interactions are competitive (positive and not too small interaction coefficient α) yields a community in which some species start to dominate others. As all species are equally abundant in the species pool, these fluctuations cannot be predicted. Community size remains approximately constant. Further increasing noise eventually leads to a community in which one species (stochastically) excludes all others.
- Increasing noise when species interactions are mutualistic or weakly competitive (negative, or positive but small interaction coefficient α) yields a community in which not only populations but also the community as a whole are subject to random fluctuations. This effect is stronger when interactions are mutualistic, as species abundances are positively correlated. Further increasing noise (again, by decreasing μ or increasing r_-) eventually leads to the collapse of the entire community.

Mutualistic interactions result in unstable communities, in which even a small amount of noise can induce high variability. However, this finding is strongly dependent on the assumption that the community-level carrying capacity K is kept constant when reducing the interaction coefficient α. If, alternatively, we keep the species-level carrying capacity K' constant, we obtain very different results. In the latter case, mutualistic interactions lead to large and stable communities, because demographic stochasticity has a relatively small impact on large populations, see Fig. 5B. It should be noted that the deterministic Lotka–Volterra model for mutualistic interactions also predicts rapidly increasing population sizes. This somewhat pathological behaviour has generally been considered as an indication that the Lotka–Volterra model is too simplistic to give a realistic description of mutualism (Wolin and Lawlor, 1984; Ringel et al., 1996).

Our model does not exactly recover Hubbell's local community model as a limiting case. The assumption of invariant community size in Hubbell's model imposes a strict regulation of the community size, which cannot be reproduced exactly by our more flexible model. However, strict community size regulation can be approximated well by community-level density dependence. Indeed, Fig. 4D shows that the distribution of the community size J is sharply peaked for $\alpha = 1$. Moreover, the stationary distribution of our model with $\alpha = 1$ conditional on a given community size J is identical to the stationary distribution of Hubbell's model, as can be seen explicitly from solution (E.5) in Appendix E. Hence, the community structure predicted by our model with $\alpha = 1$ is close to Hubbell's model predictions. For example, the population size distributions in Fig. 4D coincide with Hubbell's.

It has been observed previously that Hubbell's model is robust to the introduction of niche features. Volkov et al. (2005) noticed that the immigration process in neutral models can be reinterpreted as a particular form of species-level density dependence. Note that Hubbell's neutral model includes an implicit but strong form of community-level density dependence, as community size is kept constant over time. Surprisingly, Etienne et al. (2007) showed that dropping the invariant community size condition does not affect the stationary species abundance distribution (conditional on community size), so that Hubbell's model can be interpreted as a neutral model without any density-dependence. Other papers introduced a class of neutral-like models with community-level density dependence (Haegeman and Etienne, 2008; Allouche and Kadmon, 2009), for which the stationary distribution (conditional on community size) is identical to Hubbell's. In fact, our model with $\alpha = 1$ belongs to this class of models.

Since Hubbell's neutral model is embedded in our neutral-niche model as a limiting case when $\alpha = 1$, we can ask how the community structure predicted by neutral theory changes when taking into account niche processes, i.e., when going from $\alpha = 1$ to $\alpha < 1$. Our model shows that the change in community structure is relatively limited. For example, the community properties in Fig. 5A have a rather smooth transition from $\alpha = 1$ to $\alpha < 1$. Only the Simpson diversity index $\mathbb{E}[D]$ changes more abruptly, because the stochastic exclusion of species is prevented by a limited amount of niche differentiation. Although the set of abundance distributions is larger for $\alpha < 1$ than for $\alpha = 1$, community structure when $\alpha < 1$ is typically well approximated by community structure for $\alpha = 1$, possibly with different parameters μ and r_-. It seems therefore difficult to infer the set of model parameters from species abundance data.

Volkov et al. (2009) studied a stochastic community model with birth-death events and species interactions, which therefore has some relationship to our model. They used their model to estimate interaction coefficients based on empirical abundance data. Specifically, they divided a large plot of tropical forest into a myriad of small quadrats, and they argued that the resulting replicated data sets suffice to reliably infer species interactions. This approach might provide a connection between our model and empirical data. Alternatively, parameter estimation could be based on spatial and/or dynamical data, or data from different environmental conditions (e.g., varying immigration rates). A rigorous investigation of the parameter inference problem requires further work and falls outside the scope of this paper.

Species interactions in our model have a particular structure, as we assumed equal interaction strength between all species pairs. This symmetry assumption for species interactions is a natural extension of the neutrality assumption, in which birth, death and immigration rates of all species are assumed to be equal. On the other hand, it is a rather uncommon assumption in niche models. Symmetric interactions correspond to a niche space with identical overlaps between all species pairs, which is only possible in a high-dimensional niche space (e.g., as many dimensions as there are species). The more common interaction structure, in which species are sorted along a one-dimensional niche axis, and which is often considered in neutral-niche simulation models (Tilman, 2004; Schwilk and Ackerly, 2005; Gravel et al., 2006; Zillio and Condit, 2007), is not compatible with species symmetry. Note that symmetric interactions can also be interpreted as an approximation in which individuals effectively interact with all other species grouped together, analogous to the mean field approximation in physics.

A description of species sorting along a niche axis requires species-specific interaction coefficients, complicating model analysis. Species differences can be introduced more straightforwardly as species-specific birth and/or death rates (Zhang and Lin, 1997; Fuentes, 2004; Zhou and Zhang, 2008). Studies that have done so showed that even small demographic differences can perturb neutral community patterns, such as species abundance distributions. It would be interesting to see how the relative fragility of neutral models with respect to species differences, as found in these studies, interacts with the relative robustness of neutral model predictions with respect to the addition of niche processes, as we found in this work. Conceptual models suggest that the effect of species differences will be mitigated by niche processes (Chesson, 2000; Adler et al., 2007).

Our model analysis is restricted to the stationary structure of a single local community. It would be interesting to look at spatial and dynamical properties of our model, and see how niche processes affect neutral community behaviour. Although a direct analysis might be difficult, moment closure techniques might be helpful (Bolker et al., 2000). These techniques have been used successfully to study the stochastic logistic population model (Nåsell, 2003), and are known to be exact for Hubbell's neutral community model (Vanpeteghem et al., 2008; Vanpeteghem and Haegeman, 2010). Alternatively, the spatial and dynamical behaviour can be studied using the linear approximation. We have indicated the parameter region in which the linear approximation predicts the stationary distribution accurately. Note that the linear approximation has been used to study the effect of environmental stochasticity on community structure (Ives et al., 1999; Loreau and de Mazancourt, 2008).

Finally, it is worthwhile to note that we have constructed our stochastic community model using a mathematically natural construction. The only choice we had to make was how to distribute density dependence over birth and death rates. Our study of the population model, however, suggested that the details of this choice have little effect on the model's stationary properties. Apart from this peculiarity, our model shares the genericity of the Lotka–Volterra model and of Hubbell's neutral model. Also, we exploited a number of analytical tools (linear approximation, exact and approximate stationary distribution, community properties) to obtain a rather complete picture of the model behaviour. We hope that these tools and the model's genericity will be instrumental in narrowing the conceptual gap between niche and neutral theories in community ecology.

Acknowledgments

Michel Loreau acknowledges support from the Natural Sciences and Engineering Council of Canada and the Canada Research Chair programme.

Appendix A. Random variables and stochastic models

Stochastic models require, compared to deterministic models, some dedicated notation, which we define in this appendix. We use a simplified notation in the main text, to keep it as readable as possible; we use a more specialized notation in the appendices, to clearly present the mathematical arguments.

We use bold capital letters to denote random variables. We distinguish, for example, a particular value N for the population size, and the corresponding random variable $\boldsymbol{N}$. This distinction allows us to write expressions as $\mathbb{P}[\boldsymbol{N}=N]$, which stands for the probability that the random population size $\boldsymbol{N}$ takes the value N. We use the simplified notation $\mathbb{P}[N]$ in the main text, or when confusion is impossible.

The expectation (or average value) of the random variable $\boldsymbol{N}$ is denoted by $\mathbb{E}[\boldsymbol{N}]$,

$$\mathbb{E}[\boldsymbol{N}] = \sum_{N=0}^{\infty} N\mathbb{P}[N].$$

Similarly, we use $\mathrm{Var}[\boldsymbol{N}]$ for the variance,

$$\mathrm{Var}[\boldsymbol{N}] = \mathbb{E}[(\boldsymbol{N}-\mathbb{E}[\boldsymbol{N}])^2]$$

and $\mathrm{CV}[\boldsymbol{N}]$ for the coefficient of variation,

$$\mathrm{CV}[\boldsymbol{N}] = \frac{\sqrt{\mathrm{Var}[\boldsymbol{N}]}}{\mathbb{E}[\boldsymbol{N}]}.$$

The covariance between, for example, the abundance $\boldsymbol{N}_i$ of species i and the abundance $\boldsymbol{N}_j$ of species j is denoted by $\mathrm{Cov}[\boldsymbol{N_i},\boldsymbol{N_j}]$,

$$\mathrm{Cov}[\boldsymbol{N_i},\boldsymbol{N_j}] = \mathbb{E}[(\boldsymbol{N_i}-\mathbb{E}[\boldsymbol{N_i}])(\boldsymbol{N_j}-\mathbb{E}[\boldsymbol{N_j}])].$$

We use standard notation to denote the conditioning of one random variable on the value of another random variable. For example, the probability that the abundance $\boldsymbol{N}_i$ of species i takes the value N_i given that the community size $\boldsymbol{J}$ equals J is denoted by $\mathbb{P}[\boldsymbol{N}_i=N_i|\boldsymbol{J}=J]$, or simply $\mathbb{P}[N_i|J]$. Similarly, we use the notation $\mathbb{E}[\boldsymbol{N}_i|\boldsymbol{J}=J]$ or $\mathbb{E}[N_i|J]$ for the conditional expectation, and $\mathrm{Var}[\boldsymbol{N}_i|\boldsymbol{J}=J]$ or $\mathrm{Var}[N_i|J]$ for the conditional variance.

The dynamical variables of stochastic models are random variables. Different formalisms exist to describe the dynamics of (continuous-time Markovian) stochastic models (Van Kampen, 1997; Gardiner, 2004). The one we mainly use, the master equation formalism, is based on a dynamical equation (called the master equation) for the probability distribution of the random dynamical variables. For example, a stochastic model for the population size $\boldsymbol{N}$ is described by a system of differential equations for $\mathbb{P}[\boldsymbol{N}=N]$ (one equation for each value N). Examples of the master equation formalism are (C.1) and (E.1).

Alternatively, one can use the formalism of stochastic differential equations, which are dynamical equations for the random variables directly (and not for their probability distribution). To get an intuitive idea of these equation, consider a small time interval δt, during which the random dynamical variable $\boldsymbol{N}$ changes by an amount $\delta\boldsymbol{N}$. The stochastic differential equation describes the dependence of $\delta\boldsymbol{N}$ on the current value of the dynamical variable $\boldsymbol{N}$, together with new randomness appearing in the time interval δt. Examples of stochastic differential equations are (B.3) and (D.5).

Appendix B. Linear approximation of population model

We decompose the population model of Section 2 into a deterministic part and a purely stochastic part. We use this decomposition to derive a linear approximation, which allows us to quantify the impact of stochasticity on the deterministic population model (1).

To define the transition rates of the population model, consider the change $\delta\boldsymbol{N}$ of population size in a small time interval δt. The probability of making a transition in this time interval is proportional to δt, and the transition rates are the constants of proportionality. More precisely,

$$\mathbb{P}[\delta\boldsymbol{N}=+1|\boldsymbol{N}=N] = q_+(N)\delta t,$$

$$\mathbb{P}[\delta\boldsymbol{N}=-1|\boldsymbol{N}=N] = q_-(N)\delta t,$$

$$\mathbb{P}[\delta\boldsymbol{N}=0|\boldsymbol{N}=N] = 1-(q_+(N)+q_-(N))\delta t.$$

We compute the mean of $\delta\boldsymbol{N}$ conditioned on $\boldsymbol{N}=N$,

$$\begin{aligned}\mathbb{E}[\delta\boldsymbol{N}|\boldsymbol{N}=N] &= (+1)q_+(N)\delta t+(-1)q_-(N)\delta t\\ &= (q_+(N)-q_-(N))\delta t,\end{aligned} \tag{B.1}$$

and the variance of $\delta\boldsymbol{N}$ conditioned on $\boldsymbol{N}=N$,

$$\begin{aligned}\mathrm{Var}[\delta\boldsymbol{N}|\boldsymbol{N}=N] &= \mathbb{E}[(\delta\boldsymbol{N})^2|\boldsymbol{N}=N]-\mathbb{E}[\delta\boldsymbol{N}|\boldsymbol{N}=N]^2\\ &\approx \mathbb{E}[(\delta\boldsymbol{N})^2|\boldsymbol{N}=N]\\ &= (+1)^2q_+(N)\delta t+(-1)^2q_-(N)\delta t\\ &= (q_+(N)+q_-(N))\delta t,\end{aligned} \tag{B.2}$$

where we dropped terms in $(\delta t)^2$.

Hence, in the small time interval δt, the population size $\boldsymbol{N}$ undergoes a deterministic change given by (B.1) with a stochastic fluctuation superposed on it. The mean of this stochastic fluctuation equals zero, and its variance is given by (B.2). Formally, this decomposition can be written as

$$\begin{aligned}\delta\boldsymbol{N} &= \mathbb{E}[\delta\boldsymbol{N}]+\sqrt{\mathrm{Var}[\delta\boldsymbol{N}]}\boldsymbol{E}\\ &=(q_+(\boldsymbol{N})-q_-(\boldsymbol{N}))\delta t+\sqrt{q_+(\boldsymbol{N})+q_-(\boldsymbol{N})}\sqrt{\delta t}\boldsymbol{E},\end{aligned} \tag{B.3}$$

with $\boldsymbol{E}$ an appropriate random variable, with mean zero and variance one. Neglecting the purely stochastic second term, we get the corresponding deterministic dynamical system, given by the differential equation

$$\frac{\mathrm{d}N}{\mathrm{d}t}=q_+(N)-q_-(N).$$

The stochastic differential equation (B.3) is difficult to analyze in general. A useful approximation consists in (a) linearizing the deterministic part around a stable equilibrium point N^*, and (b) replacing the stochastic part by a Gaussian random variable with mean zero and constant variance, equal to the variance of the full equation at the equilibrium point N^*. We get

$$\begin{aligned}\delta\boldsymbol{N} &= \left(\frac{\mathrm{d}q_+}{\mathrm{d}N}(N^*)-\frac{\mathrm{d}q_-}{\mathrm{d}N}(N^*)\right)(\boldsymbol{N}-N^*)\delta t+\sqrt{q_+(N^*)+q_-(N^*)}\sqrt{\delta t}\boldsymbol{G}\\ &=\lambda(\boldsymbol{N}-N^*)\delta t+\sigma\sqrt{\delta t}\boldsymbol{G},\end{aligned} \tag{B.4}$$

with $\boldsymbol{G}$ a Gaussian random variable with mean zero and variance one, λ the slope of the deterministic equation at N^* ($\lambda<0$ for a stable equilibrium point), and σ^2 the variance of the full stochastic equation at N^*. The linear stochastic differential equation (B.4) is known as an auto-regressive model (in discrete time), or a Ornstein–Uhlenbeck process (in continuous time). The stationary distribution for the population size $\boldsymbol{N}$ is Gaussian with mean N^* and variance $\sigma^2/2|\lambda|$ (Gardiner, 2004).

Another approximation for (B.3) is based on the same linearization ideas, but uses the logarithmic population size $\boldsymbol{L}=\ln\boldsymbol{N}$ as model variable. We have

$$\delta\boldsymbol{L}=\ln(\boldsymbol{N}+\delta\boldsymbol{N})-\ln(\boldsymbol{N})=\ln\left(1+\frac{\delta\boldsymbol{N}}{\boldsymbol{N}}\right)\approx\frac{\delta\boldsymbol{N}}{\boldsymbol{N}}$$

for small changes $\delta\boldsymbol{N}$, which is satisfied in a continuum approximation. The linearized stochastic differential equation reads

$$\delta\boldsymbol{L}=\left(\frac{\mathrm{d}q_+}{\mathrm{d}N}(N^*)-\frac{\mathrm{d}q_-}{\mathrm{d}N}(N^*)\right)(\boldsymbol{L}-L^*)\delta t+\frac{1}{N^*}\sqrt{q_+(N^*)+q_-(N^*)}\sqrt{\delta t}\boldsymbol{G}, \tag{B.5}$$

with equilibrium logarithmic population size $L^*=\ln N^*$. The linearizations (B.4) and (B.5) are similar, with one notable difference. Whereas Eq. (B.4) has a Gaussian stationary distribution for the population size $\boldsymbol{N}$, Eq. (B.5) has a Gaussian stationary distribution for the logarithmic population size $\boldsymbol{L}$, and thus a lognormal stationary distribution for the population size $\boldsymbol{N}$.

Appendix C. Stationary distribution of population model

We consider the population model of Section 2, and derive the stationary distribution of the population size $\boldsymbol{N}$ using the master equation formalism. The master equation is a differential equation for the distribution $\mathbb{P}[\boldsymbol{N}=N]=\mathbb{P}[N]$ of the population size $\boldsymbol{N}$, see Van Kampen (1997). It reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}[N]=q_+(N-1)\mathbb{P}[N-1]+q_-(N+1)\mathbb{P}[N+1]-(q_+(N)+q_-(N))\mathbb{P}[N] \tag{C.1}$$

and expresses how the probability $\mathbb{P}[N]$ changes as a function of time: $\mathbb{P}[N]$ increases by transitions $N-1\rightarrow N$ (first term in right-hand side) and $N+1\rightarrow N$ (second term in right-hand side); $\mathbb{P}[N]$ decreases by transitions $N\rightarrow N+1$ and $N\rightarrow N-1$ (last term in right-hand side).

We are looking for the stationary solution of (C.1), i.e., the solution of the set of equations obtained by putting the right-hand side to zero,

$$\begin{aligned}&q_+(N-1)\mathbb{P}[N-1]+q_-(N+1)\mathbb{P}[N+1]\\ &\quad=(q_+(N)+q_-(N))\mathbb{P}[N]\quad\text{for all N.}\end{aligned}$$

This equation says that in stationary regime, the transitions arriving in state N (left-hand side) are compensated by the transitions leaving state N (right-hand side). A stronger condition, called detailed balance, is

$$q_+(N-1)\mathbb{P}[N-1]=q_-(N)\mathbb{P}[N]\quad\text{for all N}$$

stating that the transition $N-1\rightarrow N$ is directly compensated by the transition $N\rightarrow N-1$. If there exists a solution of the detailed-balance condition, then this solution is necessarily the stationary distribution (Van Kampen, 1997). For the population model, the solution of the detailed-balance condition exists and can be constructed explicitly. To do so, we express $\mathbb{P}[N]$ in terms of $\mathbb{P}[N-1]$, and by iterating we get $\mathbb{P}[N]$ in terms of $\mathbb{P}[0]$,

$$\mathbb{P}[N]=\mathbb{P}[0]\prod_{k=1}^{N}\frac{q_+(k-1)}{q_-(k)}. \tag{C.2}$$

We obtain $\mathbb{P}[0]$ by requiring that the distribution is normalized,

$$1=\sum_{N=0}^{\infty}\mathbb{P}[N]=\mathbb{P}[0]\left(1+\sum_{N=1}^{\infty}\prod_{k=1}^{N}\frac{q_+(k-1)}{q_-(k)}\right). \tag{C.3}$$

Eqs. (C.2)–(C.3) determine the stationary population size distribution $\mathbb{P}[N]$.

Substituting transition rates (6) in the stationary distribution (C.2), we get

$$\mathbb{P}[N]=\mathbb{P}[0]\frac{(a)_N}{(b)_N}\frac{c^N}{N!},$$

where we introduced the dimensionless paramaters a, b and c,

$$a=\frac{\mu}{r_+},\quad b=\frac{r_-}{r}K+1,\quad c=\frac{r_+}{r}K,$$

and we used the Pochhammer notation,

$$(a)_N=a(a+1)\cdots(a+N-1).$$

The normalization condition (C.3) can be written in terms of the hypergeometric function $\Phi_{a,b}(c)$ (sometimes called confluent hypergeometric function, or also Kummer's function),

$$\Phi_{a,b}(c)=\sum_{N=0}^{\infty}\frac{(a)_N}{(b)_N}\frac{c^N}{N!} \tag{C.4}$$

so that the stationary distribution can be written as

$$\mathbb{P}[N]=\frac{1}{\Phi_{a,b}(c)}\frac{(a)_N}{(b)_N}\frac{c^N}{N!}. \tag{C.5}$$

Analogously, the stationary distribution for the alternative choice (7) can be obtained by substituting (7) in (C.2)–(C.3).

Appendix D. Linear approximation of community model

We decompose the community model of Section 3 into a deterministic part and a purely stochastic part. We compute a linear approximation, and quantify the impact of stochasticity on the deterministic community model (11).

First, we define the transition rates of the community model. In a small time interval δt, one of the components of the abundance vector $\vec{\boldsymbol{N}} = (\boldsymbol{N}_1, \boldsymbol{N}_2, \ldots, \boldsymbol{N}_{S_T})$ can increase or decrease by one individual. Hence, the vector of random abundance changes $\delta\vec{\boldsymbol{N}} = (\delta\boldsymbol{N}_1, \delta\boldsymbol{N}_2, \ldots, \delta\boldsymbol{N}_{S_T})$ can increase or decrease by the unit vector $\vec{e}_i$,

$$\vec{e}_i = (0,0,\ldots,0,1,0,\ldots,0,0) \tag{D.1}$$

a vector with all components equal to zero, except component i which is equal to one. The transition rates $q_{i+}(\vec{N})$ and $q_{i-}(\vec{N})$ are given by

$$\mathbb{P}[\delta\vec{\boldsymbol{N}} = +\vec{e}_i | \vec{\boldsymbol{N}} = \vec{N}] = q_{i+}(\vec{N})\delta t,$$

$$\mathbb{P}[\delta\vec{\boldsymbol{N}} = -\vec{e}_i | \vec{\boldsymbol{N}} = \vec{N}] = q_{i-}(\vec{N})\delta t.$$

We compute the mean of $\delta\boldsymbol{N}_i$,

$$\mathbb{E}[\delta\boldsymbol{N}_i | \vec{\boldsymbol{N}} = \vec{N}] = (q_{i+}(\vec{N}) - q_{i-}(\vec{N}))\delta t, \tag{D.2}$$

the variance of $\delta\boldsymbol{N}_i$,

$$\mathrm{Var}[\delta\boldsymbol{N}_i | \vec{\boldsymbol{N}} = \vec{N}] \approx \mathbb{E}[(\delta\boldsymbol{N}_i)^2 | \vec{\boldsymbol{N}} = \vec{N}] = (q_{i+}(\vec{N}) + q_{i-}(\vec{N}))\delta t \tag{D.3}$$

and the covariance of $\delta\boldsymbol{N}_i$ and $\delta\boldsymbol{N}_j$,

$$\mathrm{Cov}[\delta\boldsymbol{N}_i, \delta\boldsymbol{N}_j | \vec{\boldsymbol{N}} = \vec{N}] \approx \mathbb{E}[\delta\boldsymbol{N}_i \delta\boldsymbol{N}_j | \vec{\boldsymbol{N}} = \vec{N}] = 0, \tag{D.4}$$

where we dropped terms in $(\delta t)^2$. Note that the absence of correlations is only valid instantaneously, i.e., for the abundance changes during a single event, but it holds both in and out of the stationary regime.

Eqs. (D.2)–(D.4) suggest a decomposition into a deterministic part and a purely stochastic part. Stochastic fluctuations act on a deterministic dynamical system given by

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = q_{i+}(\vec{N}) - q_{i-}(\vec{N}), \quad i = 1,2,\ldots,S_T.$$

The full stochastic model can be analyzed using a linear approximation (Gardiner, 2004). We linearize the deterministic equation at a stable equilibrium point $\vec{N}^*$, yielding the coefficient matrix A,

$$A_{ij} = \frac{\partial}{\partial N_j}(q_{i+}(\vec{N}) - q_{i-}(\vec{N}))\Big|_{\vec{N} = \vec{N}^*}.$$

Stability means that all eigenvalues λ_i of the matrix A have negative real part. The stochastic fluctuations act additively on the components $\delta\boldsymbol{N}_i$, with mean zero and variance σ_i^2, see (D.3),

$$\sigma_i^2 = q_{i+}(\vec{N}^*) + q_{i-}(\vec{N}^*)$$

and without correlation between different components $\delta\boldsymbol{N}_i$ and $\delta\boldsymbol{N}_j$, see (D.4). Hence, the linear stochastic differential equation is

$$\delta\vec{\boldsymbol{N}} = A(\vec{\boldsymbol{N}} - \vec{N}^*)\delta t + B\sqrt{\delta t}\vec{\boldsymbol{G}}, \tag{D.5}$$

with B a diagonal matrix with components σ_i, and $\vec{\boldsymbol{G}}$ a vector of mutually independent Gaussian random variables with mean zero and variance one.

The linear stochastic differential equation (D.5) is well known (auto-regressive model or Ornstein–Uhlenbeck process). Its stationary distribution for the abundance vector $\vec{\boldsymbol{N}}$ is Gaussian with mean $\vec{N}^*$ and covariance matrix C, which is the solution of the Lyapunov equation (Gardiner, 2004):

$$AC + CA^{\mathrm{T}} + B^2 = 0, \tag{D.6}$$

where A^{T} stands for the transpose of the matrix A.

For the community model with transition rates (13)–(14), the Lyapunov equation (D.6) can be solved explicitly (Ives et al., 1999; Loreau and de Mazancourt, 2008). The matrix A is given by

$$A_{ij} = \begin{cases} a_1 & \text{if } i = j, \\ a_2 & \text{if } i \neq j, \end{cases}$$

with

$$a_1 = r\left(1 - \frac{2 + \alpha(S_T - 1)}{1 + \alpha(S_T - 1)} \frac{S_T N^*}{K}\right),$$

$$a_2 = -r\frac{\alpha}{1 + \alpha(S_T - 1)} \frac{S_T N^*}{K},$$

with equilibrium abundance N^* given by (12). The matrix B is given by

$$B_{ij} = \begin{cases} b_1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

with

$$b_1 = \begin{cases} (r_+ + r_-)N^* + r\dfrac{S_T N^*}{K} + \mu & \text{if } \alpha \geq 0, \\ (r_+ + r_-)N^* + r\dfrac{1 - \alpha(S_T + 1)}{1 + \alpha(S_T - 1)} \dfrac{S_T N^*}{K} + \mu & \text{if } \alpha \leq 0. \end{cases}$$

Both matrices A and B have a special structure: all diagonal components are equal, and all off-diagonal components are equal. The correlation matrix C has the same structure, and is explicitly given by

$$C_{ij} = \begin{cases} -\dfrac{(a_1 + (S_T - 2)a_2)b_1}{2(a_1 - a_2)(a_1 + (S_T - 1)a_2)} & \text{if } i = j, \\ \dfrac{a_2 b_1}{2(a_1 - a_2)(a_1 + (S_T - 1)a_2)} & \text{if } i \neq j. \end{cases}$$

Hence, in the linear approximation, the variance of a species abundance $\boldsymbol{N}_i$ is

$$\mathrm{Var}[\boldsymbol{N}_i] = \mathrm{Var}[\boldsymbol{N}_1] = C_{11} = -\frac{(a_1 + (S_T - 2)a_2)b_1}{2(a_1 - a_2)(a_1 + (S_T - 1)a_2)}$$

and the variance of the community size $\boldsymbol{J}$ is

$$\mathrm{Var}[\boldsymbol{J}] = \sum_{ij} C_{ij} = S_T C_{11} + S_T(S_T - 1)C_{12} = -\frac{S_T b_1}{2(a_1 + (S_T - 1)a_2)}.$$

Note that, although there are no correlations between the components of $\delta\vec{\boldsymbol{N}}$, see (D.4), the dynamics generate correlations between the components of the stationary abundance vector $\vec{\boldsymbol{N}}$ (i.e., $C_{ij} \neq 0$ for $i \neq j$).

Appendix E. Stationary distribution of community model

We consider the community model of Section 3, and derive the stationary distribution of the species abundance vector $\vec{\boldsymbol{N}}$. As for the population model, we use the master equation, which is a differential equation for the probability distribution $\mathbb{P}[\vec{\boldsymbol{N}} = \vec{N}] = \mathbb{P}[\vec{N}]$, see Van Kampen (1997). It reads

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}[\vec{N}] = {} & \sum_i q_{i+}(\vec{N} - \vec{e}_i)\mathbb{P}[\vec{N} - \vec{e}_i] + \sum_i q_{i-}(\vec{N} + \vec{e}_i)\mathbb{P}[\vec{N} + \vec{e}_i] \\ & - \sum_i (q_{i+}(\vec{N}) + q_{i-}(\vec{N}))\mathbb{P}[\vec{N}], \end{aligned} \tag{E.1}$$

with $\vec{e}_i$ the i-th unit vector, see (D.1). Explicit expressions for the transition rates $q_{i+}(\vec{N})$ and $q_{i-}(\vec{N})$ are given in (13)–(14).

We give an exact solution of the stationary distribution for the cases $\alpha = 0$ and $\alpha = 1$. For the cases $\alpha < 0$ and $0 < \alpha < 1$, we present an approximation method that reproduces accurately simulation results.

E.1. Community model with $\alpha=0$

The community model with $\alpha=0$ corresponds to a community in which species are mutually independent. There is species-level density dependence: the growth rate of species i is limited by the other individuals of species i, but not by the individuals of another species $j\neq i$. Indeed, when $\alpha=0$, transition rates $q_{i+}(\vec{N})$ and $q_{i-}(\vec{N})$ only depend on $\vec{N}$ through N_i,

$$q_{i+}(\vec{N})=q_{i+}(N_i) \quad \text{and} \quad q_{i-}(\vec{N})=q_{i-}(N_i)$$

so that the master equation (E.1) decouples into S_T master equations for one-species abundance distributions. For species i,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}[N_i]={}&q_{i+}(N_i-1)\mathbb{P}[N_i-1]+q_{i-}(N_i+1)\mathbb{P}[N_i+1]\\&-(q_{i+}(N_i)+q_{i-}(N_i))\mathbb{P}[N_i],\end{aligned}$$

which is identical to the master equation (C.1) for the population model. The stationary distribution $\mathbb{P}[N_i]$ is given by (C.5),

$$\mathbb{P}[N_i]=\frac{1}{\Phi_{a,b}(c)}\frac{(a)_{N_i}}{(b)_{N_i}}\frac{c^{N_i}}{N_i!},$$

with the dimensionless parameters

$$a=\frac{\mu}{r_+}, \quad b=\frac{K}{S_\mathrm{T}}\frac{r_-}{r}+1, \quad c=\frac{K}{S_\mathrm{T}}\frac{r_+}{r}.$$

The stationary distribution for the full abundance vector $\vec{N}$ is a product of one-species abundance distributions,

$$\mathbb{P}[\vec{N}]=\frac{1}{\Phi_{a,b}(c)^{S_\mathrm{T}}}\prod_{i=1}^{S_\mathrm{T}}\frac{(a)_{N_i}}{(b)_{N_i}}\frac{c^{N_i}}{N_i!}. \tag{E.2}$$

E.2. Community model with $\alpha=1$

The community model with $\alpha=1$ corresponds to community-level density dependence: the growth rate of species i is limited by individuals of species i and species $j\neq i$ alike. Mathematically, when $\alpha=1$, transition rate $q_{i+}(\vec{N})$ depends only on N_i, and transition rate $q_{i-}(\vec{N})$ depends both on N_i and on $J=\sum_i N_i$, with linear dependence on N_i,

$$q_{i+}(\vec{N})=q_{i+}(N_i) \quad \text{and} \quad q_{i-}(\vec{N})=N_i\overline{q}_{i-}(J).$$

We use detailed balance to compute the stationary distribution (recall that a solution of the detailed-balance condition, if it exists, is necessarily the stationary distribution, Van Kampen, 1997),

$$q_{i+}(\vec{N}-\vec{e}_i)\mathbb{P}[\vec{N}-\vec{e}_i]=q_{i-}(\vec{N})\mathbb{P}[\vec{N}] \tag{E.3}$$

or,

$$q_{i+}(N_i-1)\mathbb{P}[\vec{N}-\vec{e}_i]=N_i\overline{q}_{i-}(J)\mathbb{P}[\vec{N}].$$

One can check that the detailed-balance condition is satisfied by

$$\begin{aligned}\mathbb{P}[\vec{N}]&=\frac{q_{i+}(N_i-1)}{N_i\overline{q}_{i-}(J)}\mathbb{P}[\vec{N}-\vec{e}_i]\\&=\left[\prod_{i=1}^{S_\mathrm{T}}\left(\frac{1}{N_i!}\prod_{k=0}^{N_i-1}q_{i+}(k)\right)\right]\left[\prod_{k=1}^{J}\frac{1}{\overline{q}_{i-}(k)}\right]\mathbb{P}[\vec{0}],\end{aligned}$$

with $\vec{0}$ the abundance vector of a community without individuals. Using the transition rates (13), normalization can be computed explicitly using the hypergeometric function (C.4). The resulting multi-species abundance distribution is

$$\mathbb{P}[\vec{N}]=\frac{1}{\Phi_{S_\mathrm{T}a,b}(c)}\left[\prod_{i=1}^{S_\mathrm{T}}\frac{(a)_{N_i}}{N_i!}\right]\frac{c^J}{(b)_J}, \tag{E.4}$$

with the dimensionless parameters

$$a=\frac{\mu}{r_+}, \quad b=K\frac{r_-}{r}+1, \quad c=K\frac{r_+}{r}.$$

The stationary multi-species abundance distribution conditioned on the total number of individuals J is

$$\mathbb{P}[\vec{N}|J]=\frac{J!}{(S_\mathrm{T}a)_J}\prod_{i=1}^{S_\mathrm{T}}\frac{(a)_{N_i}}{N_i!}. \tag{E.5}$$

The stationary distribution for the total number of individuals J is given by

$$\mathbb{P}[J]=\frac{1}{\Phi_{S_\mathrm{T}a,b}(c)}\frac{(S_\mathrm{T}a)_J c^J}{J!(b)_J}.$$

E.3. Community model for general α

The community model for general α contains both species-level and community-level density dependence. The detailed-balance trick we used to solve the cases $\alpha=0$ and $\alpha=1$ does not work in the general case: the detailed-balance equation (E.3) cannot be satisfied simultaneously. As we are not able to solve the full set of stationarity equations of (E.1), we introduce an approximation method to obtain the stationary distribution, and show using simulations that this approximation is accurate.

The approximation strategy consists in first, finding an approximation $Q(\vec{N}|J)$ for the distribution $\mathbb{P}[\vec{N}|J]$, the multi-species abundance distribution conditioned on the total number of individuals J, and next, using this approximate distribution $Q(\vec{N}|J)$ to compute an approximation $R(J)$ for the distribution $\mathbb{P}(J)$, the stationary distribution for the total number of individuals J. By combining the two, we get an approximate multi-species abundance distribution,

$$\mathbb{P}[\vec{N}]=\mathbb{P}[\vec{N}|J]\mathbb{P}[J]\approx Q(\vec{N}|J)R(J). \tag{E.6}$$

First, we consider the case $\alpha>0$ with transition rates (13). Transition rate $q_{i+}(\vec{N})$ depends only on N_i, but transition rate $q_{i-}(\vec{N})$ depends both on N_i and on $J=\sum_i N_i$. As we look for an approximation of the distribution $\mathbb{P}[\vec{N}|J]$ conditioned on $\boldsymbol{J}=J$, we expect to make a small error by assuming that the number of individuals $\boldsymbol{J}$ appearing in transition rate $q_{i-}(\vec{N})$ takes the fixed value J. Thus, we consider a modified stochastic community model with transition rates $q_{i+}(\vec{N})$ and $\tilde{q}_{i-}(\vec{N})$,

$$\tilde{q}_{i-}(\vec{N})=\left(r_-+r\frac{(1-\alpha)N_i+\alpha J}{K'}\right)N_i=\left(r_-+\frac{r\alpha J}{K'}\right)N_i+\frac{r(1-\alpha)}{K'}N_i^2 \tag{E.7}$$

in which J is no longer a variable but a parameter. Because transition rates $q_{i+}(\vec{N})$ and $\tilde{q}_{i-}(\vec{N})$ depend only on the abundance N_i, we can apply detailed balance to compute the stationary multi-species abundance distribution $Q(\vec{N})$ for the modified transition rate $\tilde{q}_{i-}(\vec{N})$. Conditioning this distribution on $\boldsymbol{J}=J$, we get the approximation $Q(\vec{N}|J)$ for $\mathbb{P}[\vec{N}|J]$.

The computation of the stationary distribution $Q(\vec{N})$ for transition rates $q_{i+}(\vec{N})$ and $\tilde{q}_{i-}(\vec{N})$ is analogous with the case $\alpha=0$. The result is

$$Q(\vec{N})=\frac{1}{\Phi_{a,b}(c)^{S_\mathrm{T}}}\prod_{i=1}^{S_\mathrm{T}}\frac{(a)_{N_i}}{(b)_{N_i}}\frac{c^{N_i}}{N_i!}, \tag{E.8}$$

with dimensionless parameters

$$a=\frac{\mu}{r_+}, \quad b=\frac{r_-K'+r\alpha J}{r(1-\alpha)}+1, \quad c=\frac{r_+K'}{r(1-\alpha)}.$$

Note the dependence of the parameter b on the parameter J.

To condition $Q(\vec{N})$ on $\boldsymbol{J}=J$, we use the product structure of the distribution $Q(\vec{N})$. The multi-species abundance distribution $Q(\vec{N})$ is the product of S_T one-species abundance distributions $Q_1(N_i)$,

$$Q_1(N_i)=\frac{1}{\Phi_{a,b}(c)}\frac{(a)_{N_i}}{(b)_{N_i}}\frac{c^{N_i}}{N_i!}.$$

Hence, the distribution $Q(J)$ for $\boldsymbol{J}$ is given by the S_T-fold convolution product of one-species abundance distributions Q_1,

$$Q(J) = \underbrace{(Q_1 * Q_1 * \cdots * Q_1)}_{S_T \text{ times}}(J) = Q_1^{*S_T}(J).$$

As a result, we obtain the approximation

$$\mathbb{P}[\vec{N}|J] \approx Q(\vec{N}|J) = \frac{Q(\vec{N})}{Q(J)}. \tag{E.9}$$

To compute the approximation $R(J)$ for the distribution $\mathbb{P}[J]$, we consider the transition rates for the total community size J,

$$q_+(\vec{N}) = \sum_{i=1}^{S_T} q_{i+}(\vec{N}) \quad \text{and} \quad q_-(\vec{N}) = \sum_{i=1}^{S_T} q_{i-}(\vec{N}).$$

The transition rate $q_+(\vec{N})$ from J to $J+1$ depends only on $\vec{N}$ through J,

$$q_+(\vec{N}) = q_+(J) = r_+J + \mu S_T.$$

The transition rate $q_-(\vec{N})$ from J to $J-1$ depends also on the Simpson diversity of the community:

$$q_-(\vec{N}) = r_-J + \frac{r\alpha}{K'}J^2 + \frac{r(1-\alpha)}{K'}\sum_{i=1}^{S_T} N_i^2.$$

We use the approximate distribution $Q(\vec{N}|J)$ to compute an approximation $\tilde{\mathbb{E}}[q_-(\vec{N})|J]$ for the expected transition rate $\mathbb{E}[q_-(\vec{N})|J]$,

$$\mathbb{E}[q_-(\vec{N})|J] \approx \tilde{\mathbb{E}}[q_-(\vec{N})|J] = r_-J + \frac{r\alpha}{K'}J^2 + \frac{r(1-\alpha)}{K'}S_T\tilde{\mathbb{E}}[N_i^2|J]. \tag{E.10}$$

The resulting transition rate $\tilde{\mathbb{E}}[q_-(\vec{N})|J] = \tilde{\mathbb{E}}[q_-(J)]$ depends only on J. We use $q_+(J)$ and $\tilde{\mathbb{E}}[q_-(J)]$ to compute the approximation $R(J)$ for $\mathbb{P}[J]$. From detailed balance,

$$R(J) = \frac{q_+(J-1)}{\tilde{\mathbb{E}}[q_-(J)]}R(J-1) = \left[\prod_{k=1}^{J}\frac{q_+(k-1)}{\tilde{\mathbb{E}}[q_-(k)]}\right]R(0) \tag{E.11}$$

and $R(0)$ can be obtained from normalization. Substituting (E.9) and (E.11) into (E.6), we finally get the approximation for $\mathbb{P}[\vec{N}]$.

The case $\alpha < 0$ with transition rate (14) can be dealt with in a similar way. Transition rate $q_{i-}(\vec{N})$ depends only on N_i, but transition rate $q_{i+}(\vec{N})$ depends both on N_i and $J = \sum_i N_i$. We introduce the modified transition rate $\tilde{q}_{i+}(\vec{N})$,

$$\tilde{q}_{i+}(\vec{N}) = \left(r_+ + \frac{r(-\alpha)J}{K'}\right)N_i + \mu, \tag{E.12}$$

in which J is a parameter, not a variable, analogous with (E.7). The stationary distribution $Q(\vec{N})$ of the modified stochastic community model is given by (E.8) with dimensionless parameters

$$a = \frac{\mu K'}{r_+K' + r(-\alpha)J}, \quad b = \frac{r_-K'}{r(1-\alpha)} + 1, \quad c = \frac{r_+K' + r(-\alpha)J}{r(1-\alpha)}.$$

The approximation $Q(\vec{N}|J)$ follows as for the case $\alpha > 0$. To obtain an approximation $R(J)$, we consider the transition rates for the total community size J. The transition rate $q_+(\vec{N})$ depends only on $\vec{N}$ through J, but the transition rate $q_-(\vec{N})$ depends also on the Simpson diversity. We use the same trick as for the case $\alpha > 0$, and use the approximate expected transition rate $\tilde{\mathbb{E}}[q_-(\vec{N})|J]$,

$$\tilde{\mathbb{E}}[q_-(\vec{N})|J] = r_-J + \frac{r(1-\alpha)}{K'}S_T\tilde{\mathbb{E}}[N_i^2|J]$$

analogous with (E.10). The remaining computation is the same as for the case $\alpha > 0$.

Fig. 3 compares the species abundance distribution found from the approximation (E.6) with the species abundance distribution found from a stochastic simulation over a sufficiently long time (to reach stationary regime). The correspondence is excellent. Also, we recomputed part of Fig. 5A using stochastic simulations (the original version of Fig. 5A was computed using our approximation). The results were almost identical to the results obtained using the approximation.

An additional verification of the approximation (E.6) consists in studying the limits $\alpha \to 0$ and $\alpha \to 1$, for which we have exact solutions, (E.2) and (E.4). For $\alpha \to 0$, one can verify that $\mathbb{P}[\vec{N}|J] = Q(\vec{N}|J)$, i.e., approximation (E.9) is exact, and that approximation (E.11), although not exact, is very accurate. For $\alpha \to 1$, one can verify that $\mathbb{P}[\vec{N}|J] = Q(\vec{N}|J)$, i.e., approximation (E.9) is exact, and that $\mathbb{P}[J] = R(J)$, i.e., approximation (E.11) is also exact.

Finally, we summarize the computations to evaluate the approximation (E.6):

- Our goal is to compute an approximation for the stationary probability $\mathbb{P}[\vec{N}]$ for a given abundance vector $\vec{N}$, and a given set of parameters. Denote the total community size by $J = \sum_i N_i$.
- As a preliminary, we compute the functions Q_1, $Q_1^{*(S_T-1)}$ and $Q_1^{*S_T}$. We do not have an analytical expression for the convolution products, but they can be evaluated numerically using the fast Fourier transform. Recall that the formulas for Q_1 differ whether $\alpha > 0$ or $\alpha < 0$.
- We compute the approximation $Q(\vec{N}|J)$ for $\mathbb{P}[\vec{N}|J]$. This approximation is given by $\prod_i Q_1(N_i)/Q_1^{*S_T}(J)$, see (E.9).
- Similarly, we compute the approximation $Q(n|k)$ for $\mathbb{P}[n|k]$, i.e., the probability that a species has abundance n given that the community has size k. This approximation is given by $Q_1(n)Q_1^{*(S_T-1)}(k-n)/Q_1^{*S_T}(k)$. We evaluate this formula for all n and k with $n \le k$.
- We compute $\tilde{\mathbb{E}}[n^2|k] = \sum_n n^2 Q(n|k)$ for all k, i.e., the mean of n^2 when n is distributed according to $Q(n|k)$. We evaluate this formula for all k.
- We compute the transition rates $\tilde{\mathbb{E}}[q_-(k)]$ and $q_+(k)$ for all k, using $\tilde{\mathbb{E}}[n^2|k]$. Recall that the formulas for $\tilde{\mathbb{E}}[q_-(k)]$ and $q_+(k)$ differ whether $\alpha > 0$ or $\alpha < 0$.
- We compute the cumulative products $c(m) = \prod_{k=1}^{m} q_+(k-1)/\tilde{\mathbb{E}}[q_-(k)]$. We evaluate this formula for all m.
- We compute the approximation $R(J)$ for $\mathbb{P}[J]$. The formula is given by $R(J) = c(J)/(1+\sum_m c(m))$, see (E.11).
- Finally, the approximation for $\mathbb{P}[\vec{N}]$ is given by $Q(\vec{N}|J)R(J)$, see (E.6).

Appendix F. Computation of community properties

In this appendix we explain how the community properties introduced in Section 3.3 can be computed using the approximate stationary distribution of the community model (Appendix E). The approximation method provides the community size distribution $\mathbb{P}[J]$, and the population size distribution conditioned on community size $\mathbb{P}[N_i|J]$. We express the community properties in terms of these two probability distributions.

To compute the (unconditional) variance $\mathrm{Var}[\boldsymbol{N}_i]$, we use the law of total variance, or the conditional variance formula,

$$\mathrm{Var}[\boldsymbol{N}_i] = \mathbb{E}[\mathrm{Var}[\boldsymbol{N}_i|\boldsymbol{J}]] + \mathrm{Var}[\mathbb{E}[\boldsymbol{N}_i|\boldsymbol{J}]]. \tag{F.1}$$

The notation in the first term of the right-hand side should be read as follows: first, the variance of $\boldsymbol{N}_i$ is taken conditional on $\boldsymbol{J} = J$ for all J; the result is then considered as a function of the random variable $\boldsymbol{J}$, of which the expectation is taken. The notation in the second term of the right-hand side is defined similarly. From (F.1),

$$\mathrm{Var}[\boldsymbol{N}_i] = \mathbb{E}[\mathrm{Var}[\boldsymbol{N}_i|\boldsymbol{J}]] + \frac{1}{S_T^2}\mathrm{Var}[\boldsymbol{J}]$$

a formula in terms of distributions $\mathbb{P}[J]$ and $\mathbb{P}[N_i|J]$.

For the Simpson diversity index, we use the definition

$$\boldsymbol{D} = 1 - \sum_i \frac{\boldsymbol{N}_i(\boldsymbol{N}_i - 1)}{\boldsymbol{J}(\boldsymbol{J} - 1)}.$$

Taking the expectation in two steps, we get

$$\mathbb{E}[\boldsymbol{D}] = 1 - S_{\mathrm{T}}\mathbb{E}\left[\frac{\mathbb{E}[\boldsymbol{N}_i(\boldsymbol{N}_i - 1)|\boldsymbol{J}]}{\boldsymbol{J}(\boldsymbol{J} - 1)}\right],$$

where we used a notation similar to (F.1). The inner expectation can be computed using the distribution $\mathbb{P}[N_i|J]$; the outer expectation can be computed using the distribution $\mathbb{P}[J]$.

The different steps for the computation of $\mathrm{Var}[\boldsymbol{N}_i]$ and $\mathbb{E}[\boldsymbol{D}]$ are (see end of Appendix E for more details):

- Compute the functions Q_1, $Q_1^{*(S_{\mathrm{T}}-1)}$ and $Q_1^{*S_{\mathrm{T}}}$.
- Compute the approximation $Q(N_i|J)$ for $\mathbb{P}[N_i|J]$, for all N_i and J with $N_i \le J$.
- Compute $\tilde{\mathbb{E}}[N_i^2|J]$ with respect to the distribution $Q(N_i|J)$, for all J.
- Compute the approximations for $\mathrm{Var}[N_i|J]$ and $\mathbb{E}[N_i(N_i-1)|J]$, for all J.
- Compute the transition rates $\tilde{\mathbb{E}}[q_-(J)]$ and $q_+(J)$, for all J.
- Compute the cumulative products $c(J)$, for all J.
- Compute the approximation $R(J)$ for $\mathbb{P}[J]$, for all J.
- Compute the $\mathbb{E}[\mathrm{Var}[N_i|J]]$, $\mathrm{Var}[J]$ and $\mathbb{E}[\mathbb{E}[N_i(N_i-1)|J]/J(J-1)]$ with respect to the distribution $R(J)$.
- Compute the approximations for $\mathrm{Var}[N_i]$ and $\mathbb{E}[D]$.

Appendix G. Supplementary material

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.jtbi.2010.10.006.

References

Adler, P.B., HilleRisLambers, J., Levine, J.M., 2007. A niche for neutrality. Ecol. Lett. 10, 95–104.

Allouche, O., Kadmon, R., 2009. A general framework for neutral models of community dynamics. Ecol. Lett. 12, 1287–1297.

Alonso, D., Etienne, R.S., McKane, A.J., 2006. The merits of neutral theory. Trends Ecol. Evol. 21, 451–457.

Bell, G., 2000. The distribution of abundance in neutral communities. Am. Nat. 155, 606–617.

Bolker, B.M., Pacala, S.W., Levin, S.A., 2000. Moment methods for stochastic processes in continuous space and time. In: Dieckmann, U., Law, R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions: Simplifying Spatial Complexity. Cambridge University Press, Cambridge, pp. 388–411.

Caswell, H., 1976. Community structure: a neutral model analysis. Ecol. Monogr. 46, 327–354.

Chase, J.M., Leibold, M.A., 2004. Ecological Niches: Linking Classical and Contemporary Approaches. University of Chicago Press, Chicago.

Chesson, P., 2000. Mechanisms of maintenance of species diversity. Annu. Rev. Ecol. Syst. 31, 343–366.

Etienne, R.S., Alonso, D., McKane, A.J., 2007. The zero-sum assumption in neutral biodiversity theory. J. Theor. Biol. 248, 522–536.

Fuentes, M., 2004. Slight differences among individuals and the unified neutral theory of biodiversity. Theor. Popul. Biol. 66, 199–203.

Gardiner, C.W., 2004. Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences. Springer, Berlin.

Gravel, D., Canham, C.D., Beaudet, M., Messier, C., 2006. Reconciling niche and neutrality: the continuum hypothesis. Ecol. Lett. 9, 399–406.

Haegeman, B., Etienne, R.S., 2008. Relaxing the zero-sum assumption in neutral biodiversity theory. J. Theor. Biol. 252, 288–294.

Holyoak, M., Loreau, M., 2006. Reconciling empirical ecology with neutral community models. Ecology 87, 1370–1377.

Hubbell, S.P., 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton.

Ives, A.R., Gross, R., Klug, J.L., 1999. Stability and variability in competitive communities. Science 286, 542–544.

Kadmon, R., Allouche, O., 2007. Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. Am. Nat. 170, 443–454.

Lande, R., Engen, S., Saether, B.-E., 2003. Stochastic Population Dynamics in Ecology and Conservation. Oxford University Press, Oxford.

Leibold, M.A., McPeek, M.A., 2006. Coexistence of the niche and neutral perspectives in community ecology. Ecology 87, 1399–1410.

Loreau, M., 2010. From Populations to Ecosystems: Theoretical Foundations for a New Ecological Synthesis. Princeton University Press, Princeton.

Loreau, M., de Mazancourt, C., 2008. Species synchrony and its drivers: neutral and nonneutral community dynamics in fluctuating environments. Am. Nat. 172, E48–E66.

MacArthur, R.H., 1972. Geographical Ecology. Princeton University Press, Princeton.

MacArthur, R.H., Levins, R., 1967. The limiting similarity, convergence, and divergence of coexisting species. Am. Nat. 101, 377–385.

Matis, J.H., Kiffe, T.R., 2000. Stochastic Population Models. Springer, New York.

Mouquet, N., Loreau, M., 2003. Community patterns in source–sink metacommunities. Am. Nat. 162, 544–557.

Nåsell, I., 2003. Moment closure and the stochastic logistic model. Theor. Popul. Biol. 63, 159–168.

Pielou, E.C., 1977. Mathematical Ecology. Wiley, New York.

Renshaw, E., 1991. Modelling Biological Populations in Space and Time. Cambridge University Press, Cambridge.

Ringel, M.S., Hu, H.H., Anderson, G., 1996. The stability and persistence of mutualisms embedded in community interactions. Theor. Popul. Biol. 50, 281–297.

Schwilk, D.W., Ackerly, D.D., 2005. Limiting similarity and functional diversity along environmental gradients. Ecol. Lett. 8, 272–281.

Tilman, D., 2004. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proc. Natl. Acad. Sci. U.S.A. 101, 10854–10861.

Van Kampen, N.G., 1997. Stochastic Processes in Physics and Chemistry. Elsevier, Amsterdam.

Vanpeteghem, D., Haegeman, B., 2010. An analytical approach to spatio-temporal dynamics of neutral community models. J. Math. Biol. 61, 323–357.

Vanpeteghem, D., Zemb, O., Haegeman, B., 2008. Dynamics of neutral biodiversity. Math. Biosci. 212, 88–98.

Volkov, I., Banavar, J.R., He, F., Hubbell, S.P., Maritan, A., 2005. Density dependence explains tree species abundance and diversity in tropical forests. Nature 438, 658–661.

Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2009. Inferring species interactions in tropical forests. Proc. Natl. Acad. Sci. U.S.A. 106, 13854–13859.

Wolin, C.L., Lawlor, L.R., 1984. Models of facultative mutualism: density effects. Am. Nat. 124, 843–862.

Zhang, D.-Y., Lin, K., 1997. The effects of competitive asymmetry on the rate of competitive displacement: How robust is Hubbell's community drift model? J. Theor. Biol. 188, 361–367.

Zhou, S.-R., Zhang, D.-Y., 2008. A nearly neutral model of biodiversity. Ecology 89, 248–258.

Zillio, T., Condit, R., 2007. The impact of neutrality, niche differentiation and species input on diversity and abundance distributions. Oikos 116, 931–940.