A sampling theory for asymmetric communities

Andrew E. Noble [a,*], Nico M. Temme [b], William F. Fagan [a], Timothy H. Keitt [c]

[a] *Department of Biology, University of Maryland, College Park, MD 20742, USA*
[b] *Center for Mathematics and Computer Science, Science Park 123, 1098 XG Amsterdam, The Netherlands*
[c] *Section of Integrative Biology, University of Texas at Austin, 1 University Station C0930, Austin, TX 78712, USA*

ABSTRACT

We introduce the first analytical model of asymmetric community dynamics to yield Hubbell's neutral theory in the limit of functional equivalence among all species. Our focus centers on an asymmetric extension of Hubbell's local community dynamics, while an analogous extension of Hubbell's meta-community dynamics is deferred to an appendix. We find that mass-effects may facilitate coexistence in asymmetric local communities and generate unimodal species abundance distributions indistinguishable from those of symmetric communities. Multiple modes, however, only arise from asymmetric processes and provide a strong indication of non-neutral dynamics. Although the exact stationary distributions of fully asymmetric communities must be calculated numerically, we derive approximate sampling distributions for the general case and for nearly neutral communities where symmetry is broken by a single species distinct from all others in ecological fitness and dispersal ability. In the latter case, our approximate distributions are fully normalized, and novel asymptotic expansions of the required hypergeometric functions are provided to make evaluations tractable for large communities. Employing these results in a Bayesian analysis may provide a novel statistical test to assess the consistency of species abundance data with the neutral hypothesis.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The ecological symmetry of trophically similar species forms the central assumption in Hubbell's unified neutral theory of biodiversity and biogeography (Hubbell, 2001). In the absence of stable coexistence mechanisms, local communities evolve under zero-sum ecological drift—a stochastic process of density-dependent birth, death, and migration that maintains a fixed community size (Hubbell, 2001). Despite a homogeneous environment, migration inhibits the dominance of any single species and fosters high levels of diversity. The symmetry assumption has allowed for considerable analytical developments that draw on the mathematics of neutral population genetics (Fisher, 1930; Wright, 1931) to derive exact predictions for emergent, macro-ecological patterns (Chave, 2004; Etienne and Alonso, 2007; McKane et al., 2000, 2004; Vallade and Houchmandzadeh, 2003; Volkov et al., 2003, 2005, 2007; Etienne and Olff, 2004; Pigolotti et al., 2004; He, 2005; Hu et al., 2007; Babak and He, 2008, 2009). Among the most significant contributions are calculations of multivariate sampling distributions that relate local abundances to those in the regional meta-community (Alonso and McKane, 2004; Etienne and Alonso, 2005;

Etienne, 2005, 2007). Hubbell (2001) first emphasized the utility of sampling theories for testing neutral theory against observed species abundance distributions (SADs). Since then, Etienne and Olff (2004, 2005) have incorporated sampling distributions as conditional likelihoods in Bayesian analyses (Etienne, 2007, 2009). Recent work has shown that the sampling distributions of neutral theory remain invariant when the restriction of zero-sum dynamics is lifted (Etienne et al., 2007; Haegeman and Etienne, 2008; Conlisk et al., 2010) and when the assumption of strict symmetry is relaxed to a requirement of ecological equivalence (Etienne et al., 2007; Haegeman and Etienne, 2008; Allouche and Kadmon, 2009a,b; Lin et al., 2009).

The success of neutral theory in fitting empirical patterns of biodiversity (Hubbell, 2001; Volkov et al., 2003, 2005; He, 2005; Chave et al., 2006) has generated a heated debate among ecologists, as there is strong evidence for species asymmetry in the field (Harper, 1977; Goldberg and Barton, 1992; Chase and Leibold, 2003; Wootton, 2009; Levine and HilleRisLambers, 2009). Echoing previous work on the difficulty of resolving competitive dynamics from the essentially static observations of co-occurrence data (Hastings, 1987), recent studies indicate that interspecific tradeoffs may generate unimodal SADs indistinguishable from the expectations of neutral theory (Chave et al., 2002; Mouquet and Loreau, 2003; Chase, 2005; He, 2005; Purves and Pacala, 2005; Walker, 2007; Doncaster, 2009). These results underscore the compatibility of asymmetries and coexistence. The pioneering work of Hutchinson (1951), has inspired

* Corresponding author. Tel.: +1 607 351 6341.
 E-mail addresses: andrewenoble@gmail.com (A.E. Noble),
Nico.Temme@cwi.nl (N.M. Temme), bfagan@umd.edu (W.F. Fagan),
tkeitt@gmail.com (T.H. Keitt).

a large literature on asymmetries in dispersal ability that permit the coexistence of "fugitive species" with dominant competitors. In particular, Shmida and Wilson (1985) extended the work of Brown and Kodric-Brown (1977) by introducing the paradigm of "mass-effects", where immigration facilitates the establishment of species in sites where they would otherwise be competitively excluded. Numerous attempts have been made to reconcile such determi-nistic approaches to the coexistence of asymmetric species with the stochastic model of ecological drift in symmetric neutral theory (Zhang and Lin, 1997; Tilman, 2004; Chase, 2005; Alonso et al., 2006, 2008; Gravel et al., 2006; Pueyo et al., 2007; Walker, 2007; Ernest et al., 2008; Zhou and Zhang, 2008). Many of these attempts build on insights from the concluding chapter of Hubbell's (2001) book.

Nevertheless, the need remains for a fully asymmetric, analytical, sampling theory that contains Hubbell's model as a limiting case (Alonso et al., 2006). In this article, we develop such a theory for local, dispersal-limited communities in the main text and defer an analogous treatment of metacommunities to Appendix A. Hubbell's assumption of zero-sum dynamics is preserved, but the requirement of per capita ecological equivalence among all species is eliminated. Asymmetries are introduced by allowing for the variations in ecological fitness and dispersal ability that may arise in a hetero-geneous environment (Leibold et al., 2004; Holyoak et al., 2005). Our work expands on the numerical simulations of Zhou and Zhang (2008), where variations in ecological fitness alone were considered. Coexistence emerges from mass-effects as well as ecological equiva-lence, and both mechanisms generate unimodal SADs that may be indistinguishable. For local communities and metacommunities, we derive approximate sampling distributions for both the general case and the nearly neutral case, where symmetry is broken by a single species unique in ecological function. These approximations yield the sampling distributions of Hubbell's neutral model in the limit of functional equivalence among all species.

2. A general sampling theory for local communities

For a local community of J_L individuals and S possible species, we model community dynamics as a stochastic process, $\vec{N}(\tau)$, over the labelled community abundance vectors $\vec{n} = (n_1, \ldots, n_S)$. Consistent with zero-sum dynamics, we require all accessible states to contain J_L total individuals: $\sum_{i=1}^{S} n_i = J_L$ and $0 \le n_i \le J_L$. The number of accessible states is $A = \sum_{n_1=0}^{J_L} \cdots \sum_{n_S=0}^{J_L} \delta(J_L - n_1 - \cdots - n_S)$.

Allowed transitions first remove an individual from species i and then add an individual to species j. Removals are due to death or emigration and occur with the density-dependent probability n_i/J_L. Additions are due either to an immigration event, with probability m_j, or a birth event, with probability $1 - m_j$. We will refer to the m_j as dispersal abilities. If immigration occurs, we assume that meta-community relative abundance, x_j, determines the proportional representation of species j in the propagule rain and that the probability of establishment is weighted by ecological fitness, w_j, where high values correspond to a local competitive advantage or a superior adaptation to the local environment. Therefore, species j recruits with probability

$$\frac{w_j x_j}{\sum_{k=1}^{S} w_k x_k}, \qquad (1)$$

where $x_j \in (0,1)$, $w_j \in (0,\infty)$, and $\sum_{k=1}^{S} x_k = 1$. If immigration does not occur, we assume that local relative abundance, n_j/J_L, governs propagule rain composition such that species j recruits with probability

$$\frac{w_j n_j}{\sum_{k=1}^{S} w_k n_k - w_i}. \qquad (2)$$

In numerical simulations of an asymmetric community, Zhou and Zhang (2008) employed a similar probability for recruitment in the absence of immigration. Here, a factor of w_i is subtracted in the denominator because species i loses an individual prior to the birth event for species j. An analogous subtraction is absent from Eq. (1) because we assume an infinite metacommunity where the x_j are invariant to fluctuations in the finite, local community populations.

In sum, the nonzero transition probabilities are stationary and given by

$$
\begin{aligned}
T_{ij\vec{n}} &= \lim_{\Delta\tau \to 0} \frac{Pr\{\vec{N}(\tau+\Delta\tau) = \vec{n} - \vec{e}_i + \vec{e}_j | \vec{N}(\tau) = \vec{n}\}}{\Delta\tau} \\
&= \frac{n_i}{J_L}\left((1-m_j)\frac{w_j n_j}{\sum_{k=1}^{S} w_k n_k - w_i} + m_j \frac{w_j x_j}{\sum_{k=1}^{S} w_k x_k} \right),
\end{aligned} \qquad (3)
$$

where $\vec{e}_i$ is an S-dimensional unit vector along the ith-direction, the w_k must be sufficiently large such that $\sum_{k=1}^{S} w_k n_k - w_i > 0$, and the time, τ, is dimensionless with a scale set by the overall transition rate. The probability of state occupancy, $P_{\vec{n}}$, evolves according to the master equation

$$\frac{dP_{\vec{n}}}{d\tau} = \sum_{i=1}^{S}\sum_{j=1, j\ne i}^{S} (T_{ij\vec{n}+\vec{e}_i-\vec{e}_j}P_{\vec{n}+\vec{e}_i-\vec{e}_j} - T_{ji\vec{n}}P_{\vec{n}})\Theta_{ij}, \qquad (4)$$

where

$$\Theta_{ij} = \Theta(J_L - (n_i+1))\Theta(n_j-1), \qquad (5)$$

and we define the step-function $\Theta(x)$ to be zero for $x < 0$ and one otherwise. Eq. (4) can be recast in terms of a transition probability matrix W

$$\frac{dP_a}{d\tau} = \sum_{b=1}^{A} P_b W_{ba}, \qquad (6)$$

where $a, b \in (1, \ldots, A)$ enumerate accessible states with components $(a_1, \ldots, a_S)$, $(b_1, \ldots, b_S)$. The left eigenvector of W with zero eigenvalue yields the stationary distribution for community composition, $P_a^* \equiv \lim_{\tau\to\infty} P_a(\tau)$. Marginal distributions yield the equilibrium abundance probabilities for each species i

$$P_n^{(i)*} = \sum_{a=1}^{A} \delta_{a_i, n} P_a^*. \qquad (7)$$

From here, we calculate the stationary SAD by following the general treatment of asymmetric communities in Alonso et al. (2008):

$$S_n^* = \sum_{i=1}^{S} P_n^{(i)*}. \qquad (8)$$

The expected species richness is

$$S^* = \sum_{n=1}^{S} S_n^* < S. \qquad (9)$$

Given that the local community, with abundances n_i, is defined as a sample of the metacommunity, with relative abundances x_i, we have established the framework for a general sampling theory of local communities.

This sampling theory incorporates aspects of the mass-effects paradigm (Brown and Kodric-Brown, 1977; Shmida and Wilson, 1985; Holt, 1993; Leibold et al., 2004; Holyoak et al., 2005). Local asymmetries in ecological fitness imply environmental heteroge-neity across the metacommunity such that competitive ability peaks in the local communities where biotic and abiotic factors most closely match niche requirements (Tilman, 1982; Leibold, 1998; Chase and Leibold, 2003). Where species experience a competitive disadvantage, the mass-effects of immigration allow for persistence. Indeed, the master equation given by Eq. (4), when applied to open communities where $m_j > 0$ for all j, admits no
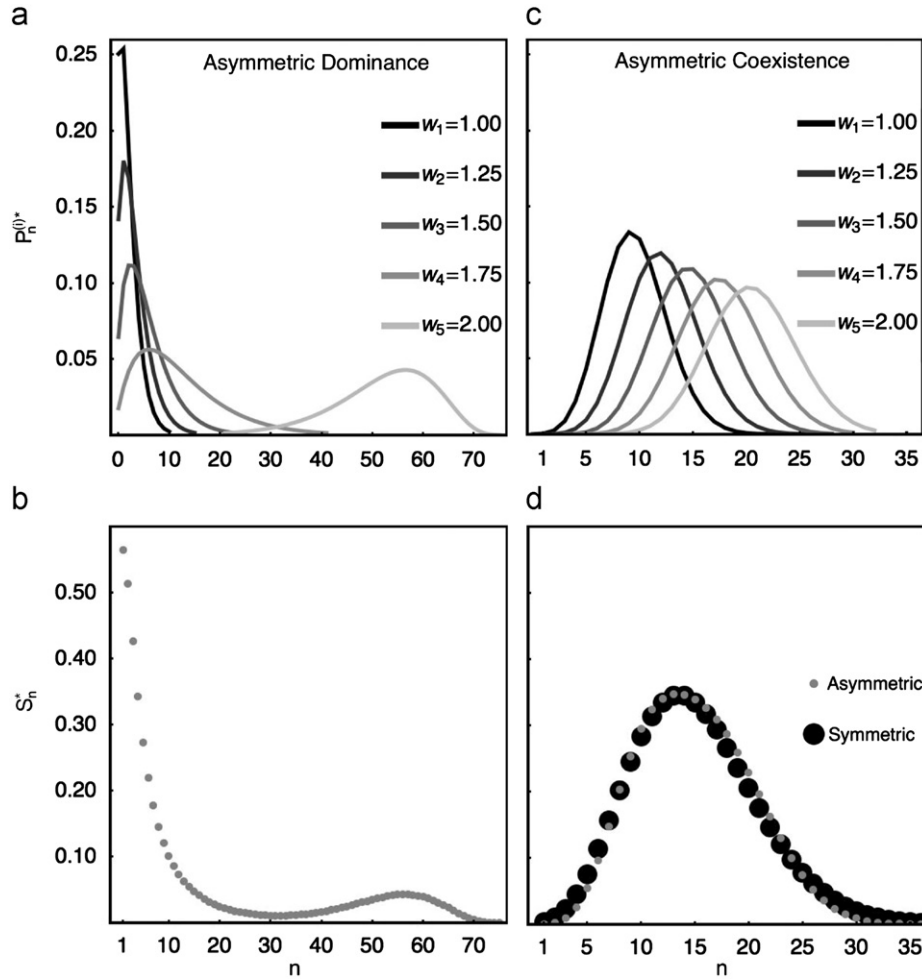
Fig. 1. Equilibrium abundance probabilities and corresponding SADs for an asymmetric community of $J_L = 75$ individuals and $S = 5$ species. Non-neutral dynamics generate local deviations from the relative metacommunity abundance, $x_i = 0.20$, common to all species. (a) Dominance by the species with highest ecological fitness given competitive asymmetries in a community of uniform dispersal abilities (the $m_i = 0.10$ for all i). (b) The resulting bimodal SAD provides a strong indicator of non-neutral dynamics. (c) Coexistence arising from mass-effects (the $m_i = 0.90$ for all i). (d) The resulting unimodal SAD closely resembles the SAD for a symmetric community of $J_L = 75$ individuals and $S = 5$ species where the $m_i = 0.35$ and the $x_i = 0.20$.

absorbing states and ensures that every species has a nonzero probability of being present under equilibrium conditions. By contrast, when Eq. (4) is applied to closed communities where $m_j = 0$ for all j, the eventual dominance of a single species is guaranteed.

Mass-effects allow for a soft breaking of the symmetry of neutral theory and provide a mechanism for multi-species coexistence. In Fig. 1, we present numerical results for the marginal equilibrium distributions of an asymmetric local community subsidized by a potentially neutral metacommunity, where the five species share a common relative abundance, $x_j = 0.2$. Although a single species may dominate due to a locally superior competitive ability (see Fig. 1a), multi-species coexistence may arise, despite significant competitive asymmetries, due to high levels of immigration that tend to align local relative abundances with those in the metacommunity (see Fig. 1c). Despite the underlying asymmetric process, coexistence via mass-effects generates unimodal SADs that, given sampling errors in field data, may be indistinguishable from SADs due to neutral dynamics, as shown in Fig. 1d. This reinforces previous conclusions that the static, aggregate data in unimodal SADs cannot resolve the individual-level rules of engagement governing the origin and maintenance of biodiversity (Chave et al., 2002; Mouquet and Loreau, 2003; Purves and Pacala, 2005; He, 2005; Chase, 2005; Walker, 2007; Doncaster, 2009).

However, SADs with multiple modes are not uncommon in nature (Dornelas and Connolly, 2008; Gray et al., 2005) and provide a strong indicator of non-neutral dynamics (Alonso et al., 2008). Fig. 1b presents a bimodal SAD for an asymmetric local community with low levels of immigration.

Each plot in Fig. 1 displays results for a relatively small community of $J_L = 75$ individuals and $S = 5$ possible species. Sparse matrix methods were used to calculate the left eigenvector with zero eigenvalue for transition matrices of rank ~ 1.5M. Obtaining stationary distributions for larger, more realistic communities poses a formidable numerical challenge. This motivates a search for analytically tractable approximations to sampling distributions of the general theory.

3. An approximation to the sampling distribution

The distribution $P_{\vec{n}}^*$ is stationary under Eq. (4) if it satisfies the condition of detailed balance

$$T_{ij\vec{n} + \vec{e}_i - \vec{e}_j} P_{\vec{n} + \vec{e}_i - \vec{e}_j}^* = T_{ji\vec{n}} P_{\vec{n}}^*, \qquad (10)$$

for all i and j such that $i \neq j$ and $\Theta_{ij} \neq 0$. For general (g) large-J_L communities where $S, w_k \ll \sum_{l=1}^{S} w_l n_l$ for all k, we will show that

detailed balance is approximately satisfied by

$$P_{\vec{n}}^{g*} = Z_g^{-1} \binom{J_L}{n_1, \ldots, n_S} \prod_{k=1}^{S} w_k^{n_k}(1-m_k)^{n_k}(\phi_{k\vec{n}} x_k)_{n_k}, \tag{11}$$

where

$$\phi_{k\vec{n}} = I_k \frac{(\sum_{l=1}^{S} w_l n_l - w_k)/(J_L - 1)}{\sum_{l=1}^{S} w_l x_l}, \tag{12}$$

where $(a)_n = \Gamma(a+n)/\Gamma(a)$ is the Pochhammer symbol, Z_g^{-1} is a normalization constant, and $I_k = m_k(J_L-1)/(1-m_k)$ is a generalization of the "fundamental dispersal number" (Etienne and Alonso, 2005). From the definition of $T_{ij\vec{n}}$ in Eq. (3), we have

$$\frac{T_{ij\vec{n}+\vec{e}_i-\vec{e}_j}}{T_{ji\vec{n}}} = \frac{n_i+1}{n_j}\frac{w_j}{w_i}\frac{1-m_j}{1-m_i}\frac{n_j-1+\phi_{ji\vec{n}} x_j}{n_i+\phi_{i\vec{n}} x_i\left(1+\frac{w_i-w_j}{\sum_{l=1}^{S} w_l n_l - w_i}\right)}, \tag{13}$$

and assuming the form of $P_{\vec{n}}^{g*}$ in Eq. (11), we find

$$\frac{P_{\vec{n}}^{g*}}{P_{\vec{n}+\vec{e}_i-\vec{e}_j}^{g*}} = \frac{n_i+1}{n_j}\frac{w_j}{w_i}\frac{1-m_j}{1-m_i}\prod_{k=1}^{S}\frac{(\phi_{k\vec{n}} x_k)_{n_k}}{\left(\phi_{k\vec{n}} x_k\left(1+\frac{w_i-w_j}{\sum_{l=1}^{S} w_l n_l - w_k}\right)\right)_{n_k+\delta_{ik}-\delta_{jk}}}. \tag{14}$$

Now, for large-J_L communities where $w_k \ll \sum_{l=1}^{S} w_l n_l$ for all k, the ratio $\varepsilon_{ijk} \equiv (w_i-w_j)/(\sum_{l=1}^{S} w_l n_l - w_k)$ is a small number. Given $(a(1+\varepsilon))_n \sim (a)_n + \mathcal{O}(\varepsilon)$, we expand the right-hand side of Eq. (14) to obtain

$$\frac{P_{\vec{n}}^{g*}}{P_{\vec{n}+\vec{e}_i-\vec{e}_j}^{g*}} = \frac{T_{ij\vec{n}+\vec{e}_i-\vec{e}_j}}{T_{ji\vec{n}}} + \sum_{k=1}^{S} \mathcal{O}(\varepsilon_{ijk}), \tag{15}$$

which validates our assertion that Eq. (11) is an approximate sampling distribution of the general theory when $S \ll \sum_{l=1}^{S} w_l n_l$. For communities of species that are symmetric (s) in ecological fitness but asymmetric in dispersal ability, Eq. (11) reduces to an exact sampling distribution

$$P_{\vec{n}}^{s*} = Z_s^{-1} \binom{J_L}{n_1, \ldots, n_S} \prod_{k=1}^{S} (1-m_k)^{n_k}(I_k x_k)_{n_k}, \tag{16}$$

that satisfies detailed balance without approximation. Analogous distributions for general and fitness-symmetric metacommunities are provided in Appendix A. However, in all of these results, the normalization constants must be calculated numerically. This limits the utility of our sampling distributions in statistical analyses. Can we find a non-neutral scenario that admits an approximate sampling distribution with an analytical expression for the normalization?

4. Sampling nearly neutral communities

As the species abundance vector evolves under Eq. (4), consider the dynamics of marginal abundance probabilities for a single focal species that deviates in ecological function from the surrounding, otherwise symmetric, community. In particular, let the first element of $\vec{N}(\tau)$ be the marginal process, $N(\tau)$, over states $n \in (0, \ldots, J_L)$, for the abundance of an asymmetric focal species with dispersal ability m, ecological fitness w, and relative metacommunity abundance x. If all other species share a common dispersal ability m_o and ecological fitness w_o, then the focal species gains an individual with probability

$$g_n \equiv \sum_{i=2}^{S} T_{i1(n,n_2,\ldots,n_S)}$$
$$= \frac{J_L-n}{J_L}\left((1-m)\frac{wn}{wn+w_o(J_L-n-1)} + m\frac{wx}{wx+w_o(1-x)}\right), \tag{17}$$

and loses an individual with probability

$$r_n \equiv \sum_{i=2}^{S} T_{1i(n,n_2,\ldots,n_S)}$$
$$= \frac{n}{J_L}\left((1-m_o)\frac{w_o(J_L-n)}{w(n-1)+w_o(J_L-n)} + m_o\frac{w_o(1-x)}{wx+w_o(1-x)}\right), \tag{18}$$

where we have used $\sum_{k=1}^{S} x_k = 1$. These marginal transition probabilities do not depend separately on w and w_o, but only on their ratio. Without loss of generality, we redefine $w \equiv w/w_o$ to be the focal species' local *advantage* in ecological fitness. Eqs. (17) and (18), which are independent of the abundances $(n_2,\ldots,n_S)$, suggest a univariate birth–death process for the marginal dynamics of the asymmetric species governed by the master equation

$$\frac{dP_n}{d\tau} = g_{n-1}\Theta(n-1)P_{n-1} + r_{n+1}\Theta(J_L-(n+1))P_{n+1}$$
$$- (g_n\Theta(J_L-(n+1)) + r_n\Theta(n-1))P_n, \tag{19}$$

and we formally derive this result from Eq. (4) in Appendix B. Given the well-known stationary distribution of Eq. (19)

$$P_n^* = P_0^* \prod_{i=0}^{n-1} \frac{g_i}{r_{i+1}}, \tag{20}$$

we find an exact result for the stationary abundance probabilities of the focal species in a nearly neutral (nn) community

$$P_n^{nn*} = Z\binom{J_L}{n}\eta^n \frac{B(\lambda+n, \xi-n)}{B(\lambda, \xi)}, \tag{21}$$

where $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta-function

$$Z^{-1} = {}_2F_1(-J_L, \lambda; 1-\xi; \eta), \tag{22}$$

and

$$\eta = w\frac{1-m+x(w-1)}{1-m_o w + x(w-1)},$$

$$\lambda = \frac{(J_L-1)mx}{1-m+x(w-1)},$$

$$\xi = 1 + \frac{(J_L-1)(1-xwm_o+x(w-1))}{1-wm_o+x(w-1)}. \tag{23}$$

For the asymmetric focal species, this is an exact result of the general model, Eq. (4), that holds for nearly neutral local communities with any number of additional species. Eq. (21) may be classified broadly as a generalized hypergeometric distribution or more specifically as an exponentially weighted Pólya distribution (Kemp, 1968; Johnson et al., 1992).

In the absence of dispersal limitation, Eq. (21) becomes

$$\lim_{m,m_o \to 1} P_n^{nn*} = \left(\frac{1}{1+x(w-1)}\right)^{J_L}\binom{J_L}{n}(wx)^n(1-x)^{J_L-n}, \tag{24}$$

where the identity $B(a,b)B(a+b,1-b) = \pi/(a\sin(\pi b))$ has been used. This is a weighted binomial distribution with expected abundance $wxJ_L/(1+x(w-1))$ and variance $wx(1-x)J_L/(1+x(w-1))^2$. In the neutral, or symmetric, limit where $w=1$, Eq. (24) reduces to a binomial sampling of the metacommunity, *sensu* Etienne and Alonso (2005).

In the presence of dispersal limitation, we evaluate $\sum_{n=1}^{J_L} nP_n^{nn*}$ to obtain the expected abundance

$$E[N^*] = \eta\frac{\partial}{\partial \eta}\log Z = \frac{J_L\lambda\eta}{\xi-1}\frac{{}_2F_1(1-J_L, 1+\lambda; 2-\xi; \eta)}{{}_2F_1(-J_L, \lambda; 1-\xi; \eta)}, \tag{25}$$

where $N^* \equiv \lim_{\tau \to \infty} N(\tau)$. The variance of the stationary distribution is given by

$$\text{Var}[N^*] = \eta \frac{\partial}{\partial \eta} \eta \frac{\partial}{\partial \eta} \log Z = E[N^{*2}] - E[N^*]^2, \qquad (26)$$

and we evaluate $\sum_{n=1}^{J_L} n^2 P_n^{\text{nn*}}$ to obtain

$$E[N^{*2}] = \frac{J_L \lambda \eta}{\xi - 1} \frac{{}_3F_2(1 - J_L, 1 + \lambda, 2; 2 - \xi, 1; \eta)}{{}_2F_1(-J_L, \lambda; 1 - \xi; \eta)}. \qquad (27)$$

In Eqs. (25) and (26), the normalization of Eq. (22) generates central moments for the abundance distribution and plays a role analogous to the grand partition function of statistical physics. Recent studies have demonstrated the utility of partition functions in extensions of Hubbell's neutral theory (O'Dwyer et al., 2009; O'Dwyer and Green, 2010).

For large-J_L communities, evaluation of the hypergeometric functions in Eqs. (21), (25), and (27) is computationally expensive. To remove this barrier, one of us (N.M.T.) has derived novel asymptotic expansions (see Appendix C). We use these expansions to plot the stationary abundance probabilities for $J_L = 1M$. In Fig. 2a, small local advantages in ecological fitness generate substantial increases in expected abundance over the neutral prediction. Hubbell (2001) found evidence for these discrepancies in Manu forest data and referred to them as "ecological dominance deviations". Hubbell (2001) also anticipated that dispersal effects would mitigate advantages in ecological fitness. The right panel of Fig. 2 demonstrates, once again, that enhanced mass-effects due to increased dispersal ability may inhibit the dominance of a locally superior competitor by compelling relative local abundance to align with relative metacommunity abundance.

An approximation to the multivariate sampling distribution of nearly neutral local communities is constructed in Appendix B

$$P_{\vec{n}}^{\text{nn*}} = Z \binom{J_L}{n, n_2, \ldots, n_S} \eta^n \frac{B(\lambda + n, \xi - n)}{B(\lambda, \xi)} \frac{1}{((1-x)\phi_{on})_{J_L - n}} \prod_{i=2}^{S} (\phi_{on} x_i)_{n_i}, \qquad (28)$$

where

$$\phi_{on} = I_o \frac{1 + n(w-1)/(J_L - 1)}{1 + x(w-1)}. \qquad (29)$$

A related approximation for the sampling distribution of nearly neutral metacommunities is derived in Appendix A. In the absence of dispersal limitation, Eq. (28) becomes

$$\lim_{m, m_o \to 1} P_{\vec{n}}^{\text{nn*}} = \left(\frac{1}{1 + x(w-1)}\right)^{J_L} \binom{J_L}{n, n_2, \ldots, n_S} (wx)^n \prod_{i=2}^{S} x_i^{n_i}, \qquad (30)$$

where we have used $(a)_n \sim a^n + \mathcal{O}(a^{n-1})$ for large a. Finally, in the symmetric limit, Eq. (30) reduces to a simple multinomial sampling of the metacommunity, as expected.

To illustrate the impacts of an asymmetric species on the diversity of an otherwise symmetric local community, Fig. 3 plots Shannon's Index of diversity

$$H = -\frac{E[N^*]}{J_L} \log \frac{E[N^*]}{J_L} - \sum_{i=2}^{S} \frac{E[N_i^*]}{J_L} \log \frac{E[N_i^*]}{J_L}, \qquad (31)$$

for various values of the ecological fitness advantage, w, and dispersal ability, m, in a nearly neutral community of $S = 5$ species and $J_L = 75$ individuals. All five species share a common relative metacommunity abundance, $x = x_i = 0.2$, so given the exact result for $E[N^*]$ in Eq. (25), we know immediately that $E[N_i^*] = (J_L - E[N^*])/(S-1)$ for the remaining symmetric species. Note that H is maximized where all abundances are equivalent, such that $E[N^*]/J_L = E[N_i^*]/J_L = x_i$. As can be seen from the next section, this relation holds in the neutral limit where $w = 1$ and $m = m_o = 0.1$, but small asymmetries in dispersal ability have a negligible impact on diversity when all species are symmetric in ecological fitness. Therefore, each curve in Fig. 3 peaks near $w = 1$ at approximately the same value of H. Away from $w = 1$, the declines in diversity are regulated by mass-effects, with more gradual declines at higher values of m.

5. Recovering the sampling distribution of neutral theory

In a perfectly symmetric local community, the stochastic dynamics for each species differ solely due to variations in relative metacommunity abundances, the x_i. In particular, if $m_j = m$ and $w_j = 1$ for all j in Eq. (3), we recover the multivariate transition probabilities for a neutral sampling theory of local communities, as suggested on p. 287 of Hubbell's book (Hubbell, 2001). Similarly, in the symmetric limit of Eq. (19) where $m = m_o$ and $w = 1$, we recover the marginal dynamics for neutral (n) local communities with
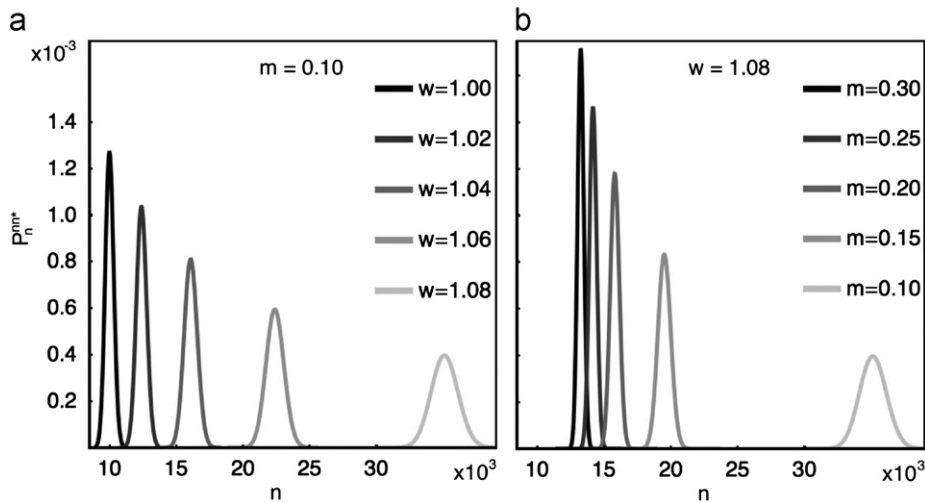

Fig. 2. Novel asymptotic expansions of hypergeometric functions have been used to plot marginal equilibrium abundance probabilities for the asymmetric focal species, with relative metacommunity abundance $x = 0.01$, in a nearly neutral local community of $J_L = 1M$ individuals. (a) Dominance with rising advantage in ecological fitness, as indicated for each curve. Here, all species are symmetric in their dispersal ability ($m = m_o = 0.10$). (b) Dispersal mitigates the advantage in ecological fitness ($w = 1.08$) of the asymmetric focal species. All other species share a common dispersal ability of $m_o = 0.10$.
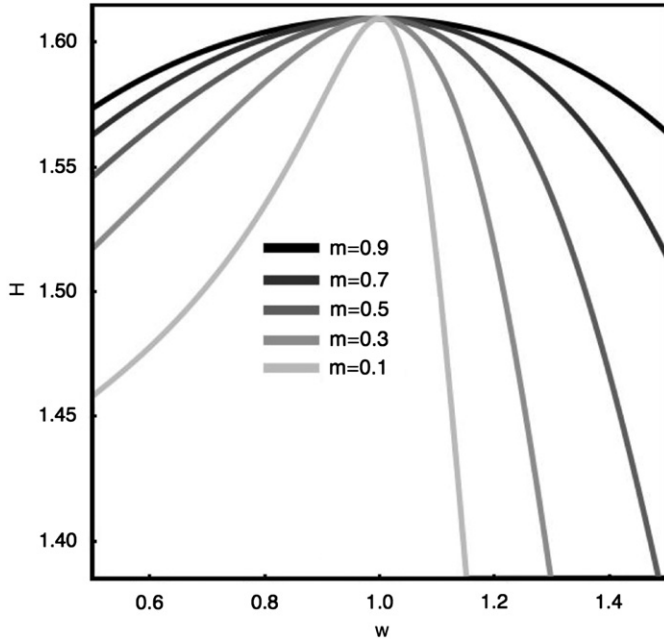
Fig. 3. Plots of the information-theoretic diversity metric, H, for a nearly neutral local community of $J_L = 75$ individuals and $S = 5$ species over various values of the ecological fitness advantage, w, and dispersal ability, m. The symmetric species share a common dispersal ability of $m_o = 0.1$, and all species share a common relative metacommunity abundance such that $x = x_i = 0.2$. Diversity peaks where expected local abundances are equivalent, and this occurs in the symmetric limit, given by $w = 1$ and $m = m_o$. The asymmetries in dispersal ability shown here have a negligible impact on diversity when all species are symmetric in ecological fitness, so each curve in this figure peaks at approximately the same value of H near $w = 1$. Away from the peak, declines in diversity are regulated by mass-effects, with more gradual declines at higher values of m.

stationary distribution (McKane et al., 2000)

$$P_n^{n*} = \binom{J_L}{n} \frac{B(Ix+n, J_L+I(1-x)-n)}{B(Ix, I(1-x))}, \tag{32}$$

where $I = m(J_L - 1)/(1-m)$. This result follows from the symmetric limit of Eq. (21) after applying the identity $\Gamma(a)\Gamma(1-a) = \pi/\sin(\pi a)$. The expected abundance and variance are obtained from the symmetric limits of Eqs. (25) and (26), respectively, after applying the identities in Eqs. (C.1.0.1) and (C.2.2.2)

$$E[N^*] = xJ_L, \tag{33}$$

$$\text{Var}[N^*] = x(1-x)J_L \frac{J_L+I}{1+I}. \tag{34}$$

Finally, the symmetric limits of Eqs. (11), (16), and (28) all yield the stationary sampling distribution for a neutral local community (Etienne and Alonso, 2005; Etienne et al., 2007)

$$P_{\vec{n}}^{n*} = \binom{J_L}{n_1, \ldots, n_S} \frac{1}{(I)_{J_L}} \prod_{i=1}^{S} (Ix_i)_{n_i}. \tag{35}$$

In the special case of complete neutrality, Eq. (35) is an exact result of the general model, Eq. (4). This sampling distribution continues to hold when the assumptions of zero-sum dynamics and stationarity are relaxed (Etienne et al., 2007; Haegeman and Etienne, 2008).

6. Discussion

We have developed a general sampling theory that extends Hubbell's neutral theory of local communities and metacommunities to include asymmetries in ecological fitness and dispersal ability. We anticipate that a parameterization of additional biological complexity, such as asymmetries in survivorship probabilities or differences between the establishment probabilities of local reproduction and immigration, may be incorporated without significant changes to the structure of our analytical results. Although the machinery is significantly more complicated for asymmetric theories than their symmetric counterparts, some analytical calculations remain tractable. We find approximate sampling distributions for general and nearly neutral communities that yield Hubbell's theory in the symmetric limit. Our fully normalized approximation in the nearly neutral case may provide a valuable statistical tool for determining the degree to which an observed SAD is consistent with the assumption of complete neutrality. To facilitate a Bayesian analysis, we have enabled rapid computation of the required hypergeometric functions by deriving previously unknown asymptotic expansions.

Acknowledgements

Appendix A. Sampling asymmetric metacommunities

The analytical insights of Etienne et al. (2007) suggest a clear prescription for translating local community dynamics into metacommunities dynamics in the context of Hubbell's (2001) unified neutral theory of biodiversity and biogeography: replace probabilities of immigration, m_j, with probabilities of speciation, v_j; assume $x_j \sim 1/S_T + \mathcal{O}(1/S_T^2)$ for all j, where S_T is the total number of species that could possibly appear through speciation events; and consider asymptotics as S_T becomes large.

Following this recipe, we translate the transition probabilities for asymmetric local communities, Eq. (3), into the transition probabilities for asymmetric metacommunities (M)

$$T_{ij\vec{n}}^M = \frac{n_i}{J_M}\left((1-v_j)\frac{w_j n_j}{\sum_{k=1}^{S_T} w_k n_k - w_i} + v_j \frac{w_j}{\sum_{k=1}^{S_T} w_k} + \mathcal{O}\left(\frac{1}{S_T}\right)\right), \tag{A.1}$$

where J_M is the number of individuals in the metacommunity, $w_j/\sum_{k=1}^{S_T} w_k$ is the probability that an individual of species j establishes following a speciation event, and

$$\Theta_{ij}^M = \Theta(J_M - (n_i+1))\Theta(n_j-1). \tag{A.2}$$

Metacommunity dynamics are governed by the master equation

$$\frac{dP_{\vec{n}}^M}{d\tau} = \sum_{i=1}^{S_T}\sum_{j=1, j\neq i}^{S_T}(T_{ij\vec{n}+\vec{e}_i-\vec{e}_j}^M P_{\vec{n}+\vec{e}_i-\vec{e}_j}^M - T_{ji\vec{n}}^M P_{\vec{n}}^M)\Theta_{ij}^M. \tag{A.3}$$

If $v_j > 0$ for all j, there are no absorbing states, so for large-S_T, there is a nonzero probability that any given species j exists. Analogous develops to those in Section 3 show that detailed balance in the general theory is approximated by

$$P_{\vec{n}}^{g,M*} = Z_{g,M}^{-1}\binom{J_M}{n_1, \ldots, n_{S_T}}\prod_{k=1}^{S_T} w_k^{n_k}(1-v_k)^{n_k}(\phi_{k\vec{n}}^M)_{n_k}, \tag{A.4}$$

where

$$\phi_{k\vec{n}}^M = \theta_k \frac{(\sum_{l=1}^{S_T} w_l n_l - w_k)/(J_M-1)}{\sum_{l=1}^{S_T} w_l}, \tag{A.5}$$

and $\theta_k = v_k(J_M-1)/(1-v_k)$ is the generalization of Hubbell's "fundamental biodiversity number" (Hubbell, 2001). The fitness-symmetric (s) distribution

$$P_{\vec{n}}^{\mathrm{s,M*}} = Z_{\mathrm{s,M}}^{-1} \binom{J_M}{n_1,\ldots,n_{S_T}} \prod_{k=1}^{S_T} (1-v_k)^{n_k} (\theta_k/S_T)_{n_k}, \tag{A.6}$$

satisfies detailed balance up to $\mathcal{O}(1/S_T)$.

For the special case of nearly neutral metacommunities, we translate the marginal dynamics for an asymmetric species in an otherwise symmetric local community into the marginal dynamics for an asymmetric species in an otherwise symmetric metacommunity. The transition probabilities are

$$g_n^M = \frac{J_M-n}{J_M}\left((1-v)\frac{wn}{J_M+n(w-1)-1}+v\frac{w}{S_T}+\mathcal{O}\left(\frac{1}{S_T}\right)\right),$$

$$r_n^M = \frac{n}{J_M}\left((1-v_o)\frac{J_M-n}{J_M+n(w-1)-w}+v_o\left(1-\frac{w}{S_T}\right)+\mathcal{O}\left(\frac{1}{S_T}\right)\right), \tag{A.7}$$

where the asymmetric focal species has speciation probability v and enjoys an ecological fitness advantage, w, over all other species, which share a common probability of speciation v_o. If $n=0$ is an accessible state, then as S_T becomes large and w remains finite, the equilibrium probability of observing the asymmetric species approaches zero. However, if we assume that the asymmetric species is identified and known to exist at nonzero abundance levels, the stationary distribution is

$$P_n^{\mathrm{nn,M*}} = Z_M \binom{J_M}{n} \eta_M^n \frac{\mathrm{B}(\lambda_M+n,\xi_M-n)}{\mathrm{B}(\lambda_M,\xi_M)}, \tag{A.8}$$

with

$$Z_M^{-1} = {}_2F_1(-J_M,\lambda_M;1-\xi_M;\eta_M)-1, \tag{A.9}$$

and

$$\eta_M = w\frac{(J_M-1)(J_M+\theta_o-1)}{(J_M-\theta_o(w-1)-1)(J_M+\theta-1)}+\mathcal{O}\left(\frac{1}{S_T}\right),$$

$$\lambda_M = \frac{\theta}{S_T}+\mathcal{O}\left(\frac{1}{S_T^2}\right),$$

$$\xi_M = 1+\frac{(J_M-1)(J_M+\theta_o-1)}{J_M-\theta_o(w-1)-1}+\mathcal{O}\left(\frac{1}{S_T}\right), \tag{A.10}$$

where $\theta = (J_M-1)v/(1-v)$ and $\theta_o = (J_M-1)v_o/(1-v_o)$ are Hubbell's "fundamental biodiversity numbers" for the asymmetric species and all other species, respectively.

An approximate multivariate stationary distribution is obtained in an identical manner to the derivation of Eq. (28)

$$P_{\vec{n}}^{\mathrm{nn,M*}} = Z_M \binom{J_M}{n,n_2,\ldots,n_{S_T}} \eta_M^n \frac{\mathrm{B}(\lambda_M+n,\xi_M-n)}{\mathrm{B}(\lambda_M,\xi_M)}$$

$$\times \frac{1}{((1-1/S_T)\phi_{n,M})_{J_M-n}} \prod_{i=2}^{S_T} (\phi_{n,M}/S_T)_{n_i}, \tag{A.11}$$

where

$$\phi_{n,M} = \theta_o\left(1+\frac{n(w-1)}{J_M-1}\right)+\mathcal{O}\left(\frac{1}{S_T}\right). \tag{A.12}$$

We now propose a modest extension to the prescription in Etienne et al. (2007) for converting multivariate distributions over labelled abundance vectors to distributions over unlabelled abundance vectors. Because the asymmetric focal species has been identified and is known to exist with abundance $n>0$, this species must be labelled, while all other species are equivalent and may be unlabelled. Therefore, we aim to transform Eq. (A.11) into a multivariate distribution over the "mostly unlabelled" states

$\hat{\vec{n}} = (n,\hat{n}_2,\ldots,\hat{n}_S)$, where S is the number of species observed in a sample and each $(\hat{n}_2,\ldots,\hat{n}_S)$ is an integer partition of J_M-n. (To provide an example, if $J_M=3$, four distinct states are accessible: (3) with $S=1$, (2,1) with $S=2$, (1,2) with $S=2$, and (1,1,1) with $S=3$.) The conversion is given by

$$P_{\hat{\vec{n}}}^{\mathrm{nn,M*}} = \frac{(S_T-1)!}{\prod_{i=0}^{J_M-n} \hat{\Phi}_i!} P_{\vec{n}}^{\mathrm{nn,M*}}, \tag{A.13}$$

where $\hat{\Phi}_i$ is the number of elements in $(\hat{n}_2,\ldots,\hat{n}_S)$ equal to i. Note that $\hat{\Phi}_0 = S_T-1-(S-1)$. Taking the leading behavior for large-S_T, we obtain a modification of the Ewens (1972) sampling distribution appropriate to nearly neutral metacommunities

$$\lim_{S_T\to\infty} P_{\hat{\vec{n}}}^{\mathrm{nn,M*}} = \lim_{S_T\to\infty} \frac{(S_T-1)!}{(S_T-S)!} \frac{J_M!}{n\prod_{i=2}^S \hat{n}_i \prod_{i=1}^{J_M-n} \hat{\Phi}_i!}$$

$$\times \frac{\eta_M^n(\theta/S_T)_n(\xi_M)_{-n}}{(n-1)!({}_2F_1(-J_M,\theta/S_T;1-\xi_M;\eta_M)-1)}$$

$$\times \frac{1}{((1-1/S_T)\phi_{n,M})_{J_M-n}} \prod_{i=2}^S \frac{(\phi_{n,M}/S_T)_{\hat{n}_i}}{(\hat{n}_i-1)!}$$

$$= \lim_{S_T\to\infty} \frac{(S_T-1)!}{(S_T-S)!}\left(\frac{1}{S_T}\right)^{S-1} \frac{J_M!}{n\prod_{i=2}^S \hat{n}_i \prod_{i=1}^{J_M-n} \hat{\Phi}_i!}$$

$$\times \frac{\theta/S_T}{{}_2F_1(-J_M,\theta/S_T;1-\xi_M;\eta_M)-1} \eta_M^n(\xi_M)_{-n} \frac{\phi_{n,M}^{S-1}}{(\phi_{n,M})_{J_M-n}}$$

$$= \hat{Z}_M \frac{J_M!}{n(J_M-n)!} \eta_M^n(\xi_M)_{-n} \frac{(J_M-n)!}{\prod_{i=2}^S \hat{n}_i \prod_{i=1}^{J_M-n} \hat{\Phi}_i!} \frac{\phi_{n,M}^{S-1}}{(\phi_{n,M})_{J_M-n}}, \tag{A.14}$$

where $(a)_0=1$ allows us to take a product over the observed species, S, rather than the total number of possible species, S_T, in the first expression; $(z)_n/(n-1)! \sim z+\mathcal{O}(z^2)$ as z approaches 0 for $n>0$ has been used to obtain the second expression; l'Hôpital's rule along with $\lim_{b\to 0}\partial_2F_1(a,b;c;z)/\partial b = az\,{}_3F_2(a+1,1,1;c+1,2;z)/c$ has been used to obtain the third expression; and

$$\hat{Z}_M^{-1} = \frac{J_M\eta_M}{\xi_M-1}\,{}_3F_2(1-J_M,1,1;2-\xi_M,2;\eta_M), \tag{A.15}$$

with asymptotics of the hypergeometric function provided in Appendices C.3 and C.4. In the neutral limit, we obtain a modification to the Ewens sampling distribution for the scenario where a single species is labelled and guaranteed to exist

$$\lim_{S_T\to\infty} P_{\hat{\vec{n}}}^{\mathrm{n,M*}} = \left(\sum_{i=1}^{J_M} \frac{\theta}{\theta+i} + \frac{J_M}{\theta+J_M}\right)^{-1} \frac{J_M!}{n\prod_{i=2}^S \hat{n}_i \prod_{i=1}^{J_M-n} \hat{\Phi}_i!} \frac{\theta^S}{(\theta)_{J_M}}. \tag{A.16}$$

Converting this result to a distribution over the "fully unlabelled" states $\hat{\vec{n}} = (\hat{n}_1,\ldots,\hat{n}_S)$, we multiply by

$$\left(\sum_{i=1}^{J_M} \frac{\theta}{\theta+i} + \frac{J_M}{\theta+J_M}\right) \frac{\prod_{i=1}^{J_M-n} \hat{\Phi}_i!}{\prod_{i=1}^{J_M} \hat{\hat{\Phi}}_i}!, \tag{A.17}$$

and recover the Ewens sampling distribution (Ewens, 1972), which is also the sampling distribution for Hubbell's metacommunity theory (Hubbell, 2001).

Appendix B. Marginal dynamics for the local community

We first demonstrate that the marginal dynamics of the asymmetric species in Eq. (19) can be derived from the multivariate dynamics of Eq. (4). Let

$$\sum^{J_L-n}\delta \equiv \sum_{n_2=0}^{J_L-n} \ldots \sum_{n_S=0}^{J_L-n} \delta(J_L-n-n_2-\cdots-n_S), \tag{B.1}$$

so that the marginal distribution for the asymmetric species is given by

$$P_n = \sum_{\vec{n}}^{J_L - n} \delta P_{\vec{n}}. \tag{B.2}$$

Applying Eq. (B.1) to both sides of Eq. (4), we obtain

$$
\begin{aligned}
\frac{dP_n}{d\tau} &= \sum^{J_L-n} \delta \sum_{i=1}^{S} \sum_{j=1,j\neq i}^{S} (T_{ij\vec{n}+\vec{e}_i-\vec{e}_j} P_{\vec{n}+\vec{e}_i-\vec{e}_j} - T_{ji\vec{n}} P_{\vec{n}}) \Theta_{ij} \\
&= \sum^{J_L-n} \delta \sum_{i=2}^{S} \sum_{j=2,j\neq i}^{S} (T_{ij\vec{n}+\vec{e}_i-\vec{e}_j} P_{\vec{n}+\vec{e}_i-\vec{e}_j} - T_{ji\vec{n}} P_{\vec{n}}) \Theta_{ij} \\
&\quad + \sum^{J_L-n} \delta \sum_{j=2}^{S} (T_{1j\vec{n}+\vec{e}_1-\vec{e}_j} P_{\vec{n}+\vec{e}_1-\vec{e}_j} - T_{j1\vec{n}} P_{\vec{n}}) \Theta_{1j} \\
&\quad + \sum^{J_L-n} \delta \sum_{i=2}^{S} (T_{i1\vec{n}+\vec{e}_i-\vec{e}_1} P_{\vec{n}+\vec{e}_i-\vec{e}_1} - T_{1i\vec{n}} P_{\vec{n}}) \Theta_{i1}.
\end{aligned}
\tag{B.3}
$$

By inspection, the first term is identically zero and the remaining terms generate the right-hand side of Eq. (19), namely

$$\sum^{J_L-n} \delta \sum_{j=2}^{S} T_{1j\vec{n}+\vec{e}_1-\vec{e}_j} P_{\vec{n}+\vec{e}_1-\vec{e}_j} \Theta_{1j} = r_{n+1} \Theta(J_L - (n+1)) P_{n+1},$$

$$\sum^{J_L-n} \delta \sum_{j=2}^{S} T_{j1\vec{n}} P_{\vec{n}} \Theta_{1j} = g_n \Theta(J_L - (n+1)) P_n,$$

$$\sum^{J_L-n} \delta \sum_{i=2}^{S} T_{i1\vec{n}+\vec{e}_i-\vec{e}_1} P_{\vec{n}+\vec{e}_i-\vec{e}_1} \Theta_{i1} = g_{n-1} \Theta(n-1) P_{n-1},$$

$$\sum^{J_L-n} \delta \sum_{i=2}^{S} T_{1i\vec{n}} P_{\vec{n}} \Theta_{i1} = r_n \Theta(n-1) P_n. \tag{B.4}$$

To provide an illustration, let $J_L = 4$, $S = 3$, and $n = 1$. Then

$$
\begin{aligned}
&\sum^{J_L-n} \delta \sum_{j=2,j\neq i}^{S} T_{1j\vec{n}+\vec{e}_1-\vec{e}_j} P_{\vec{n}+\vec{e}_1-\vec{e}_j} \Theta_{1j} \\
&= (P_{(2,0,2)}(T_{12(2,0,2)}+T_{13(2,0,2)})+P_{(2,1,1)}(T_{12(2,1,1)}+T_{13(2,1,1)}) \\
&\quad + P_{(2,2,0)}(T_{12(2,2,0)}+T_{13(2,2,0)}))\Theta(J_L-(n+1)) \\
&= r_{n+1} \Theta(J_L-(n+1))(P_{(2,0,2)}+P_{(2,1,1)}+P_{(2,2,0)}) \\
&= r_{n+1} \Theta(J_L-(n+1)) P_{n+1},
\end{aligned}
\tag{B.5}
$$

where we have used the definitions of Θ_{ij} from Eq. (5), r_n from Eq. (18), and P_n from Eq. (B.2).

We construct an approximation to the multivariate sampling distribution of a nearly neutral community, $P_{\vec{n}}^{\mathrm{nn}*}$, by following the subsample approach of Etienne and Alonso (2005) and Etienne et al. (2007) that centers on the identity

$$P_{\vec{n}}^{\mathrm{nn}*} = P_n^{\mathrm{nn}*} \prod_{f=2}^{S-1} P_{n_f|n,n_2,\ldots,n_{f-1}}^{\mathrm{nn}*}. \tag{B.6}$$

Assuming that P_n, $P_{n_2|n}$, ..., $P_{n_{f-1}|n,n_2,\ldots,n_{f-2}}$ are nonzero and stationary, we argue that conditional marginal dynamics for $P_{n_f|n,n_2,\ldots,n_{f-1}}$ are approximated by the master equation

$$
\begin{aligned}
\frac{dP_{n_f|n,n_2,\ldots,n_{f-1}}}{d\tau} =\ & g_{n_f-1|n,n_2,\ldots,n_{f-1}} \Theta(n_f-1) P_{n_f-1|n,n_2,\ldots,n_{f-1}} \\
& + r_{n_f+1|n,n_2,\ldots,n_{f-1}} \Theta(J_L-(n_f+1)) P_{n_f+1|n,n_2,\ldots,n_{f-1}} \\
& - (g_{n_f|n,n_2,\ldots,n_{f-1}} \Theta(J_L-(n_f+1)) \\
& + r_{n_f|n,n_2,\ldots,n_{f-1}} \Theta(n_f-1)) P_{n_f|n,n_2,\ldots,n_{f-1}},
\end{aligned}
\tag{B.7}
$$

where

$$
\begin{aligned}
g_{n_f|n,n_2,\ldots,n_{f-1}} &\equiv \sum_{i=f+1}^{S} T_{if(n,n_2,\ldots,n_f,\ldots,n_S)} \\
&= \frac{J_L-\tilde{n}_f}{J_L}\left((1-m_o)\frac{n_f}{J_L+n(w-1)-1}+m_o\frac{x_f}{1+x(w-1)}\right),
\end{aligned}
$$

$$
\begin{aligned}
r_{n_f|n,n_2,\ldots,n_{f-1}} &\equiv \sum_{i=f+1}^{S} T_{fi(n,n_2,\ldots,n_f,\ldots,n_S)} \\
&= \frac{n_f}{J_L}\left((1-m_o)\frac{J_L-\tilde{n}_f}{J_L+n(w-1)-1}+m_o\frac{1-\tilde{x}_f}{1+x(w-1)}\right),
\end{aligned}
\tag{B.8}
$$

and

$$\tilde{n}_f = n + \sum_{k=2}^{f} n_k,$$

$$\tilde{x}_f = x + \sum_{k=2}^{f} x_k. \tag{B.9}$$

Eq. (B.7) can be derived from the multivariate dynamics of Eq. (4) under the approximation that stochastic variables $(N(\tau), N_2(\tau), \ldots, N_{f-1}(\tau)) = (n, n_2, \ldots, n_{f-1})$ are fixed in time such that $T_{ij\vec{n}} = 0$ for $i, j < f$. In this scenario, the summations on the right-hand side of Eq. (4) may begin at f

$$\frac{dP_{\vec{n}}}{d\tau} = \sum_{i=f}^{S} \sum_{j=f,j\neq i}^{S} \left(T_{ij\vec{n}+\vec{e}_i-\vec{e}_j} P_{\vec{n}+\vec{e}_i-\vec{e}_j} - T_{ji\vec{n}} P_{\vec{n}}\right)\Theta_{ij}. \tag{B.10}$$

Given the identity

$$P_{\vec{n}} = P_n P_{n_2|n} \cdots P_{n_{f-1}|n,n_2,\ldots,n_{f-2}} P_{n_f,\ldots,n_S|n,n_2,\ldots,n_{f-1}}, \tag{B.11}$$

the stationary factor $P_n P_{n_2|n} \cdots P_{n_{f-1}|n,n_2,\ldots,n_{f-2}}$ cancels from both sides of Eq. (B.10) to yield

$$
\begin{aligned}
&\frac{dP_{\vec{n}_f|n,n_2,\ldots,n_{f-1}}}{d\tau} \\
&= \sum_{i=f}^{S} \sum_{j=f,j\neq i}^{S} (T_{ij\vec{n}+\vec{e}_i-\vec{e}_j} P_{\vec{n}_f+\vec{e}_{fi}-\vec{e}_{fj}|n,n_2,\ldots,n_{f-1}} - T_{ji\vec{n}} P_{\vec{n}_f|n,n_2,\ldots,n_{f-1}})\Theta_{ij},
\end{aligned}
\tag{B.12}
$$

where $\vec{n}_f \equiv (n_f, \ldots, n_S)$ and e_{fi} is an $(S-f+1)$-dimensional unit vector along the ith-direction. Now let

$$\sum^{J_L-\cdots-n_f} \delta \equiv \sum_{n_{f+1}=0}^{J_L-n-n_2-\cdots-n_f} \cdots \sum_{n_S=0}^{J_L-n-n_2-\cdots-n_f} \delta(J_L-n-n_2-\cdots-n_S), \tag{B.13}$$

so that

$$P_{n_f|n,n_2,\ldots,n_{f-1}} = \sum^{J_L-\cdots-n_f} \delta P_{\vec{n}_f|n,n_2,\ldots,n_{f-1}}. \tag{B.14}$$

Applying Eq. (B.13) to both sides of Eq. (B.12), we obtain

$$
\begin{aligned}
&\frac{dP_{n_f|n,n_2,\ldots,n_{f-1}}}{d\tau} \\
&= \sum^{J_L-\cdots-n_f} \delta \sum_{i=f+1}^{S} \sum_{j=f+1,j\neq i}^{S} (T_{ij\vec{n}+\vec{e}_i-\vec{e}_j} P_{\vec{n}_f+\vec{e}_{fi}-\vec{e}_{fj}|n,n_2,\ldots,n_{f-1}} \\
&\quad - T_{ji\vec{n}} P_{\vec{n}_f|n,n_2,\ldots,n_{f-1}})\Theta_{ij} \\
&\quad + \sum^{J_L-\cdots-n_f} \delta \sum_{j=f+1}^{S} (T_{fj\vec{n}+\vec{e}_f-\vec{e}_j} P_{\vec{n}_f+\vec{e}_{ff}-\vec{e}_{fj}|n,n_2,\ldots,n_{f-1}} \\
&\quad - T_{jf\vec{n}} P_{\vec{n}_f|n,n_2,\ldots,n_{f-1}})\Theta_{fj} \\
&\quad + \sum^{J_L-\cdots-n_f} \delta \sum_{i=f+1}^{S} (T_{if\vec{n}+\vec{e}_i-\vec{e}_f} P_{\vec{n}_f+\vec{e}_{fi}-\vec{e}_{ff}|n,n_2,\ldots,n_{f-1}} \\
&\quad - T_{fi\vec{n}} P_{\vec{n}_f|n,n_2,\ldots,n_{f-1}})\Theta_{if}.
\end{aligned}
\tag{B.15}
$$

By inspection, the first term is identically zero and the remaining terms generate the right-hand side of Eq. (B.7), namely

$$\sum^{J_L-\cdots-n_f} \delta \sum_{j=f+1}^{S} T_{fj\vec{n}+\vec{e}_f-\vec{e}_j} P_{\vec{n}_f+\vec{e}_{ff}-\vec{e}_{fj}|n,n_2,\ldots,n_{f-1}} \Theta_{fj}$$
$$= r_{n_f+1|n,n_2,\ldots,n_{f-1}} \Theta(J_L-(n_f+1)) P_{n_f+1|n,n_2,\ldots,n_{f-1}},$$

$$\sum^{J_L-\cdots-n_f} \delta \sum_{j=f+1}^{S} T_{jf\vec{n}} P_{\vec{n}_f|n,n_2,\ldots,n_{f-1}} \Theta_{fj}$$
$$= g_{n_f|n,n_2,\ldots,n_{f-1}} \Theta(J_L-(n_f+1)) P_{n_f|n,n_2,\ldots,n_{f-1}},$$

$$\sum^{J_L-\cdots-n_f} \delta \sum_{i=f+1}^{S} T_{if\vec{n}+\vec{e}_i-\vec{e}_f} P_{\vec{n}_f+\vec{e}_{fi}-\vec{e}_{ff}|n,n_2,\ldots,n_{f-1}} \Theta_{if}$$
$$= g_{n_f-1|n,n_2,\ldots,n_{f-1}} \Theta(n_f-1) P_{n_f-1|n,n_2,\ldots,n_{f-1}},$$

$$\sum^{J_L-\cdots-n_f} \delta \sum_{i=f+1}^{S} T_{fi\vec{n}} P_{\vec{n}_f|n,n_2,\ldots,n_{f-1}} \Theta_{if}$$
$$= r_{n_f|n,n_2,\ldots,n_{f-1}} \Theta(n_f-1) P_{n_f|n,n_2,\ldots,n_{f-1}}. \tag{B.16}$$

The stationary distribution of Eq. (B.7) is a Pólya distribution (Johnson et al., 1992)

$$P_{n_f|n,n_2,\ldots,n_{f-1}}^{\text{nn}*}$$
$$= N_{Pf} \binom{J_L-\tilde{n}_{f-1}}{n_f} \frac{\text{B}(\phi_{on}x_f+n_f,J_L-\tilde{n}_{f-1}+\phi_{on}(1-\tilde{x}_f)-n_f)}{\text{B}(\phi_{on}x_f,J_L-\tilde{n}_{f-1}+\phi_{on}(1-\tilde{x}_f))}, \tag{B.17}$$

where

$$N_{Pf}^{-1} = \frac{\text{B}(J_L-\tilde{n}_{f-1}+\phi_{on}(1-\tilde{x}_{f-1}),\phi_{on}(1-\tilde{x}_f))}{\text{B}(\phi_{on}(1-\tilde{x}_{f-1}),J_L-\tilde{n}_{f-1}+\phi_{on}(1-\tilde{x}_f))}, \tag{B.18}$$

and

$$\phi_{on} = I_o \frac{1+n(w-1)/(J_L-1)}{1+x(w-1)}. \tag{B.19}$$

Plugging Eqs. (21) and (B.17) into (B.6), we obtain the approximate sampling distribution of Eq. (28). To validate Eq. (28), we demonstrate approximate detailed balance, as defined by Eq. (10), for large-J_L nearly neutral communities where $S, w-1 \ll \sum_{l=1}^{S} w_l n_l - 1 = J_L-1+n(w-1)$ such that $\varepsilon \equiv (w-1)/(J_L-1+n(w-1))$ is a small number. For $i,j \geq 2$, detailed balance is exact. But for $i=1$ and $j \geq 2$, we have

$$\frac{P_{\vec{n}}^{\text{nn}*}}{P_{\vec{n}+\vec{e}_1-\vec{e}_j}^{\text{nn}*}} = \frac{n+1}{n_j} \frac{1}{w} \frac{1-m_o}{1-m} \frac{n_j-1+\phi_{on}x_j}{n+\phi_n x}$$
$$\times \frac{((1-x)\phi_{on+1})_{J_L-n-1}}{((1-x)\phi_{on})_{J_L-n-1}} \prod_{k=2}^{S} \frac{(\phi_{on}x_k)_{n_k-\delta_{jk}}}{(\phi_{on+1}x_k)_{n_k-\delta_{jk}}}, \tag{B.20}$$

where

$$\phi_n = I \frac{1+n(w-1)/(J_L-1)}{1+x(w-1)}. \tag{B.21}$$

Given $(a(1+\varepsilon))_n \sim (a)_n + \mathcal{O}(\varepsilon)$, we find

$$\frac{P_{\vec{n}}^{\text{nn}*}}{P_{\vec{n}+\vec{e}_1-\vec{e}_j}^{\text{nn}*}} = \frac{T_{1j\vec{n}+\vec{e}_1-\vec{e}_j}}{T_{j1\vec{n}}} + \mathcal{O}(S\varepsilon). \tag{B.22}$$

The case of $i \geq 2$ and $j=1$ is similar.

Appendix C. Asymptotic expansions for hypergeometric functions

Calculating Eqs. (21), (25), (27), (28), (A.8), (A.11), and (A.14) for large communities requires computationally intensive evaluations of hypergeometric functions. To address this problem, one of us (N.M.T.) developed previously unknown asymptotic expansions. All required expansions are summarized here. Relevant details can be found in Abramowitz and Stegun (1965), Luke (1969), Wong (2001), and Gil et al. (2007).

C.1. Expanding $_3F_2(1-J_L,1+\lambda,2;2-\xi,1;\eta)$

Using the reduction formula

$$_3F_2(a,b,2;c,1;z) = \frac{abz}{c} {}_2F_1(a+1,b+1;c+1;z) + {}_2F_1(a,b;c;z), \tag{C.1.0.1}$$

this case can be expanded with the methods of Appendix C.2.

C.2. Expanding $_2F_1(\alpha-J_L,\alpha+\lambda;\alpha+1-\xi;\eta)$

C.2.1. Notation
We write

$$a = \alpha-J_L, \quad b = \alpha+\beta+\mu J_L, \quad c = \alpha+\gamma+\rho J_L, \tag{C.2.1.1}$$

with $\alpha = 0,1,2$ and J_L a positive integer. In terms of w, m, x, and m_o we have

$$\beta = -\frac{mx}{1-m+x(w-1)}, \quad \mu = -\beta, \tag{C.2.1.2}$$

and

$$\gamma = \frac{1-xwm_o+x(w-1)}{1-wm_o+x(w-1)}, \quad \rho = -\gamma. \tag{C.2.1.3}$$

The asymptotic behaviour will be considered of the Gauss hypergeometric function

$$F = {}_2F_1(a,b;c;\eta), \tag{C.2.1.4}$$

for large-J_L, where

$$\eta = w \frac{1-m+x(w-1)}{1-wm_o+x(w-1)}, \tag{C.2.1.5}$$

and

$$w \in (0,\infty), \quad x,m,m_o \in (0,1). \tag{C.2.1.6}$$

C.2.2. The neutral case: $w=1$, $m=m_o$
In this case

$$\eta = 1, \quad \mu = \frac{mx}{1-m}, \quad \rho = -\frac{1-mx}{1-m}. \tag{C.2.2.1}$$

The exact relation

$$_2F_1(-n,b;c;1) = \frac{(c-b)_n}{(c)_n} = \frac{\Gamma(c)\Gamma(c-b+n)}{\Gamma(c+n)\Gamma(c-b)}, \quad n = 0,1,2,\ldots, \tag{C.2.2.2}$$

can be used, together with the asymptotic estimate of the ratio of gamma functions

$$\frac{\Gamma(x+n)}{\Gamma(y+n)} = n^{x-y}(1+\mathcal{O}(1/n)), \quad n \to \infty. \tag{C.2.2.3}$$

C.2.3. Critical values
Considered as functions of w, μ and ρ become unbounded at $w=w_{c_\mu}$ and $w=w_{c_\rho}$, respectively, where

$$w_{c_\mu} = \frac{m+x-1}{x}, \quad w_{c_\rho} = \frac{1-x}{m_o-x}. \tag{C.2.3.1}$$

The case $w \to w_{c_\mu}$: In this case η becomes small, b becomes unbounded, but the product $b\eta$ remains finite. The kth term of the

standard power series of F becomes (see also (C.2.2.3))

$$\frac{(a)_k(b)_k}{k!(c)_k}\eta^k \sim \frac{(a)_k}{k!(c_0)_k}z^k, \tag{C.2.3.2}$$

with

$$z = \lim_{w \to w_{c_\mu}} b\eta = u + vJ_L, \quad c_0 = \lim_{w \to w_{c_\mu}} c = \gamma_0 + \rho_0 J_L, \tag{C.2.3.3}$$

where

$$u = -\frac{mx(m+x-1)}{mx-m_o(m+x-1)}, \quad v = -u, \tag{C.2.3.4}$$

and

$$\gamma_0 = \frac{x(m(1-m_o)+m_o(1-x))}{mx-m_o(m+x-1)}, \quad \rho_0 = -\gamma_0. \tag{C.2.3.5}$$

It follows that F approaches a confluent hypergeometric function:

$$_2F_1(a,b;c;\eta) \to {}_1F_1(a;c_0;z). \tag{C.2.3.6}$$

Further action is needed to obtain an asymptotic approximation of the $_1F_1$-function.

The case $w \to w_{c_\rho}$: In this case η and c become unbounded, but the ratio η/c remains finite. The kth term of the standard power series of F becomes

$$\frac{(a)_k(b)_k}{k!(c)_k}\eta^k \sim \frac{(a)_k(b_0)_k}{k! \, z^k}, \tag{C.2.3.7}$$

with

$$z = \lim_{w \to w_{c_\rho}} c/\eta = u + vJ_L, \quad b_0 = \lim_{w \to w_{c_\rho}} b = \beta_0 + \mu_0 J_L, \tag{C.2.3.8}$$

where

$$u = \frac{m_o(m_o-x)(1-x)}{mx-m_o(m+x-1)}, \quad v = -u, \tag{C.2.3.9}$$

and

$$\beta_0 = -\frac{mx(m_o-x)}{mx-m_o(m+x-1)}, \quad \mu_0 = -\beta_0. \tag{C.2.3.10}$$

It follows that F approaches a $_2F_0$ hypergeometric function

$$_2F_1(a,b;c;\eta) \to {}_2F_0(a,b_0;-;1/z) = \sum_{k=0}^{-a}\frac{(a)_k(b_0)_k}{k!z^k}, \tag{C.2.3.11}$$

because a is a negative integer. This function can be expressed in terms of the Kummer U-function

$$_2F_0(a,b_0;-;1/z) = (-z)^a U(a,1+a-b_0,-z). \tag{C.2.3.12}$$

Further action is needed to obtain an asymptotic approximation of the U-function.

C.2.4. Expansion A

An integral representation is

$$_2F_1(a,b;c;\eta) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-t\eta)^{-a}\,dt, \tag{C.2.4.1}$$

valid for $c > b > 0, \eta < 1$. This integral can be used when $\rho > \mu > 0, \eta < 1$.

As an example, consider

$$r = 3, \quad m = \tfrac{1}{2}, \quad m_o = \tfrac{1}{2}, \quad x = \tfrac{1}{3}. \tag{C.2.4.2}$$

This gives

$$b = \alpha + \tfrac{1}{11}(J_L-1), \quad c = \alpha + 5(J_L-1), \quad \mu = \tfrac{1}{11}, \quad \rho = 5, \quad \eta = -11. \tag{C.2.4.3}$$

In this case the integrand becomes small at $t=0$ and $t=1$, and there is a maximum of the integrand at $t=t_1$, with $t_1 \in (0,1)$. This point gives the main contribution.

Write (C.2.4.1) as

$$_2F_1(a,b;c;\eta) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{\alpha+\beta-1}(1-t)^{\gamma-\beta-1}(1-t\eta)^{-\alpha}e^{-J_L\phi(t)}\,dt, \tag{C.2.4.4}$$

where

$$\phi(t) = -\mu\ln(t)-(\rho-\mu)\ln(1-t)-\ln(1-t\eta). \tag{C.2.4.5}$$

The saddle points t_0 and t_1 are the zeros of $\phi'(t)$. For the example (C.2.4.2) this gives

$$t_0 = -0.01169\ldots, \quad t_1 = 0.1178\ldots, \tag{C.2.4.6}$$

and

$$\phi(t_1) = -0.02136\ldots, \quad \phi''(t_1) = 35.83\ldots. \tag{C.2.4.7}$$

An asymptotic approximation follows from the substitution

$$\phi(t)-\phi(t_1) = \tfrac{1}{2}\phi''(t_1)s^2, \quad \text{sign}(t-t_1) = \text{sign}(s), \tag{C.2.4.8}$$

which gives

$$_2F_1(a,b;c;\eta) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}e^{-J_L\phi(t_1)}\int_{-\infty}^{\infty}f(s)e^{-(1/2)J_L\phi''(t_1)s^2}\,ds, \tag{C.2.4.9}$$

where

$$f(s) = t^{\alpha+\beta-1}(1-t)^{\gamma-\beta-1}(1-t\eta)^{-\alpha}\frac{dt}{ds}. \tag{C.2.4.10}$$

Because locally at $t=t_1$ (or $s=0$), $t = t_1+s+\mathcal{O}(s^2)$, we have $dt/ds=1$ at $s=0$, and

$$f(0) = t_1^{\alpha+\beta-1}(1-t_1)^{\gamma-\beta-1}(1-t_1\eta)^{-\alpha}. \tag{C.2.4.11}$$

This gives the first order approximation

$$_2F_1(a,b;c;\eta) \sim \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}e^{-J_L\phi(t_1)}f(0)\int_{-\infty}^{\infty}e^{-(1/2)J_L\phi''(t_1)s^2}\,ds, \tag{C.2.4.12}$$

that is

$$_2F_1(a,b;c;\eta) \sim \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}e^{-J_L\phi(t_1)}f(0)\sqrt{\frac{2\pi}{J_L\phi''(t_1)}}, \quad J_L \to \infty. \tag{C.2.4.13}$$

C.2.5. Expansion B

Another integral representation is

$$_2F_1(a,b;c;\eta) = \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)}\int_0^{\infty} t^{b-1}(t+1)^{c-b-1}(1+t\eta)^{-a}\,dt, \tag{C.2.5.1}$$

which is only valid for $a = 0,-1,-2,\ldots$ and $c < a+1$. It can be verified by expanding $(1+t\eta)^{-a}$ in powers of η.

We have $\mu > 0$ and $\rho < -1$, and because (see (C.2.1.2), (C.2.1.3) and (C.2.1.5))

$$\eta = -\frac{mx}{(1-m_o x)}\frac{\rho}{\mu}, \tag{C.2.5.2}$$

we see that $\eta \geq 0$.

As an example, consider

$$r = \tfrac{1}{3}, \quad m = \tfrac{1}{2}, \quad m_o = \tfrac{1}{2}, \quad x = \tfrac{1}{3}. \tag{C.2.5.3}$$

This gives

$$b = \alpha + 3(J_L-1), \quad c = \alpha + \tfrac{15}{13}(1-J_L), \quad \mu = 3, \quad \rho = -\tfrac{15}{13}, \quad \eta = \tfrac{1}{13}. \tag{C.2.5.4}$$

Write (C.2.5.1) as

$$_2F_1(a,b;c;\eta) = \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)} \int_0^\infty t^{\alpha+\beta-1}(t+1)^{\gamma-\beta-1}(1+t\eta)^{-\alpha}e^{-J_L\psi(t)}\,dt, \quad (C.2.5.5)$$

where

$$\psi(t) = -\mu\ln(t) - (\rho-\mu)\ln(t+1) - \ln(1+t\eta). \quad (C.2.5.6)$$

The saddle points t_0 and t_1 are for the example (C.2.5.3)

$$t_0 = -74.89\ldots, \quad t_1 = 3.385\ldots, \quad (C.2.5.7)$$

and

$$\psi(t_1) = 2.251\ldots, \quad \psi''(t_1) = 0.04951\ldots. \quad (C.2.5.8)$$

An asymptotic approximation follows from the substitution:

$$\psi(t)-\psi(t_1) = \tfrac{1}{2}\psi''(t_1)s^2, \quad \text{sign}(t-t_1) = \text{sign}(s), \quad (C.2.5.9)$$

which gives

$$_2F_1(a,b;c;\eta) = \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)} e^{-J_L\psi(t_1)} \int_{-\infty}^\infty g(s)e^{-(1/2)J_L\psi''(t_1)s^2}\,ds, \quad (C.2.5.10)$$

where

$$g(s) = t^{\alpha+\beta-1}(1+t)^{\gamma-\beta-1}(1+t\eta)^{-\alpha}\frac{dt}{ds}. \quad (C.2.5.11)$$

Because locally at $t=t_1$ (or $s=0$), $t = t_1+s+\mathcal{O}(s^2)$, we have $dt/ds=1$ at $s=0$, and

$$g(0) = t_1^{\alpha+\beta-1}(1+t_1)^{\gamma-\beta-1}(1+t_1\eta)^{-\alpha}. \quad (C.2.5.12)$$

This gives the first order approximation

$$_2F_1(a,b;c;\eta) \sim \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)} e^{-J_L\psi(t_1)}g(0) \int_{-\infty}^\infty e^{-(1/2)J_L\psi''(t_1)s^2}\,ds, \quad (C.2.5.13)$$

that is

$$_2F_1(a,b;c;\eta) \sim \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)} e^{-J_L\psi(t_1)}g(0)\sqrt{\frac{2\pi}{J_L\psi''(t_1)}}, \quad J_L\to\infty. \quad (C.2.5.14)$$

C.2.6. Expansion C

If $\mu < \rho < -1$ and $\eta < 0$, apply the transformation

$$_2F_1(a,b;c;\eta) = (1-\eta')^a\,_2F_1(a,b';c;\eta'), \quad (C.2.6.1)$$

where

$$b' = c-b = \beta'+\mu'J_L, \quad \beta' = \gamma-\beta, \quad \mu' = \rho-\mu, \quad \eta' = \frac{\eta}{\eta-1}. \quad (C.2.6.2)$$

Now,

$$\mu' > 0, \quad \rho < -1, \quad \eta' > 0, \quad (C.2.6.3)$$

and it follows that Expansion B, Appendix C.2.5, applies to the Gauss function on the right-hand side of (C.2.6.1).

C.2.7. General cases for all non-critical values

1. $w_{c_\mu}, w_{c_\rho} < 0$: For all $w > 0$, we have $\mu > 0$, $\rho < -1$, and $\eta > 0$, so use Expansion B, Appendix C.2.5.
2. $w_{c_\mu} > 0$, $w_{c_\rho} < 0$: For all $w_{c_\mu} > w > 0$, we have $\mu < -1$, $\rho < -1$, and $\eta < 0$, so use Expansion C, Appendix C.2.6.
   For all $w > w_{c_\mu}$, we have $\mu > 0$, $\rho < -1$, and $\eta > 0$, so use Expansion B, Appendix C.2.5.
3. $w_{c_\mu} < 0$, $w_{c_\rho} > 0$: For all $w_{c_\rho} > w > 0$, we have $\mu > 0$, $\rho < -1$, and $\eta > 0$, so use Expansion B, Appendix C.2.5.

For all $w > w_{c_\rho}$, we have $\rho > \mu > 0$ and $\eta < 0$, so use Expansion A, Appendix C.2.4.

4. $w_{c_\rho} > w_{c_\mu} > 0$: For all $w_{c_\rho} > w > 0$, we have $\mu < -1$, $\rho < -1$, and $\eta < 0$, so use Expansion C, Appendix C.2.6.
   For all $w_{c_\rho} > w > w_{c_\mu}$, we have $\mu > 0$, $\rho < -1$, and $\eta > 0$, so use Expansion B, Appendix C.2.4.
   For all $w > w_{c_\rho}$, we have $\rho > \mu > 0$ and $\eta < 0$, so use Expansion A, Appendix C.2.4.

C.3. Expanding $_2F_1(1-J_M,1;2-\xi_M;\eta_M)$

C.3.1. Notation

We write

$$a = 1-J_M, \quad b = 1, \quad c = \sigma+\tau J_M, \quad (C.3.1.1)$$

with

$$\sigma = 1+\frac{1}{1-wv_o}, \quad \tau = -\frac{1}{1-wv_o}. \quad (C.3.1.2)$$

The asymptotic behaviour will be considered of the Gauss hypergeometric function

$$F = \,_2F_1(a,b;c;\eta_M) \quad (C.3.1.3)$$

for large-J_M, where

$$\eta_M = \frac{w(1-v)}{1-wv_o}, \quad (C.3.1.4)$$

and

$$w \in (0,\infty), \quad v,v_o \in (0,1). \quad (C.3.1.5)$$

C.3.2. The neutral case: $w=1$, $v=v_o$

In this case $\eta_M = 1$ and (C.2.2.2) can be used to get an exact result in terms of gamma functions.

C.3.3. The critical case $w_{c_{v_o}} = 1/v_o$

In this case we have (see also Appendix C.2.3)

$$_2F_1(a,b;c;\eta_M) \to \,_2F_0(a,b;-;1/z) = (-z)^a U(a,1+a-b,-z), \quad (C.3.3.1)$$

where

$$z = -\frac{(J_M-1)v_o}{1-v}. \quad (C.3.3.2)$$

C.3.4. The case $0 < w < w_{c_{v_o}}$

Use the integral representation

$$_2F_1(a,b;c;\eta_M) = \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)} \int_0^\infty t^{b-1}f(t)e^{-J_M\phi(t)}\,dt, \quad (C.3.4.1)$$

where

$$f(t) = (1+t)^{\sigma-b-1}(1+\eta_M t)^{-1}, \quad \phi(t) = -\tau\ln(1+t) - \ln(1+\eta_M t). \quad (C.3.4.2)$$

The saddle point t_0 follows from solving $\phi'(t) = 0$. This gives

$$\phi'(t) = -\frac{\tau}{1+t} - \frac{\eta_M}{1+\eta_M t}, \quad t_0 = -\frac{\tau+\eta_M}{\eta_M(\tau+1)}. \quad (C.3.4.3)$$

In terms of w and v

$$t_0 = -\frac{(1-wv_o)(1-w+wv)}{v_o(1-v)w^2}. \quad (C.3.4.4)$$

1. If $0 < w < 1/(1-v) \equiv w_{c_v}$, then the saddle point is negative, and we can substitute $s = \phi(t)$, giving

$$_2F_1(a,b;c;\eta_M) = \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)} \int_0^\infty s^{b-1}g(s)e^{-J_M s}\,ds, \quad (C.3.4.5)$$

where

$$g(s) = f(t)\left(\frac{t}{s}\right)^{b-1}\frac{dt}{ds} = \left(\frac{t}{s}\right)^{b-1}\frac{f(t)}{\phi'(t)}. \tag{C.3.4.6}$$

Apply Watson's lemma by expanding $g(s) = \sum_{k=0}^{\infty} g_k s^k$ to obtain

$$_2F_1(a,b;c;\eta_M) \sim \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)}\sum_{k=0}^{\infty}\frac{\Gamma(b+k)\,g_k}{J_M^{b+k}}. \tag{C.3.4.7}$$

To compute the coefficients g_k we first expand $t = \sum_{k=1}^{\infty} t_k s^k$. The coefficients t_k follow from inverting the expansion

$$s = -\tau\ln(1+t) - \ln(1+\eta_M t) = \sum_{k=1}^{\infty} s_k t^k, \quad s_1 = -\tau - \eta_M. \tag{C.3.4.8}$$

This gives

$$t_1 = -\frac{1}{\tau+\eta_M} = \frac{1-wv_o}{1-w+wv}, \tag{C.3.4.9}$$

and for the first coefficient in the expansion (C.3.4.7) $g_0 = g(0) = t_1^b$. This gives

$$_2F_1(a,b;c;\eta_M) \sim \frac{\Gamma(1+b-c)}{\Gamma(1-c)}\left(\frac{t_1}{J_M}\right)^b. \tag{C.3.4.10}$$

2. If $w > w_{c_v}$, then t_0 is positive, and Laplace's method can be used, as in Appendices C.2.4 and C.2.5. We substitute

$$\frac{1}{2}\phi''(t_0)s^2 = \phi(t) - \phi(t_0), \quad \phi''(t_0) = \frac{\eta_M^2(\tau+1)^3}{\tau(\eta_M-1)^2}, \tag{C.3.4.11}$$

and obtain

$$_2F_1(a,b;c;\eta_M) \sim \frac{\Gamma(1+b-c)}{\Gamma(b)\Gamma(1-c)}\sqrt{\frac{2\pi}{J_M\phi''(t_0)}}e^{-J_M\phi(t_0)}t_0^{b-1}f(t_0), \tag{C.3.4.12}$$

where t_0 is given in (C.3.4.3).

3. If $w = w_{c_v}$, then $t_0 = 0$ and Laplace's method on a half-infinite interval can be used.

C.3.5. The case $w > w_{c_{v_o}}$

Use the integral representation

$$_2F_1(a,b;c;\eta_M) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 f(t)e^{-J_M\phi(t)}\,dt, \tag{C.3.5.1}$$

where

$$f(t) = t^{b-1}(1-t)^{\sigma-b-1}(1-\eta_M t)^{-1}, \quad \phi(t) = -\tau\ln(1-t) - \ln(1-\eta_M t). \tag{C.3.5.2}$$

The saddle point t_0 follows from solving $\phi'(t) = 0$. This gives

$$\phi'(t) = \frac{\tau}{1-t} + \frac{\eta_M}{1-\eta_M t}, \quad t_0 = \frac{\tau+\eta_M}{\eta_M(\tau+1)}. \tag{C.3.5.3}$$

In terms of w and v

$$t_0 = \frac{(1-wv_o)(1-w+wv)}{v_o(1-v)w^2}. \tag{C.3.5.4}$$

1. If $w < w_{c_v}$, then $t_0 < 0$ and Watson's lemma should be used. The result is

$$_2F_1(a,b;c;\eta_M) \sim \frac{\Gamma(c)}{\Gamma(c-b)}\left(\frac{t_1}{J_M}\right)^b, \quad t_1 = \frac{1}{\eta_M+\tau}. \tag{C.3.5.5}$$

2. If $w > w_{c_v}$, then the saddle point t_0 is always inside the interval $(0,1)$, with $t_0 \to 1$ if $w \to \infty$. Laplace's method should be used.

This gives

$$_2F_1(a,b;c;\eta_M) \sim \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\sqrt{\frac{2\pi}{J_M\phi''(t_0)}}e^{-J_M\phi(t_0)}t_0^{b-1}f(t_0), \tag{C.3.5.6}$$

where f, ϕ and t_0 are given in (C.3.5.2)–(C.3.5.3) and

$$\phi''(t_0) = \frac{\eta_M^2(\tau+1)^3}{\tau(\eta_M-1)^2}. \tag{C.3.5.7}$$

3. If $w = w_{c_v}$, then $t_0 = 0$ and Laplace's method on a half-infinite interval can be used.

C.4. Expanding $_3F_2(1-J_M,1,1;2,2-\xi_M;\eta_M)$

C.4.1. Notation
We write

$$a = 1 - J_M, \quad c = \sigma + \tau J_M, \tag{C.4.1.1}$$

with

$$\sigma = 1 + \frac{1}{1-wv_o}, \quad \tau = -\frac{1}{1-wv_o}, \tag{C.4.1.2}$$

The asymptotic behaviour will be considered of the hypergeometric function

$$F = _3F_2(a,1,1;c,2;\eta_M), \tag{C.4.1.3}$$

for large-J_M, where

$$\eta_M = \frac{w(1-v)}{1-wv_o}, \tag{C.4.1.4}$$

and

$$r \in (0,\infty), \quad v, v_o \in (0,1). \tag{C.4.1.5}$$

Note that

$$_3F_2(a,1,1;c,2;\eta_M) = \sum_{k=0}^{\infty}\frac{(a)_k(1)_k(1)_k}{k!\,(c)_k(2)_k}\eta_M^k = \sum_{k=0}^{\infty}\frac{(a)_k}{(c)_k}\frac{\eta_M^k}{k+1}, \tag{C.4.1.6}$$

and because a is a negative integer this series terminates at $k = -a$.

C.4.2. Representation in terms of a Laplace integral
We substitute

$$\frac{1}{k+1} = \int_0^{\infty} e^{-(k+1)w}\,dw, \tag{C.4.2.1}$$

and obtain

$$_3F_2(a,1,1;c,2;\eta_M) = \int_0^{\infty} e^{-w}\,_2F_1(a,1;c,z)\,dw, \quad z = \eta_M e^{-w}. \tag{C.4.2.2}$$

For this representation we use the results of Appendices C.3.4 and C.3.5 when, with η_M replaced by $\eta_M e^{-w}$, the saddle points t_0 of (C.3.4.3) and (C.3.5.3) are negative. This gives two cases.

1. *The case $w < w_{c_{v_o}}$ and $w < w_{c_v}$.* For this case we use the results in (C.3.4.5)–(C.3.4.10) with $b = 1$. We have

$$_2F_1(a,1;c;z) = \frac{1-c}{J_M}g_0 + \mathcal{O}(1/J_M), \tag{C.4.2.3}$$

where $g_0 = t_1$ and $t_1 = -1/(\tau+\eta_M e^{-w})$. This gives

$$_3F_2(a,1,1;c,2;\eta_M) = \frac{c-1}{J_M}\int_0^{\infty} e^{-w}\frac{dw}{\tau+\eta_M e^{-w}} + \mathcal{O}(1/J_M). \tag{C.4.2.4}$$

Evaluating the integral we obtain

$$_3F_2(a,1,1;c,2;\eta_M) = \frac{c-1}{\tau J_M}\frac{\ln(1+\eta_M/\tau)}{\eta_M/\tau} + \mathcal{O}(1/J_M). \tag{C.4.2.5}$$

2. *The case $w > w_{c_{v_0}}$ and $w < w_{c_v}$.* In this case we use the results in (C.3.5.1)–(C.3.5.5), again, with $b=1$ and η_M replaced with $\eta_M e^{-w}$. We use Watson's lemma for (C.3.4.11) by substituting $s = \phi(t)$ and obtain

$$_2F_1(a,1;c;z) = (c-1)\int_0^\infty e^{-J_M s}g(s)\,ds, \quad g(s) = \frac{(1-t)^{\sigma-2}}{1-te^{-w}}\frac{dt}{ds}.$$
(C.4.2.6)

Expanding g at $s=0$ we have $g(s) = g_0 + \mathcal{O}(s)$, with $g_0 = 1/(\tau + \eta_M e^{-w})$ and, as in the above case, we derive

$$_3F_2(a,1,1;c,2;\eta_M) = \frac{c-1}{\tau J_M}\frac{\ln(1+\eta_M/\tau)}{\eta_M/\tau} + \mathcal{O}(1/J_M).$$
(C.4.2.7)

C.4.3. Summing the series by integration

We replace the Pochhammer symbols in (C.4.1.6) by representations in terms of the gamma functions

$$(\alpha)_k = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} = (-1)^k\frac{\Gamma(1-\alpha)}{\Gamma(1-\alpha-k)}.$$
(C.4.3.1)

and replace the gamma functions with large positive argument by their asymptotic forms that follow from

$$\Gamma(az+b) \sim \sqrt{2\pi}\,e^{-az}(az)^{az+b-1/2}, \quad z\to\infty, \quad a>0.$$
(C.4.3.2)

This gives the remaining two cases.

1. *The case $w < w_{c_{v_0}}$ and $w > w_{c_v}$.* In this case $\tau < -1$ and $\eta_M > 0$. We replace the Pochhammer symbols in (C.4.1.6) with the second form in (C.4.3.1). This gives

$$_3F_2(a,1,1;c,2;\eta_M) = \sum_{k=0}^\infty F(k),$$
(C.4.3.3)

where

$$F(k) = \frac{\Gamma(1-a)}{\Gamma(1-c)}\frac{\eta_M^k}{k+1}\frac{\Gamma(1-c-k)}{\Gamma(1-a-k)},$$
(C.4.3.4)

and we replace the summation in (C.4.3.3) by integration, invoking Euler's summation formula, or the compound trapezoidal rule

$$_3F_2(a,1,1;c,2;\eta_M) \sim \frac{\Gamma(1-a)}{\Gamma(1-c)}\int_0^{J_M}\frac{\eta_M^k}{k+1}\frac{\Gamma(1-c-k)}{\Gamma(1-a-k)}\,dk.$$
(C.4.3.5)

In Euler's summation formula additional terms occur but in the present case they can be neglected.
We replace the gamma functions by their asymptotic estimates following from (C.4.3.2). This gives

$$_3F_2(a,1,1;c,2;\eta_M) \sim \frac{\Gamma(1-a)}{\Gamma(1-c)}\int_0^{J_M}\frac{e^{-\phi(k)}}{k+1}\frac{\sqrt{1-a-k}}{\sqrt{1-c-k}}\,dk,$$
(C.4.3.6)

where

$$\phi(k) = -k\ln\eta_M - (1-c-k)\ln(1-c-k) + (1-a-k)\ln(1-a-k).$$
(C.4.3.7)

Then,

$$\phi'(k) = -\ln\eta_M + \ln(1-c-k) - \ln(1-a-k)),$$
(C.4.3.8)

and the derivative vanishes for $k = k_s$, where

$$k_s = \frac{c-1+\eta_M(1-a)}{1-\eta_M} = k_0 + k_1 J_M, \quad k_0 = \frac{\sigma-1}{\eta_M-1}, \quad k_1 = \frac{\tau+\eta_M}{\eta_M-1},$$
(C.4.3.9)

and $0 < k_1 < 1$. The dominant point of the integral in (C.4.3.6) is $k = k_s$, and we apply Laplace's method. We substitute

$$\frac{1}{2}\phi''(k_s)s^2 = \phi(k) - \phi(k_s), \quad \phi''(k_s) = -\frac{(\eta_M-1)^2}{J_M\eta_M(\tau+1)} + \mathcal{O}(1/J_M^2).$$
(C.4.3.10)

This gives

$$_3F_2(1-J_M,1,1;\sigma+\tau J_M,2;\eta_M) \sim \sqrt{\frac{2\pi}{\phi''(k_s)}}F(k_s),$$
(C.4.3.11)

where $F(k)$ is given in (C.4.3.4). After using (C.4.3.2) we obtain

$$_3F_2(1-J_M,1,1;\sigma+\tau J_M,2;\eta_M)$$
$$\sim \sqrt{\frac{2\pi}{\phi''(k_s)}}\frac{\eta_M-1}{(\eta_M+\tau)J_M}\left(-\frac{\eta_M}{\tau}\right)^{1/2-c}\left(\frac{1+\tau}{1-\eta_M}\right)^{a-c}.$$
(C.4.3.12)

2. *The case $w > w_{c_{v_0}}$ and $w > w_{c_v}$.* In this case $\tau > 0$ and $\eta_M < 0$. We use (C.4.1.6), replacing the Pochhammer symbol $(a)_k$ by the second form of (C.4.3.1) and $(c)_k$ by the first. This gives

$$_3F_2(a,1,1;c,2;\eta_M) = \sum_{k=0}^\infty F(k),$$
(C.4.3.13)

where

$$F(k) = \frac{\Gamma(J_M)\Gamma(c)(-\eta_M)^k}{(k+1)\Gamma(J_M-k)\Gamma(c+k)},$$
(C.4.3.14)

and we replace the summation in (C.4.3.13) by integration,

$$_3F_2(a,1,1;c,2;\eta_M) \sim \int_0^{J_M} F(k)\,dk.$$
(C.4.3.15)

Applying the asymptotic estimates of gamma functions in (C.4.3.2), we obtain

$$F(k) = \frac{\Gamma(J_M)\Gamma(c)e^{c+J_M}}{2\pi(k+1)}\sqrt{(J_M-k)(c+k)}e^{-\phi(k)},$$
(C.4.3.16)

where

$$\phi(k) = (c+k)\ln(c+k) + (J_M-k)\ln(J_M-k) - k\ln(-\eta_M).$$
(C.4.3.17)

We have

$$\phi'(k) = \ln(c+k) - \ln(J_M-k) - \ln(-\eta_M),$$
(C.4.3.18)

and the saddle point given by

$$k_s = \frac{c+\eta_M J_M}{\eta_M-1} = k_0 + k_1 J_M, \quad k_0 = \frac{\sigma}{\eta_M-1}, \quad k_1 = \frac{\tau+\eta_M}{\eta_M-1},$$
(C.4.3.19)

where, again, $0 < k_1 < 1$. The dominant point of the integral in (C.4.3.15) is $k = k_s$, and we apply Laplace's method to this integral. We substitute

$$\frac{1}{2}\phi''(k_s)s^2 = \phi(k) - \phi(k_s), \quad \phi''(k_s) = -\frac{(\eta_M-1)^2}{J_M\eta_M(\tau+1)} + \mathcal{O}(1/J_M^2),$$
(C.4.3.20)

where $\phi(k)$ is given in (C.4.3.17). This gives

$$_3F_2(1-J_M,1,1;\sigma+\tau J_M,2;\eta_M) \sim \sqrt{\frac{2\pi}{\phi''(k_s)}}F(k_s),$$
(C.4.3.21)

where $F(k)$ is given in (C.4.3.16). After using (C.4.3.2) we obtain, as in the case above,

$$_3F_2(1-J_M,1,1;\sigma+\tau J_M,2;\eta_M)$$
$$\sim \sqrt{\frac{2\pi}{\phi''(k_s)}}\frac{\eta_M-1}{(\eta_M+\tau)J_M}\left(-\frac{\eta_M}{\tau}\right)^{\frac{1}{2}-c}\left(\frac{1+\tau}{1-\eta_M}\right)^{a-c}.$$
(C.4.3.22)

References

Abramowitz, M., Stegun, I.A. (Eds.), 1965. Handbook of Mathematical Functions. Dover, New York, 1965.

Allouche, O., Kadmon, R., 2009a. Demographic analysis of Hubbell's neutral theory of biodiversity. Journal of Theoretical Biology 258, 274–280.

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

Alonso, D., Etienne, R.S., McKane, A., 2006. The merits of neutral theory. Trends in Ecology & Evolution 21, 452–456.

Alonso, D., McKane, A., 2004. Sampling Hubbell's neutral theory of biodiversity. Ecology Letters 7, 901–910.

Alonso, D., Ostling, A., Etienne, R.S., 2008. The implicit assumption of symmetry and the species abundance distribution. Ecology Letters 11, 93–105.

Babak, P., He, F., 2008. Species abundance distribution and dynamics in two locally coupled communities. Journal of Theoretical Biology 253, 739–748.

Babak, P., He, F., 2009. A neutral model of edge effects. Theoretical Population Biology 75, 76–83.

Brown, J.H., Kodric-Brown, A., 1977. Turnover rates in insular biogeography: effect of immigration on extinction. Ecology 58, 445–449.

Chase, J.M., 2005. Towards a really unified theory for metacommunities. Ecology 19, 182–186.

Chase, J.M., Leibold, M.A., 2003. Ecological Niches. University of Chicago Press, Chicago.

Chave, J., 2004. Neutral theory and community ecology. Ecology Letters 7, 241–253.

Chave, J., Alonso, D., Etienne, R.S., 2006. Comparing models of species abundance. Nature 441, E1–E2.

Chave, J., Muller-Landau, H., Levin, S., 2002. Comparing classical community models: theoretical consequences for patterns of diversity. American Naturalist 159, 1–23.

Conlisk, J., Conlisk, E., Harte, J., 2010. Hubbell's local abundance distribution: insights from a simple colonization rule. Oikos 119, 379–383.

Doncaster, C.P., 2009. Ecological equivalence: a realistic assumption for niche theory as a testable alternative to neutral theory. PLoS ONE 4, e7460.

Dornelas, M., Connolly, S.R., 2008. Multiple modes in a coral species abundance distribution. Ecology Letters 11, 1008–1016.

Ernest, S.K.M., Brown, J.H., Thibault, K.M., White, E.P., Goheen, J.R., 2008. Zero sum, the niche, and metacommunities: long-term dynamics of community assembly. American Naturalist 172, E257–E269.

Etienne, R.S., 2005. A new sampling formula for neutral biodiversity. Ecology Letters 8, 253–260.

Etienne, R.S., 2007. A neutral sampling formula for multiple samples and an 'exact' test of neutrality. Ecology Letters 10, 608–618.

Etienne, R.S., 2009. Maximum likelihood estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. Journal of Theoretical Biology 257, 510–514.

Etienne, R.S., Alonso, D., 2005. A dispersal-limited sampling theory for species and alleles. Ecology Letters 8, 1147–1156.

Etienne, R.S., Alonso, D., 2007. Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. Journal of Statistical Physics 128, 485–510.

Etienne, R.S., Alonso, D., McKane, A., 2007. The zero-sum assumption in neutral biodiversity theory. Journal of Theoretical Biology 248, 522–536.

Etienne, R.S., Olff, H., 2004. A novel genealogical approach to neutral biodiversity theory. Ecology Letters 7, 170–175.

Etienne, R.S., Olff, H., 2005. Confronting different models of community structure to species-abundance data: a Bayesian model comparison. Ecology Letters 8, 493–504.

Ewens, W., 1972. The sampling theory of selectively neutral alleles. Theoretical Population Biology.

Fisher, R., 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Gil, A., Segura, J., Temme, N.M., 2007. Numerical Methods for Special Functions. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Goldberg, D., Barton, A., 1992. Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. American Naturalist 139, 771–801.

Gravel, D., Canham, C.D., Beaudet, M., Messier, C., 2006. Reconciling niche and neutrality: the continuum hypothesis. Ecology Letters 9, 399–409.

Gray, J., Bjoergesaeter, A., Ugland, K., 2005. The impact of rare species on natural assemblages. Journal of Animal Ecology 74, 1131–1139.

Haegeman, B., Etienne, R.S., 2008. Relaxing the zero-sum assumption in neutral biodiversity theory. Journal of Theoretical Biology 252, 288–294.

Harper, J., 1977. Population Biology of Plants. Academic Press, London.

Hastings, A., 1987. Can competition be detected using species co-occurrence data? Ecology 68, 117–123.

He, F., 2005. Deriving a neutral model of species abundance from fundamental mechanisms of population dynamics. Functional Ecology 19, 187–193.

Holt, R.D., 1993. Ecology at the mesoscale: the influence of regional processes on local communities. In: Ricklefs, R., Schluter, D. (Eds.), Species Diversity in Ecological Communities. University of Chicago Press, Chicago, pp. 77–88.

Holyoak, M., Leibold, M.A., Holt, R.D. (Eds.), 2005. Metacommunities. The University of Chicago Press.

Hu, X., He, F., Hubbell, S.P., 2007. Species diversity in local neutral communities. American Naturalist 170, 844–853.

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

Hutchinson, G., 1951. Copepodology for the onithologist. Ecology 32, 571–577.

Johnson, N.L., Kotz, S., Kemp, A.W., 1992. Univariate Discrete Distributions. Wiley Series in Probability and Mathematical Statistics, second ed. John Wiley & Sons Inc., New York, NY.

Kemp, A.W., 1968. Studies in univariate discrete distribution theory based on the generalized hypergeometric function and associated differential equations. Ph.D. Thesis, The Queen's University of Belfast, Belfast.

Leibold, M.A., 1998. Similarity and local co-existence of species in regional biotas. Evolutionary Ecology 12, 95–110.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M., Gonzalez, A., 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7, 601–613.

Levine, J.M., HilleRisLambers, J., 2009. The importance of niches for the maintenance of species diversity. Nature 461, 254–257.

Lin, K., Zhang, D., He, F., 2009. Demographic trade-offs in a neutral model explain death-rate–abundance-rank relationship. Ecology 90, 31–38.

Luke, Y.L., 1969. The special functions and their approximations, vol. I. Mathematics in Science and Engineering, vol. 53. Academic Press, New York.

McKane, A., Alonso, D., Solé, R., 2000. A mean field stochastic theory for species-rich assembled communities. Physical Review E 62, 8466–8484.

McKane, A., Alonso, D., Solé, R., 2004. Analytic solution of Hubbell's model of local community dynamics. Theoretical Population Biology 65, 67–73.

Mouquet, N., Loreau, M., 2003. Community patterns in source-sink metacommunities. American Naturalist 162, 544–556.

O'Dwyer, J., Green, J., 2010. Field theory for biogeography: a spatially explicit model for predicting patterns of biodiversity. Ecology Letters 13, 87–95.

O'Dwyer, J., Lake, J., Ostling, A., Savage, V., 2009. An integrative framework for stochastic, size-structured community assembly. Proceedings of the National Academy of Sciences 106, 6170–6175.

Pigolotti, S., Flammini, A., Maritan, A., 2004. A stochastic model for the species abundance problem in an ecological community. Physical Review E 70, 011916.

Pueyo, S., He, F., Zillio, T., 2007. The maximum entropy formalism and the idiosyncratic theory of biodiversity. Ecology Letters 10, 1017–1028.

Purves, D., Pacala, S., 2005. Ecological drift in niche-structured communities: neutral pattern does not imply neutral process. In: Burslem, D., Pinard, M., Hartley, S. (Eds.), Biotic Interactions in the Tropics. Cambridge University Press, Cambridge, pp. 107–138.

Shmida, A., Wilson, M., 1985. Biological determinants of species diversity. Journal of Biogeography 12, 1–20.

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

Tilman, D., 2004. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences 101, 10854–10861.

Vallade, M., Houchmandzadeh, B., 2003. Analytical solution of a neutral model of biodiversity. Physical Review E 68, 061902.

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., 2003. Neutral theory and relative species abundance in ecology. Nature 424, 1035–1037.

Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2007. Patterns of relative species abundance in rainforests and coral reefs. Nature 450, 45–49.

Walker, S., 2007. When and why do non-neutral metacommunities appear neutral? Theoretical Population Biology 71, 318–331.

Wong, R., 2001. Asymptotic Approximations of Integrals. Classics in Applied Mathematics, vol. 34. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA Corrected reprint of the 1989 original.

Wootton, J., 2009. Field parameterization and experimental test of the neutral theory of biodiversity. Nature 433, 309–312.

Wright, S., 1931. Evolution in Mendelian populations. Genetics 16, 97–159.

Zhang, D., Lin, K., 1997. The effects of competitive asymmetry on the rate of competitive displacement: how robust is Hubbell's community drift model? Journal of Theoretical Biology 188, 361–367.

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