



UNIVERSITY OF LEEDS

This is a repository copy of *Statistical mechanics of ecological systems: Neutral theory and beyond*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/94121/>

Version: Accepted Version

Article:

Azaele, S orcid.org/0000-0002-5153-4833, Suweis, S, Grilli, J et al. (3 more authors) (2016) *Statistical mechanics of ecological systems: Neutral theory and beyond*. *Reviews of Modern Physics*, 88 (3). 035003. ISSN 0034-6861

<https://doi.org/10.1103/RevModPhys.88.035003>

(c) 2016, American Physical Society. This is an author produced version of a paper published in *Reviews of Modern Physics*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Statistical Mechanics of Ecological Systems: Neutral Theory & Beyond

Sandro Azaele^{1,*}, Samir Suweis^{2,*}, Jacopo Grilli³, Igor Volkov⁴, Jayanth R. Banavar⁴,
Amos Maritan²

¹ *Department of Applied Mathematics,
School of Mathematics,
University of Leeds,
Leeds LS2 9JT,
United Kingdom.*

² *Dipartimento di Fisica ‘G. Galilei’ & CNISM,
INFN, Università di Padova,
Via Marzolo 8,
35131 Padova, Italy.*

³ *Department of Ecology and Evolution,
University of Chicago,
1101 East 57th Street,
Chicago, Illinois 60637, USA.*

⁴ *Department of Physics,
University of Maryland,
College Park,
Maryland 20742, USA.*

** These authors contributed equally to this work. Email: suweis@pd.infn.it;
S.Azaele@leeds.ac.uk*

Abstract

The simplest theories often have much merit and many limitations, and in this vein, the value of Neutral Theory (NT) of biodiversity has been the subject of much debate over the past 15 years. NT was proposed at the turn of the century by Stephen Hubbell to explain several patterns observed in the organization of ecosystems. Among ecologists, it had a polarizing effect: There were a few ecologists who were enthusiastic, and there were a larger number who firmly opposed it. Physicists and mathematicians, instead, welcomed the theory with excitement. Indeed, NT spawned several theoretical studies that attempted to explain empirical data and predicted trends of quantities that had not yet been studied. While there are a few reviews of NT oriented towards ecologists, our goal here is to review the quantitative aspects of NT and its extensions for physicists who are interested in learning what NT is, what its successes are and what important problems remain unresolved. Furthermore, we hope that this review could also be of interest to theoretical ecologists because many potentially interesting results are buried in the vast NT literature. We propose to make these more accessible by extracting them and presenting them in a logical fashion. The focus of this review is broader than NT: we also discuss new, more recent approaches for studying ecological systems and how one might introduce realistic non-neutral models.

CONTENTS

I. Introduction	5
BOX 1: Macro-ecological patterns	11
BOX 2: Glossary	13
II. Neutral theory at stationarity	15
A. Markovian modeling of neutral ecological communities	17
1. The Master Equation of birth and death	17
2. Physics Ensembles	20
3. Ecological Ensembles	21
B. Coalescent approach to Neutral Theory	28
III. Dynamics in neutral theory	29
A. The continuum limit of the Master Equation	32
B. Boundary conditions	37
C. Stationary and time dependent solutions	38
D. The species turnover distribution	40
E. Persistence or lifetime distributions	41
1. Discrete Population Dynamics	42
2. Continuum limit	43
IV. Neutral spatial models and environmental fragmentation	44
A. Phenomenological models	47
B. Spatial Stochastic Processes	49
1. Spatially implicit model	49
2. Spatially explicit model	50
C. Environmental Fragmentation and Habitat Loss	53
V. Beyond neutrality.	55
A. Reconciling Neutral and Niche Theory	55
B. Emergent neutrality	57
C. Maximum Entropy Models	58

VI. Open problems and future challenges	60
A. Analytical spatially explicit models.	60
B. Linking different Macro-Ecological Patterns	61
C. Environmental noise	62
D. Multi-Trophic Ecosystems: Ecological Networks	63
VII. Conclusions	63
VIII. Acknowledgments	64
Appendix A	64
A. Density dependence, and assembling of local communities.	64
Appendix B	66
B. Persistence or lifetime distributions	66
Appendix C	68
C. The species turnover distribution with absorbing boundary conditions	68
References	69
Figures	77

I. INTRODUCTION

It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other in so complex a manner, have been all produced by laws acting around us. In this celebrated text from the Origin of Species (Darwin, 1909), Darwin eloquently conveys his amazement for the underlying laws of Nature: despite the striking diversity of shapes and forms, it exhibits deep commonalities that have emerged over wide scales of space, time and organizational complexity. For more than fifty years now, ecologists have collected census data for several ecosystems around the world from diverse communities such as tropical forests, coral reefs, plankton, etc. However, despite the contrasting biological and environmental conditions in these ecological communities, some macro-ecological patterns can be detected that reflect strikingly similar characteristics in very different communities (see BOX 1). This suggests that there are ecological mechanisms that are insensitive to the details of the systems and that can structure general patterns. Although the biological properties of individual species and their interactions retain their importance in many respects, it is likely that the processes that generate such macro-ecological patterns are common to a variety of ecosystems and they can therefore be considered to be universal. The question then is to understand how these patterns arise from just a few simple key features shared by all ecosystems. Contrary to inanimate matter, living organisms adapt and evolve through the key elements of inheritance, mutation and selection.

This fascinating intellectual challenge fits perfectly into the way physicists approach scientific problems and their style of inquiry. Statistical physics and thermodynamics have taught us an important lesson, that not all microscopic ingredients are equally important if a macroscopic description is all one desires. Consider for example a simple system like a gas. In the case of an ideal gas, the assumptions are that the molecules behave as point-like particles that do not interact and that only exchange energy with the walls of the container in which they are kept at a given temperature. Despite its vast simplifications, the theory yields amazingly accurate predictions of a multitude of phenomena, at least in a low-density regime and/or at not too low temperatures. Just as statistical mechanics provides a framework to relate the microscopic properties of individual atoms and molecules to the

macroscopic or bulk properties of materials, ecology needs a theory to relate key biological properties at the individual scale, with macro-ecological properties at the community scale. Nevertheless, this step is more than a mere generalization of the standard statistical mechanics approach. Indeed, in contrast to inanimate matter, for which particles have a given identity with known interactions that are always at play, in ecosystems we deal with entities that evolve, mutate and change, and that can turn on or off as well as tune their interactions with partners. Thus the problem at the core of the statistical physics of ecological systems is to identify the key elements one needs to incorporate in models in order to reproduce the known emergent patterns and eventually discover new ones.

Historically, the first models defining the dynamics of interacting ecological species were those of Lotka & Volterra, which describe asymmetrical interactions between predator-prey or resource-consumers systems. The Lotka & Volterra equations have provided much theoretical guidance. For instance, MacArthur developed a model for studying interactions among consumers which exploit common resources (MacArthur, 1970). By making use of different time scales, he showed how resources can be included in the Lotka & Volterra equations. He also derived the formulæ for the competition matrix, which provided hints about how it can be measured empirically (MacArthur, 1970). Seemingly, this suggested a viable way to measure niche overlap between species. Soon after, May and MacArthur studied the problem of how similar competing species can be and yet coexist in a community (May and Mac Arthur, 1972). According to Gause's competitive exclusion principle (Gause, 1934), two species cannot occupy the same niche in the same environment for a long time (see BOX 2). They found that environmental fluctuations limit niche overlap and therefore species' similarity. All these studies, and many other variations and generalizations, provided a robust theoretical basis for understanding ecosystems. Their limitations were identified as more research was carried out (Roughgarden *et al.*, 1989). These approaches share the idea that species can be well described by deterministic models which are shaped by the fundamental concept of niche, which, however, can be appropriately defined only *a-posteriori*. This theoretical focus is not eminently suited for studying a wide range of empirical patterns. In fact, all such models have several drawbacks: 1) They are mostly deterministic models and often do not take into account stochastic effects in the demographic dynamics (May, 2001); 2) As the number of species in the system increases, they become analytically intractable and computationally expensive; 3) They have a lot of parameters that are difficult to estimate

from ecological data or experiments; 4) It is very difficult to draw generalizations that include spatial degrees of freedom; and 5) While time series of abundance are easily analyzed, it remains challenging to study analytically the macroecological patterns they generate and thus, their universal properties.

A pioneering attempt to explain macro-ecological patterns as a dynamic equilibrium of basic and universal ecological processes - and that also implicitly introduced the concept of neutrality in ecology - was made by MacArthur and Wilson in the famous monograph of 1967 titled “The theory of island biogeography” (MacArthur and Wilson, 1967). In this work, the authors proposed that the number of species present on an island (and forming a local community) changes as the result of two opposing forces: on the one hand, species not yet present on the island can reach the island from the mainland (where there is a meta-community); and on the other hand, the species already present on the island may become extinct. MacArthur and Wilson’s model implies a radical departure from the then main current of thought among contemporary ecologists for at least three reasons: 1) Their theory stresses that demographic and environmental stochasticity can play a role in structuring the community as part of the classical principle of competitive exclusion; 2) The number of coexisting species is the result of a dynamic balance between the rates of immigration and extinction; 3) No matter which species contribute to this dynamic balance between immigration and extinction on the island, all the species are treated as identical. Therefore, they introduced a model that is neutral at the level of species (see BOX 2), even though they did not think of it as such.

Just a few years later, the American ecologist, H. Caswell, proposed a model in which the species in a community are essentially a collection of non-interacting entities and their abundance is driven solely by random migration/immigration. In contrast to the mainstream vision of niche community assembly, where species persist in the community because they adapt to the habitat, Caswell stressed the importance of random dispersal in shaping ecological communities. Although the model was unable to correctly describe the empirical trends observed in a real ecosystem, it is important because it pictured ecosystems as an open system, within which various species have come together by chance, past history and random diffusion.

In 2001, greatly inspired by the theory of island biogeography and the dispersal limitation concept (see BOX 2), Hubbell published an influential monograph titled “The Unified

Neutral Theory of Biodiversity and Biogeography” (however, the debate dates back to 1979 (Hubbell, 1979)). Unlike the niche theory and the approach adopted by Lotka & Volterra, the neutral theory (NT) aims to only model species on the same trophic level (monotrophic communities, see BOX 2), species that therefore compete with each other because they all feed on the same pool of limited resources. For instance, competition arises among plant species in a forest because all of them place demands on similar resources like carbon, light or nitrate. Other examples include species of corals, bees, hoverflies, butterflies, birds and so on. The NT is an ecological theory within which organisms of a community have identical per capita probabilities of giving birth, dying, migrating and speciating, regardless of the species they belong to. Thus, from an ecological point of view, the originality of Hubbell’s NT lies in the combination of several factors: *i*) it assumes competitive equivalence among interacting species; *ii*) it is an individual-based stochastic theory founded on mechanistic assumptions about the processes controlling the origin and interaction of biological populations at the individual level (i.e. speciation, birth, death and migration); *iii*) it can be formulated as a dispersal limited sampling theory; *iv*) it is able to describe several macro-ecological patterns through just a few fundamental ecological processes, such as birth, death and migration (Bell, 2000; Butler and Goldenfeld, 2009; Chave, 2004; Hubbell, 2001). Although the theory has been highly criticized by many ecologists as being unrealistic (Clark, 2009; Leigh, 2007; McGill, 2003; Nee, 2005; Ricklefs, 2006; Ricklefs and Renner, 2012), it does provide very good results when describing observed ecological patterns and it is simple enough to allow analytical treatment (Hubbell, 2005; Muneeppeerakul *et al.*, 2008; Volkov *et al.*, 2005, 2003, 2007). However, such precision does not necessarily imply that communities are truly neutral and indeed, non-neutral models can also produce similar patterns (Adler *et al.*, 2007; Du *et al.*, 2011). Yet the NT does call into question approaches that are either more complex or equally unrealistic (Noble *et al.*, 2011b; Purves and Turnbull, 2010). Moreover, NT is not only a useful tool to reveal universal patterns but also it is a framework that provides valuable information when it fails. One of the strengths of NT is that one can, in fact, falsify one or more of its assumptions, and thereby actually test the theory. Few models in community ecology meet this gold standard. These features have made NT an important approach in the study of biodiversity (Alonso *et al.*, 2006; Black and McKane, 2012; Chave, 2004; Hubbell, 2006; J.Harte, 2003; Rosindell *et al.*, 2011, 2012; Walker, 2007).

From a physicist’s perspective, NT is appealing as it represents a sort of “thermodynamic” theory of ecosystems. Similar to the kinetic theory of ideal gases in physics, NT is a basic theory that provides the essential ingredients to further explore theories that involve more complex assumptions. Indeed, NT captures the fundamental approach of physicists, which can be summarized by Einstein’s celebrated quote “Make everything as simple as possible, but not simpler”. Finally, it should be noted that the NT of biodiversity is basically the analogue of the theory of neutral evolution in population genetics (Kimura, 1985) and indeed, several results obtained in population genetics can be mapped to the corresponding ecological case (Blythe and McKane, 2007).

Statistical physics is contributing decisively to our understanding of biological and ecological systems by providing powerful theoretical tools and innovative steps (Goldenfeld and Woese, 2007, 2010) to understand empirical data about emerging patterns of biodiversity. The aim of this review is not to present a complete and exhaustive summary of all the contributions to this field in recent years - a goal that would be almost impossible in such an active and broad interdisciplinary field - but rather, we would like to introduce this exciting new field to physicists that have no background in ecology and yet are interested in learning about NT. Thus, we will focus on what has already been done and what issues must be addressed most urgently in this nascent field, that of statistical physics applied to ecological systems. A nice feature of this field is the availability of ecological data that can be used to falsify models and highlight their limitations. At the same time, we will see how the development of a quantitative theoretical framework will enable one to better understand the multiplicity of empirical experiments and ecological data.

This review is organized into six main sections. Sec. II is an attempt to review several important results that have been obtained by solving neutral models at stationarity. In particular, we will present the theoretical framework based on Markovian assumptions to model ecological communities, where different models may be seen as the results of different NT ensembles. We will also show how NT, despite its simplicity, can describe patterns observed in real ecosystems. In Sec. III, we will present more recent results on dynamic quantities related to NT. In particular, we will discuss the continuum limit approximation of the discrete Markovian framework, paying special attention to boundary conditions, a subtle aspect of the time dependent solution of the NT. In Sec. IV, we will provide examples of how space plays an essential role in shaping the organization of an ecosystem. We will discuss

both phenomenological, and spatially implicit and explicit NT models. A final subsection will be devoted to the modeling of environmental fragmentation and habitat loss. In Secs. V and VI we will propose some emerging topics in this fledgling field, and present the problems currently being faced. Finally, we will close the review with a section dedicated to conclusions.

BOX 1: Macro-ecological patterns

- **Alpha-diversity:** the number of species found in a given area, regardless of their abundance and spatial distributions. Sometimes it is more appropriate to assess the diversity of a local community by taking into account their abundances as well. For this purpose two alternative indexes have been used, the Shannon (H) and Simpson (D) index (Simpson, 1949)
- **Beta-diversity:** the probability that two individuals at distance r belong to the same species (i.e., they are conspecific). Several alternative definitions have been used in the literature, including the pair (or two-point) correlation function (PCF), the Sørensen similarity index (SSI) or the Jaccard similarity index (JSI). However, the general purpose of beta-diversity is to describe the turnover of multiple species in space.
- **Pair Correlation Function (PCF):** the correlation in species' abundance between pairs of samples as a function of their distance.
- **Relative Species Abundance (RSA):** the probability that a species has n individuals in a given region. When multiplied by the total number of species in the region, this gives the number of species with n individuals (see section II). This is sometimes called the Species Abundance Distribution (SAD).
- **Species (time) Turnover Distribution (STD):** the probability density function that the ratio of the future to the current population sizes of any species has a value λ in a given ecological community (see section III).
- **Persistence or lifetime distribution (SPT):** the probability density function of the time interval between the emergence and local extinction of any species within a given area (see section III).
- **Species Area Relationship (SAR):** the function that relates the mean number of species, S , to the area, A , they live in. On a relatively large range of spatial scales it is well approximated by a power law, $S(A) = cA^z$ where $0 < z < 1$ (typical values of z are 0.2 or 0.3). It has been shown that the SAR in the log-log

scale has three qualitatively different behaviours from local to continental spatial scales: approximately linear-like at very small and very large scales, and power-law-like for intermediate scales. This is referred to as a triphasic SAR (see section IV).

- **Endemic Area Relationship (EAR):** The mean number of species that are present in the area A but not outside it (see section IV).

BOX 2: Glossary

- **Trophic level:** the set of all species belonging to the same level in the food chain. Individuals of species belonging to higher levels feed upon those in the lower ones, while individuals belonging to the same trophic level compete for the same pool of resources. Neutral theory is an ecological theory for species in one specific trophic level, whereas Niche theory (see below) can also deal with species at different trophic levels.
- **Gause's competitive exclusion principle:** a pair of species cannot stably co-exist if they feed upon exactly the same resources under the same environmental conditions (Gause, 1934).
- **Niche theory:** species can stably coexist in an ecological community if their characteristics (or traits) allow them to specialize on one particular set of resources or environment conditions (niches) in which they are superior to their competitors. In other words, they occupy different niches (Hutchinson, 1957). Such niche separation is deemed to enhance trade-offs and facilitate co-existence, even though there is no *a priori* method to identify the correct niches that favor co-existence. The underlying rationale of the theory is Gause's exclusion principle, and the classical model is Lotka-Volterra's set of differential equations. In contrast to Niche Theory, Neutral theory claims that niche differences are not essential to co-existence.
- **Species-level neutral models:** these models assume that all species are equivalent as they all have the same probability of immigration, extinction and speciation. The only state variable of these models is the number of species in a community and thus, it cannot predict the distribution of the population sizes across species (see RSA below) (MacArthur and Wilson, 1967).
- **Individual-level neutral models:** these models assume that all species are equivalent at the individual level, having the same birth, death, immigration and speciation rates regardless of their identity. Therefore, although species can have different abundances, they are competitively equivalent. The state variable of these models is the population of any species in a region and therefore, such models can be used to understand species

richness and abundance (Hubbell, 2001).

- **Symmetric models:** any model whose outcomes are invariant when exchanging species identities. The family of symmetric models is larger than the neutral ones: for instance, effects of environmental stochasticity or density dependence of individuals on the per capita birth and death rates can be accommodated in symmetric – but not neutral – models. Some authors do not make a distinction between symmetric and neutral models.
- **Dispersal limited process:** any process that constrains offspring to disperse in the vicinity of its parents.
- **Multispecies Voter Model with Speciation (MVM):** a spatially explicit neutral model in which each individual is located in a regular lattice and belongs to a species. Given a spatial configuration at time t , the configuration at time $t + 1$ is obtained as follows: an individual is chosen at random and is replaced by a copy of one of its nearest neighbours (chosen at random) with a probability $1 - \nu$, or by an individual belonging to a new species (not already present in the lattice) with probability ν . ν is the speciation parameter, although it may incorporate immigration effects.

II. NEUTRAL THEORY AT STATIONARITY

Neutral theory deals with ecological communities within a single trophic level, i.e. communities whose species compete for the same pool of resources (see BOX 2). This means that neutral models will generally be tested on data describing species that occupy the same position in the food chain, like trees in a forest, breeding birds in a given region, butterflies in a landscape, plankton, etc.. Therefore, ecological food webs with predator-prey type interactions are not suitable to be studied with standard neutral models.

As explained in the introduction, ecologists have been studying an array of biodiversity descriptors over the last sixty years (see BOX 1), including relative species abundance distributions (RSA), species-area relationships (SAR) and spatial pair correlation function (PCF). For instance, the RSA represents one of the most commonly used static measures to summarize information on ecosystem diversity. The analysis of this pattern reveals that the RSA distributions in tropical forests share similar shapes, regardless of the type of ecosystem, geographical location or the details of species interactions (see Fig. 2). Therefore, the functional form of the RSA (see seminal papers by Fisher (Fisher *et al.*, 1943) and Preston (Preston, 1948) for a theoretical explanation of its origins) has been one of the great problems studied by ecologists. Indeed, a great deal of attention has been devoted to the precise functional forms of these patterns. A meta-analysis revealed that RSA distributions have basically three shapes: more often unimodal (lognormal-like) in fully sampled communities, and either power-law or without any mode (logseries-like) within incompletely censused regions (Ulrich *et al.*, 2010). Also, it is difficult to tease apart, from the RSA alone, the nature of the basic processes driving communities (Pueyo, 2006), because neutral and non-neutral mechanisms coexist in nature.

It is instructive to derive some of these functional forms by starting with a very simple but extreme neutral model, which assumes that species are independent and randomly distributed in space. This null model tells us what we should expect when the observed macroecological patterns are only driven by randomness, with no underlying ecological mechanisms. If the density of individuals in a very large region is ρ , then the probability that a species has n individuals within an area a is well approximated by a Poisson distribution with a mean ρa . As defined in Box 1, this is the RSA for the area a . However, the empirical data are not well described by a Poisson distribution and as we shall see later on, better fits are

usually given by log-series, gamma or log-normal distributions.

The SAR curve can also be calculated within a slightly more accurate model, which still assumes that species are independent and randomly situated in space. Let us now suppose that a region with area A_0 contains S_{tot} species in total (α -diversity - see Box 1), and that the species i has n_i individuals in A_0 . If we consider a smaller area A within the region, then the probability that an individual will not be found in such area is $1 - A/A_0$, while the probability that the whole species i is not present therein is $(1 - A/A_0)^{n_i} = 1 - p_i$. If we now consider the random variable $I_i(A)$, which is 1 when the species i is found within the area A and 0 if not, then $\langle I_i(A) \rangle = p_i = 1 - (1 - A/A_0)^{n_i}$, because $I_i(A)$ is a Bernoulli random variable with an expectation value p_i . Therefore, the mean number of species in the area A (i.e. the SAR) is simply $S(A) = \sum_i \langle I_i(A) \rangle$, which is

$$S(A) = S_{tot} - \sum_{i=1}^{S_{tot}} (1 - A/A_0)^{n_i}. \quad (1)$$

Although this model was originally studied by Coleman (Coleman, 1981), we now know that it significantly overestimates species diversity at almost all spatial scales (Plotkin *et al.*, 2000).

Beta-diversity (see Box 1) can be estimated under the assumptions we have mentioned. Now, regardless of the spatial distance between two individuals, the probability that two of them belong to the same species i is $n_i(n_i - 1)/[N(N - 1)]$, where N is the total number of individuals in the community, i.e. $N = \sum_i n_i$. Therefore, the probability to find any pair of con-specific individuals is $(1/S^{tot}) \sum_{i=1}^{S^{tot}} n_i(n_i - 1)/[N(N - 1)]$. This means that the random placement model with independent species predicts that beta-diversity should not depend on the distance between two individuals. Again, we now have clear evidence that the probability that two individuals at distance r belong to the same species is a decaying function of r (Morlon *et al.*, 2008).

The failure of the random placement model to capture the RSA, SAR and beta-diversity is a clear indication that ecological patterns are driven by non-trivial mechanisms that need to be appropriately identified. Thus, we shall assess to what extent the NT at stationarity can provide predictions in agreement with empirical data.

There are two related, but distinct analytical frameworks that have been used to mathematically formulate the NT of biodiversity at stationarity for both local and meta-

communities: the first fixes the total population, whereas the second fixes the average total population of a community. From an ecological point of view, a local community is defined as a group of potentially interacting species sharing the same environment and resources. Mathematically, when modeling a local community the total community population abundance remains fixed. Alternatively, a meta-community can be considered a set of interacting communities that are linked by dispersal and migration phenomena. In this case, it is the average total abundance of the whole meta-community that is held constant. From the physical point of view, these roughly correspond to the micro/canonical (fixed total abundance) and the grand canonical (fixed average total abundance) ensembles, respectively. The micro/canonical ensemble or so-called zero sum dynamics when death and birth events always occur as a pair originates from the sampling frameworks in population genetics pioneered by Warren Ewens and Ronald Fisher (Fisher *et al.*, 1943). It should be noted that even though a fixed size sample is one way to analyze available data, for the majority of cases (apart from very small size samples), the grand canonical ensemble approach is that used routinely in statistical physics and it provides a very precise yet largely simplified description of the system. The ultimate reason for this lies in the surprising accuracy of the asymptotic expansion of the gamma function (the mathematical framework heavily uses combinatorials and factorials etc.). The Stirling approximation can be used for very large values of the gamma function, nevertheless, it is quite accurate even for values of the arguments of the order of 20. The advantages of the master equation (see below) and the grand canonical ensemble approach stem from their computational simplicity, which make the results more intuitively transparent.

We now introduce the mathematical tools of stochastic processes that will be used extensively in the rest of the article.

A. Markovian modeling of neutral ecological communities

1. The Master Equation of birth and death

Let c be a configuration of an ecosystem that could be as detailed as the characteristics of all individuals in the ecosystem, including their spatial locations or as minimal as the abundance of a specified subset of species. Let $P(c, t|c_0, t_0)$ be the probability that a

configuration c is seen at time t , given that the configuration at time t_0 was c_0 (referred to as $P(c, t)$ for simplicity). For our applications, the configurations c are typically species abundance (denoted by n).

Assuming that the stochastic dynamics are Markovian, the time evolution of $P(c, t)$ is given by the Master Equation (ME) (Gardiner, 2009; Kampen, 2007)

$$\frac{\partial P(c, t)}{\partial t} = \sum_{c'} \left(T[c|c']P(c', t) - T[c'|c]P(c, t) \right), \quad (2)$$

where $T[c|c']$ is the transition rate from the configuration c' to configuration c . Under suitable and very plausible conditions (Kampen, 2007), $P(c, t)$ approaches a stationary solution at long times, $P_s(c)$, which satisfies the following equation

$$\sum_{c'} \left(T[c|c']P_s(c') - T[c'|c]P_s(c) \right) = 0. \quad (3)$$

This equation is typically intractable with analytical tools, because it involves a sum of all the configurations. If each term in the summation is zero, i.e.

$$T[c|c']P_s(c') - T[c'|c]P_s(c) = 0, \quad (4)$$

detailed balance is said to hold. A necessary and sufficient condition for the validity of detailed balance is that for *all* possible cycles in the configuration space, the probability of walking through it in one direction is equal to the probability of walking through it in the opposite direction. Given a cycle $\{c_1, c_2, \dots, c_{n-1}, c_1\}$, detailed balance holds if and only if, for every such cycle,

$$T[c_1|c_2]T[c_2|c_3] \dots T[c_{n-2}|c_{n-1}]T[c_{n-1}|c_1] = T[c_1|c_{n-1}]T[c_{n-1}|c_{n-2}] \dots T[c_3|c_2]T[c_2|c_1]. \quad (5)$$

This condition evidently corresponds to a time-reversible condition.

Now let us apply these mathematical tools to the study of community dynamics that is driven by random demographical drift. Consider a well-mixed local community. This is equivalent to saying that the distribution of species in space is not relevant, which should hold for an ecosystem with a linear size smaller than, or of the same order as the seed dispersal range. In this case one can use $c = (n_1, \dots, n_S) = \mathbf{n}$ where n_i is the population of the i -th species and the system contains S species. We can rewrite eq. 2 in the following way:

$$\frac{\partial P(\mathbf{n}, t)}{\partial t} = \sum_{\mathbf{n}' \neq \mathbf{n}} T[\mathbf{n}|\mathbf{n}']P(\mathbf{n}', t) - \sum_{\mathbf{n}' \neq \mathbf{n}} T[\mathbf{n}'|\mathbf{n}]P(\mathbf{n}, t). \quad (6)$$

In this case, $T[\mathbf{n}'|\mathbf{n}]$ can take into account birth and death, as well as immigration from a meta-community. The simplest hypothesis is that $T[\mathbf{n}'|\mathbf{n}]$ is the result of S elementary birth and death processes that occur independently for each of the S species i.e.

$$T[\mathbf{n}'|\mathbf{n}] = \sum_{k=1}^S \prod_{i \neq k} \delta_{n'_i, n_k} (\delta_{n'_k, n_k+1} T_k(n+1|n) + \delta_{n'_k, n_k-1} T_k(n-1|n)), \quad (7)$$

where $\delta_{k',k}$ is a Kronecker delta,

$$T_k(n+1|n) = b(n, k) \quad (8)$$

is the birth rate and

$$T_k(n-1|n) = d(n, k), \quad (9)$$

the death rate. This particular choice corresponds to a sort of mean-field (ME) approach (Alonso and McKane, 2004; McKane *et al.*, 2004; Pigolotti *et al.*, 2004; Vallade and Houchmandzadeh, 2003; Volkov *et al.*, 2003, 2007; Zillio *et al.*, 2008).

Our many-body ecological system can also be formulated in a language more familiar to statistical physicists, where we consider the distribution of balls into boxes. The “boxes” are the species and the “balls” are the individuals. Birth/death processes correspond to adding/removing a ball to/from one of the boxes using a rule as dictated by Eqs. (8) and (9). Eq. (6) with the choice (7) can be simplified if one assumes that the initial condition is factorized as $P(\mathbf{n}, t=0) = \prod_{k=1}^S P_k(n_k, t=0)$. In this case the solution is again factorized as $P(\mathbf{n}, t) = \prod_{k=1}^S P_k(n_k, t)$ where each P_k satisfies the following ME in one degree of freedom n_k

$$\frac{\partial P_k(n_k, t)}{\partial t} = P_k(n_k-1, t)b(n_k-1, k) + P_k(n_k+1, t)d(n_k+1, k) - (b(n_k, k) + d(n_k, k)) P_k(n_k, t). \quad (10)$$

The stationary solution of eq.(10) is easily seen to satisfy detailed balance (Gardiner, 2009; Kampen, 2007).

First, we note that because of the neutrality hypothesis, species are assumed to be demographically identical and therefore, we can drop the k dependence factor from the equations (8)-(9). In other words, we can concentrate on the probability $P(n)$ that a given species (box) has an abundance n (balls). In this case, species do not interact and thus, in our calculation we can follow a particular species (the boxes are taken to be independent). Following equation (4), it is easy to see that the solution of the birth-death ME (2) that satisfies

the detailed balance condition and that thus corresponds to equilibrium, is (Gardiner, 2009; Kampen, 2007)

$$P_s(n) = P_0 \prod_{z=1}^n \frac{b(z-1)}{d(z)} \quad (11)$$

where P_0 can be calculated by the normalization condition $\sum_n P_s(n) = 1$, and assuming that all the rates are positive. In particular, $b(0) > 0$ and $d(0) = 0$.

When there are S boxes, all satisfying the same birth-death rules, the general and unique equilibrium solution is

$$P_s(n_1, n_2, \dots, n_S) = \prod_{k=1}^S P_s(n_k) \quad (12)$$

Depending on the functional form of $b(n)$ and $d(n)$, one can readily work out the desired $P_s(\mathbf{n})$. We will start with some cases that are familiar to physicists (Volkov *et al.*, 2006), and then move onto more ecologically meaningful cases (Volkov *et al.*, 2003, 2007).

2. Physics Ensembles

The Random walk and Bose-Einstein Distribution. If one chooses $b(n) = b_0$ and $d(n) = d_1$ for $n \geq 1$, and $d(n) = 0$ otherwise (or alternatively, $b(n) = (n+1)b_0$ and $d(n) = d_1 n$), one obtains a pure exponential distribution $P(n) = r^n(1-r)$ where $r = b_0/d_1$. Note that in both these cases, $b(n-1)/d(n)$ is the same. Substituting this in equation (12) gives us the well known Bose-Einstein distribution for non-degenerate energy levels, i.e.

$$P(n_1, n_2, \dots, n_S) = r^N(1-r)^S, \quad (13)$$

where $N = \sum_k n_k$. Here, r corresponds to $e^{-\beta(\epsilon-\mu)}$ in the grand canonical ensemble.

The Fermi-Dirac Distribution. If $d(n) = d_1 n$ and $b(n) = 0$ for any n other than 0, and equal to b_0 otherwise, then $P(n) = r^n/(1+r)$ for $n = 0$ or $n = 1$ and $P(n) = 0$ for other values of n , and accordingly the Fermi-Dirac distribution is achieved

$$P(n_1, n_2, \dots, n_S) = r^{\sum_k n_k} (1+r)^{-S} \quad \text{for } n_k = 0 \text{ or } n_k = 1 \quad \forall i \quad (14)$$

$$P(n_1, n_2, \dots, n_S) = 0 \quad \text{otherwise.} \quad (15)$$

Boltzmann counting If $d(n) = d_1 n$ and $b_n = b_0$ for all n , then one obtains a Poisson

distribution $P(n) = e^{-r}r^n/n!$, and this leads to

$$P(n_1, n_2, \dots, n_S) \propto \frac{r^N}{\prod_{k=1}^S n_k!}. \quad (16)$$

This is the familiar grand-canonical ensemble Boltzmann counting in physics, where r plays the role of fugacity. It is noteworthy that, unlike the conventional classical treatment (Huang, 2001), where an additional factor of $N!$ is obtained, here one gets the correct Boltzmann counting and thereby avoids the well-known Gibbs paradox (Huang, 2001). Thus, if one were to ascribe energy values to each of the boxes and enforce a fixed average total energy, the standard Boltzmann result would be obtained whereby the probability of occupancy of an energy level ε is proportional to $e^{-\beta\varepsilon}$, where β is proportional to the inverse of the temperature.

3. Ecological Ensembles

Density independent dynamics. We now consider the dynamic rules of birth, death and speciation that govern the population of an individual species. The most simple ecologically meaningful case is to consider $d(n) = dn$ and $b(n) = bn$ for $n > 0$, and $b(0) > 0$, $d(0) = b(-1) = 0$. We now define $r = b/d$. Moreover, in order to ensure that the community will not become extinct at longer times, speciation may be introduced by ascribing a non-zero probability of the appearance of an individual from a new species, i.e. $b(0) = b_0 = \nu$. In this case the probability for a species of having n individuals at stationarity is:

$$P(n) = \tilde{\nu}[1 - \tilde{\nu} \ln(1 - r)]^{-1}r^n/n, \quad (17)$$

where $\tilde{\nu} = \nu/b$ and $n > 0$ and $P(0) = 1/(1 - \tilde{\nu} \ln(1 - r))$.

The RSA $\langle\phi(n)\rangle$ is the average number of species with a population n , and this is simply (Volkov *et al.*, 2003)

$$\langle\phi(n)\rangle = SP(n) = \theta r^n/n. \quad (18)$$

This is the celebrated Fisher log-series distribution, i.e. the distribution Fisher proposed as that describing the empirical RSA in real ecosystems (Fisher *et al.*, 1943). The parameter $\theta = S\tilde{\nu}/[1 - \tilde{\nu} \ln(1 - r)]$ is known as the Fisher number or biodiversity parameter.

In 1948, another great ecologist of the inductive school, F.W. Preston, published a paper (Preston, 1948) challenging Fisher's point of view. He showed that the Log-series is not

a good description for the data from a large sample of birds. In fact, he observed an internal mode in the RSA that was absent in a Log-series distribution. In particular, Preston introduced a way to plot the experimental RSA data by octave abundance classes (i.e. $[2^k, 2^{k+1}]$, for $k = 1, 2, 3, \dots$), showing that a good fit of the data was represented by a Log-Normal distribution. Indeed, there are several examples of RSA data that display this internal mode feature (see Fig. 2). The intuition of Preston was that the shape of the RSA must depend on the sampling intensity or size of the community. Conventionally, when studying ecological communities, ecologists separate them into two distinct classes: small local communities (e.g., on a island) and meta-communities of much larger communities or those composed of several smaller local communities (see Fig. 3). The neutral modeling schemes for these two cases - that we will denote by the sub-script L and M respectively - are not the same as the ecological processes involved differ. In fact, the immigration rate (m) in a local community is a crucial parameter as the community is mainly structured by dispersal limited mechanisms, and the speciation rate (ν) can be neglected. On the other hand, in a meta-community immigration does not occur (species colonize within the community) and the community is shaped by birth-death-competition processes, although speciation ν also plays an important role. A meta-community can also be thought of as consisting of many small semi-isolated local communities, each of which receives immigrants from other local communities. When considering meta-community dynamics, the natural choice is to put a soft constraint on J_M , i.e. the total number of individuals is free to fluctuate around the average population $\langle J_M \rangle$. Indeed, one finds that at the largest J_M limit, the results obtained with a hard constraint on J_M are equivalent to those with a soft constraint. On the other hand, when considering a local community, it is safer to place a hard constraint on the total population J_L . In both these cases, S represents the total number of species that may potentially be present in the community, while the average number of species observed in the community is denoted by $\langle S \rangle$.

There are several ecological meaningful mechanisms that can generate a bell shaped Preston-like RSA. The first of these involves density dependent effects on birth and death rates. The second involves considering a Fisher log-series as the RSA of a meta-community acting as a source of immigrants to a local community embedded within it. The dynamics of the local community are governed by births, deaths and immigration, whereas the meta-community is characterized by births, deaths and speciation. This leads to a local

community RSA with an internal mode (McKane *et al.*, 2000; Vallade and Houchmandzadeh, 2003; Volkov *et al.*, 2003, 2007). A third way is to incrementally aggregate several local communities (see Appendix A)

Local Dynamics with Density Dependent Rates.

One of the most fundamental and long debated questions in community ecology is to understand what maintains species diversity within communities (Chave *et al.*, 2002; Hutchinson, 1959). Mechanisms which are able to promote diversity include competitive trade-offs among species because of species-specific traits, balance between speciation and extinction, frequency- or density-dependence and environmental variability. For instance, processes that hold the abundance of a common species in check inevitably lead to rare-species advantages, given that the space or resources freed up by density-dependent death can be exploited by less-common species. Therefore, inter-species frequency dependence is the community-level consequence of intra-species density dependence, and thus, they may be thought of as two different manifestations of the same phenomenon (Volkov *et al.*, 2005). Because these processes are capable of driving community-level patterns such as SAR or RSA, one might hope to identify mechanisms by delving into observed patterns. However, communities with different governing processes can unexpectedly show similar patterns, thus suggesting that their shape cannot be, in general, used to identify specific mechanisms (Chave *et al.*, 2002; Purves *et al.*, 2005).

In this review we focus on two of the most prominent hypotheses which explain species coexistence through frequency and density dependence: the Janzen-Connell (Connell, 1971; Janzen, 1970) and the Chesson-Warner hypotheses (Chesson and Warner, 1981). These mechanisms generally predict the reproductive advantage of a rare species due to ecological factors and they can be readily captured in a common mathematical framework that will be presented below.

The Janzen-Connell hypothesis postulates that host-specific pathogens or predators act in the vicinity of the maternal parent. Thus, seeds that disperse further away from the mother are more likely to escape mortality. This spatially structured mortality effect suppresses the uncontrolled population growth of locally abundant species relative to uncommon species, thereby producing a reproductive advantage to a rare species. The Chesson-Warner storage hypothesis explores the consequences of a variable external environment and it relies on

three empirically validated observations: species respond in a species-specific manner to the fluctuating environment; there is a covariance between the environment, and intra- and inter-species competition; life history stages buffer the growth of population against unfavorable conditions. Such conditions prevail when species have similar per capita rates of mortality but they reproduce asynchronously and there are overlapping generations.

We begin by noting that the mean number of species with n individuals, $\langle \phi_n \rangle$, is not determined by the absolute rates of birth or death but rather, by their ratio, $r_{i,k} = \frac{b(i,k)}{d(i+1,k)}$. This follows from the observation that $\langle \phi_n \rangle$ is proportional to $\langle r_{1,k} r_{2,k} \dots r_{n-1,k} \rangle_k$, where the average $\langle \dots \rangle_k$ is obtained from all the species. This simple formulation (Volkov *et al.*, 2005) is sufficiently general to represent the communities of either symmetric species (in which all the species have the same demographic birth and death rates) or the case of asymmetric or distinct species. The more general asymmetric situation captures niche differences and/or differing immigration fluxes that might arise from the different relative abundances of distinct species in the surrounding meta-community (see BOX 2).

The density-dependence arising from the Janzen-Connell effect is spatially explicit – in plant communities at least – because it is caused by interactions among neighboring individuals. Here, instead we formulate a density-dependence which is spatially implicit and only abundance-driven. We therefore introduce a modified symmetric theory that incorporates rare species advantage or common species disadvantage by making r_n a decreasing function of abundance. The equations of density dependence in the per capita birth and death rates for an arbitrary species of abundance n are: $\frac{b(n)}{n} = b \cdot [1 + \frac{b_1}{n} + o(\frac{1}{n^2})]$ and $\frac{d(n)}{n} = d \cdot [1 + \frac{d_1}{n} + o(\frac{1}{n^2})]$, for $n > 0$ as the leading term of a power series in $(1/n)$, $\frac{b(n)}{n} = b \cdot \sum_{l=0}^{\infty} b_l n^{-l}$ and $\frac{d(n)}{n} = d \cdot \sum_{l=0}^{\infty} d_l n^{-l}$, where b_l and d_l are constants. This expansion captures the essence of density-dependence by ensuring that the per-capita rates decrease and approach a constant value for a large n , given that the higher order terms are negligible. As noted earlier, the quantity that controls the RSA distribution is the ratio b_n/d_{n+1} . Thus, the birth and death rates, b_n and d_n , can be defined up to $f(n+1)$ and $f(n)$ respectively, where f is any arbitrary well-behaved function.

Strikingly, any relative abundance data can be considered as arising from effective density dependent processes in which the birth and death rates are given by the above expressions. Thus, one would expect that the per capita birth rate or fecundity drops as the abundance increases, whereas mortality ought to increase with abundance. Indeed, the per capita death

rate can be arranged to be an increasing function of n , as observed in nature, by choosing an appropriate function f and appropriately adjusting the birth rate so that the ratio b_n/d_{n+1} remains the same. This then ensures that the RSA does not vary.

The mathematical formulation of density dependence may seem unusual to ecologists familiar with the logistic or Lotka-Volterra systems of equations, wherein density dependence is typically described as a polynomial expansion of powers of n truncated at the quadratic level. Such an expansion is valid when the characteristic scale of n is determined by a fixed carrying capacity. Conversely, here the range of n is from 1 to an arbitrarily large value and not to some carrying capacity. Therefore, an expansion in terms of powers of $(1/n)$ is more appropriate. For this symmetric model (Volkov *et al.*, 2005), and bearing in mind that $\langle \phi_n \rangle = SP_0 \prod_{i=1}^{n-1} \frac{b(i)}{d(i+1)}$, one readily arrives at the following relative species-abundance relationship:

$$\langle \phi_n \rangle = \theta \frac{x^n}{n+c}, \quad (19)$$

where $x = b/d$ and for parsimony, we make the simple assumption that $b_1 = d_1 = c$ and the higher order coefficients, $b_2, d_2, b_3, d_3, \dots$, are all 0. The biodiversity parameter, θ , is the normalization constant that ensures the total number of species in the community is S , and it is given by $\theta = S^{\frac{1+c}{cx}} / F(1+c, 2+c, x)$, where $F(1+c, 2+c, x)$ is the standard hypergeometric function. The parameter c measures the strength of the symmetric density- and frequency dependence in the community, and it controls the shape of the RSA distribution. This simple model (Volkov *et al.*, 2005) does a good job of matching the patterns of abundance distribution observed in the tropical forest communities throughout the world (see Fig. 2). Note that when $c \rightarrow 0$ (the case of no density dependence), one obtains the Fisher log-series. In this case θ captures the effects of speciation.

Local community with immigrants from a meta-community. Various analytical solutions for the Markovian (ME) approach have been suggested in the literature. Based on pioneering work by McKane, Alonso and Solé (McKane *et al.*, 2000), an analytical solution was proposed for the ME when the size of both the local community and the meta-community was fixed (McKane *et al.*, 2000; Vallade and Houchmandzadeh, 2003; Volkov *et al.*, 2003). Here, we will refer to the work of (Vallade and Houchmandzadeh, 2003), in whose model the birth

and death rates in the meta-community are given by

$$T(n+1|n) = b_M(n) = \frac{(J_M - n)}{J_M} \frac{n}{J_M - 1} (1 - \nu) \quad (20)$$

$$T(n-1|n) = d_M(n) = \frac{n[J_M - n + (n-1)\nu]}{J_M(J_M - 1)}. \quad (21)$$

During a unit of time, the population $(n \neq 0, J_M)$ of the species under consideration can increase by one individual if we uniformly pick at random an individual that does not belong to that species (this occurs with probability $(J_M - n)/J_M$), we substitute it with an individual randomly picked from the considered species (with probability $n/(J_M - 1)$) and, finally, there is no speciation (prob. $1 - \nu$). Multiplying all these terms, we get the birth rate in Eq. (20). On the other hand, in the same unit of time the population of the species under consideration can decrease by one individual if we uniformly pick at random an individual that belongs to that species (prob. n/J_M) and we substitute it either with a completely new individual (speciation with prob. ν) or with an individual which does not belong to the considered species (with probability $(J_M - n)/(J_M - 1)$). These terms provide the death rate in Eq. (21).

Let $\langle \phi_n(t) \rangle_M$ designate the average number of species of abundance n in the meta-community. Then, in the stationary limit ($t \rightarrow \infty$), the species that contribute to $\langle \phi_n(t) \rangle_M$ are those which entered the system by mutation at time $t - \tau$ and that have reached size n at time t (Suweis *et al.*, 2012b; Vallade and Houchmandzadeh, 2003). As speciation is a Poissonian event (of rate ν) and due to neutrality, all species have the same probability $p(\tau)d\tau = \nu d\tau$ of appearing between the time interval $[\tau, \tau + d\tau]$. Thus, the time evolution of $\langle \phi_n(t) \rangle$ is then simply given by $\langle \phi_n(t) \rangle = \int_0^t P(n, t - \tau) p(\tau) d\tau = \nu \int_0^t P(n, \tau) d\tau$, where $P(n, t)$ is the solution of the ME (6) with transition rates given by equations (20)-(21) and with the initial condition $P(n, 0) = \delta_{n,1}$ (δ represents the Kronecker delta). Using Laplace transforms, the stationary RSA for the meta-community of size J_M is (Vallade and Houchmandzadeh, 2003):

$$\langle \phi_k \rangle_M = \frac{\theta \Gamma(J_M + 1) \Gamma(J_M + \theta - k)}{k \Gamma(J_M + 1 - k) \Gamma(J_M + \theta)}, \quad (22)$$

where $\theta = (J_M - 1)\nu/(1 - \nu)$.

Therefore the density of species of relative abundance $\omega = k/J_M$ in the meta community can be written as:

$$g_M(\omega) = \frac{\theta(1 - \omega)^{\theta-1}}{\omega}, \quad (23)$$

where $g_M(\omega) = \lim_{J_M \rightarrow \infty} J_M \langle \phi_k \rangle_M$ and with ω and θ kept fixed (implying that ν tends to zero and k tends to infinity).

Let us now consider a local community (of size J_L) in contact with the aforementioned meta-community. As $J_M \gg J_L$, then mutations within the local community can be neglected ($\nu \rightarrow 0$) and the equilibrium distribution in the meta-community is negligibly modified by migration from or towards the community. Nevertheless, there is migration from the meta-community to the local community and thus, a migration rate m must be included in the transition rates for the local community dynamics:

$$b_L(n) = \frac{J_L - n}{J_L} \left[\frac{n}{J_L - 1} (1 - m) + m\omega \right] \quad (24)$$

$$d_L(n) = \frac{n}{J_L} \left[\frac{J_L - n}{J_L - 1} (1 - m) + m(1 - \omega) \right], \quad (25)$$

where $\omega = k/J_M$ is the relative abundance of the species in the meta-community and k is the abundance of the tracked species within the meta-community.

The stationary solution $P_n(\omega)$ of the corresponding ME can be calculated exactly using Eq. (11), considering m as the immigration rate (Vallade and Houchmandzadeh, 2003):

$$P_n(\omega) = \binom{J_L}{n} \frac{(\mu\omega)_n [\mu(1 - \omega)]_{J_L - n}}{(\mu)_{J_L}}, \quad (26)$$

where $\binom{J_L}{n}$ is the binomial coefficient, $(a)_n = \Gamma(a + n)/\Gamma(a)$ is the Pochhammer coefficient, and $\mu = (J_L - 1)m/(1 - m)$.

Finally, the average number of species with abundance n in the local community, $\langle \phi_n \rangle_L$, is given by the following expression:

$$\langle \phi_n \rangle_L = \sum_{k=1}^{J_M} P_n(k/J_M) \langle \phi_k \rangle_M, \quad (27)$$

which displays an internal mode for appropriate values of the model parameters. A generalization of this result to a system of two communities of arbitrary yet fixed sizes that are subject to both speciation and migration (Vallade and Houchmandzadeh, 2006) has also been carried out.

Alonso and McKane (Alonso and McKane, 2004) suggested that the log-series solution for the species abundance distribution in the meta-community is applicable only for species-rich communities, and that it does not adequately describe species-poor meta-communities. An application of the ME approach to the asymmetric species case was considered (Alonso

et al., 2008) and it was further demonstrated that the species abundance distribution has exactly the same sampling formula for both zero-sum and non-zero-sum models within the neutral approximation (Etienne *et al.*, 2007; Haegeman and Etienne, 2008). The simplest mode of speciation – a point mutation – has been the one most commonly used to derive the species abundance distribution of the meta-community. A Markovian approach incorporating various modes of speciation such as random fission was also presented by Haegeman and Etienne (Haegeman and Etienne, 2009, 2010). An innovative way to model speciation was used in (Rosindell *et al.*, 2010). The authors introduced protracted speciation, i.e., a gradual process whereby new species are created with a few individuals, instead of being an instantaneous process starting with exactly one individual. This gradual speciation improves predictions of species lifetimes, speciation rates and the number of rare species.

B. Coalescent approach to Neutral Theory

An exciting alternative approach to compute the distribution of individuals in a community is based on coalescent theory. The idea of this approach consists in tracing each member of the community back to their ancestors that first immigrated into the community (Etienne, 2005; Etienne and Olf, 2004b). A succinct comparison between the coalescent approach and other method is presented in (Chave *et al.*, 2006; Etienne and Alonso, 2005).

Within this framework, the local community consists of a fixed number of individuals J that undergo zero-sum dynamics. At each timestep a randomly chosen individual dies and is replaced by a new individual that is an offspring of a randomly chosen individual in the community (of size $J - 1$), or by one of I potential immigrants from the meta-community. The immigration rate is thus $m = I/(I + J - 1)$, where the parameter I is merely a proxy for the immigration rate. The meta-community is governed by the same dynamics, with speciation replacing immigration.

Two observations are crucial to derive the sampling formula for the species abundance distribution. Firstly, because there is no speciation in the local community, each individual is either an immigrant or a descendent of an immigrant. Thus, the information pertaining to the species abundance distribution is specified by considering “the ancestral tree” (tracing back each individual to its immigrant ancestor - which is somewhat similar to the

phylogenetic tree construction in genetics) and the species composition of the set of all the ancestors. Secondly, this set of all ancestors can be considered as a random sample from the meta-community and its species abundance distribution is provided by the well-known multivariate Ewens distribution, corresponding to the Fisher's log-series in the limit of large sample size, that describes neutral population genetics models with speciation and no immigration. As we do not know the ancestor of each individual, the final formula for the species-abundance composition of the data comes from summing the probabilities of all possible combinations of ancestors that give rise to the observed data.

In a tour de force calculation, the resulting sampling formula is found to be (Etienne, 2005)

$$P[D|\theta, m, J] = \frac{J!}{\prod_{i=1}^S n_i \prod_{j=1}^J \Phi_j!} \frac{\theta^S}{(I)_J} \sum_{A=S}^J K(D, A) \frac{I^A}{(\theta)_A}, \quad (28)$$

where $P[D|\theta, m, J]$ is the probability of observing the empirical data D consisting of S species with abundances n_i , $i \in [1, S]$ given the value of the fundamental biodiversity number θ of the meta-community, the immigration rate m and the local community size J . Φ_j is the number of species with abundance j ,

$$K(D, A) = \sum_{\{a_1, \dots, a_S | \sum_{i=1}^S a_i = A\}} \prod_{i=1}^S \frac{\bar{s}(n_i, a_i) \bar{s}(a_i, 1)}{\bar{s}(n_i, 1)}, \quad (29)$$

where \bar{s} is the unsigned Stirling number of the first kind. Finally, $(x)_y$ is the Pochhammer symbol $\prod_{i=1}^y (x + i - 1)$.

These expressions are simplified versions of the original results presented by the authors and as expected, they provide a comparable fit to the ME approach for the data from the tropical forests. It is heartening that equivalent results can be derived using very different approaches.

III. DYNAMICS IN NEUTRAL THEORY

So far we have focused on the implications of neutral theory when models describing neutral patterns reach a steady state. The stationary condition allows one to take advantage of a variety of different mathematical techniques to obtain analytical expressions of ecological patterns. However, stationarity is not always a good approximation, either because the ecosystem is still in a state of flux, or because the assumption may hide different

and important processes that lead to the same final steady state. Furthermore, one can calculate time dependent correlation functions (Hashitsume *et al.*, 1991) (see, for example, (46) below). It is therefore essential to understand the temporal behavior of ecosystems in order to discriminate between ecological processes that would otherwise be indistinguishable at stationarity. Statistical comparisons between time-dependent patterns are usually more difficult than those between stationary patterns because they require more data and long empirical time series that are rarely available. In addition, although time-dependent solutions facilitate stronger tests when confronted with data, they are more difficult to obtain and this is the reason why only a few studies have investigated the temporal behavior of neutral models.

An important method to study the time dependence is van Kampen's system size expansion. Assuming that the total population of the system, N , is very large and the initial population is of the order of N , one expects that, at short times at least, the probability distribution of the system, $P(n, t)$, would peak sharply around the macroscopic value $n = N\phi(t)$ with a width of order $\sim \sqrt{N}$. The function $\phi(t)$ is chosen to describe the evolution of the peak. These considerations lead to the Ansatz $n = N\phi(t) + \sqrt{N}\xi$, where ξ is a new random variable. When expressing the ME in terms of ξ and expanding it in \sqrt{N} , at order \sqrt{N} one recovers the deterministic evolution of the macroscopic state of the system, and at order N^0 one obtains a linear Fokker-Planck equation whose coefficients depend on time through the function $\phi(t)$. Therefore, fluctuations around the peak are Gaussian in a first approximation (details of the method can be found in Ref.(Kampen, 2007)). This expansion has some attractive features but it is usually only a good approximation of the temporal evolution of the original ME at short or intermediate temporal scales, since the Gaussian behavior is usually lost over longer periods. The time dependence of mean-field neutral models was investigated by means of van Kampen's system size expansion in (McKane *et al.*, 2000) and (McKane *et al.*, 2004) and a good agreement with simulations was found at early times.

Some neutral models can also be formulated in terms of discrete-time Markov processes that have the form $p_i(t+1) = \sum_j Q_{ij}p_j(t)$, where $p_i(t)$ is the probability that a species has i individuals at time t , and Q_{ij} is the probability that a species changes its abundance from j to i in one time step (i.e. the transition matrix). If the community has N individuals in total and either one birth or one death is allowed in one time step, then Q is a tridiagonal

matrix with dimension $N + 1$. Solving these models in time is basically equivalent to finding the eigenvalues along with the left and right eigenvectors of the matrix Q , which is usually a non-trivial task. Using a theorem derived in population genetics, Chisholm (Chisholm, 2011) was able to obtain the eigen decomposition of the transition matrix of a neutral model that describes the behavior of the local community in the original Hubbell model (Hubbell, 2001). Also, he was able to show that the first two eigenvectors are sufficient to provide a good approximation of the full time-dependent solution.

Other authors have focused on specific temporal patterns that can sometimes be calculated more easily. For instance, a measure that has been used to quantify the diversity of a community is the Simpson index (Simpson, 1949), D , which is defined as the probability that two individuals drawn randomly from a well-mixed community belong to the same species. If the individuals are drawn with replacement, then one obtains $D = \sum_i n_i^2 / N^2$, where n_i is the number of individuals of species i (we have already encountered this index in Section II when we calculated the beta-diversity in the random placement model). This index is also a good indicator of the relative importance of spatial effects in a community. The dynamic evolution of D is usually simple enough to allow analytical calculations (Vanpeteghem *et al.*, 2008). Other interesting dynamic properties of neutral models have been studied, where species' extinction and monodominance were investigated by looking into the first-passage properties and fixation times of the process (Babak, 2006; Babak and He, 2009). Empirical power-law relationships between the temporal mean and variance of population fluctuations (the so-called Taylor's power law) are also in good agreement with neutral predictions (Keil *et al.*, 2010).

Although it is not mandatory from the neutral assumption, models often assume that demographic stochasticity is the main source of fluctuations in stochastic neutral dynamics (Hu *et al.*, 2006; Ricklefs, 2003). Demographic randomness originates from the intrinsic stochastic nature of birth and death events within a discrete population of individuals. However, other sources may be important, such as environmental stochasticity that, by contrast, encompasses effects of abiotic and biotic environmental variables on all individuals (strictly speaking, we should more correctly refer to symmetric – instead of neutral – models (see BOX2) when these incorporate environmental fluctuations, because the per-capita vital rates vary across individuals; however, many authors do not appreciate this distinction). There is theoretical and empirical evidence that forest dynamics exhibit signatures of environmen-

tal variance (Chisholm *et al.*, 2014; Kalyuzhny *et al.*, 2015; Lande *et al.*, 2003). Usually, neutral models only based on demographic stochasticity tend to overestimate the expected times to extinction for abundant species (Ricklefs, 2006), whose temporal fluctuations will also be underestimated (Mutshinda *et al.*, 2008). Incorporating an environmental source of randomness as well as more realistic forms of speciation have made some dynamic aspects of NT more realistic (Allen and Savage, 2007).

Here we will focus on the dynamic aspects of the species abundance distribution (SAD) (Azaele *et al.*, 2006) which, under appropriate assumptions, can be studied in detail (Lehningk, 1993; Masoliver and Perello, 2012) and whose predictions can be benchmarked against empirical data. We will also review recent progress in modeling dynamic patterns, including the species turnover distribution (Azaele *et al.*, 2006) as well as species' persistence-times (or lifetimes) distributions (Bertuzzo *et al.*, 2011; Pigolotti *et al.*, 2005; Pinto and Muñoz, 2011; Suweis *et al.*, 2012a) (see BOX 1).

A. The continuum limit of the Master Equation

The microscopic description of a system starts by correctly identifying the variables that define all the possible configurations of the system. Having decided how to describe the states of the system, the next step is to consider the transition rates among different states. The configurations of an ecological community can be described by different variables according to different levels of coarsening of the spatio-temporal scales. However, for the sake of simplicity, we will focus on ecosystems comprising S species and N individuals in total, in which the configurations are specified fully by the variable $\mathbf{n} = (n_1, \dots, n_S)$, where n_i is the population of the i -th species. Therefore, for the time being, we will ignore any spatial effects, which will be considered in section IV.

If the ecological community at time t_0 was in the configuration \mathbf{n}_0 , then the probability that the system will be in configuration \mathbf{n} at time $t > t_0$ is $P(\mathbf{n}, t | \mathbf{n}_0, t_0)$ ($P(\mathbf{n}, t)$ for brevity). Assuming a Markovian stochastic dynamics, the evolution of $P(\mathbf{n}, t)$ is governed by the ME (6) with transition rates $T[\mathbf{n} | \mathbf{n}']$ from the states \mathbf{n}' to \mathbf{n} . As seen in section II, the specification of a model is tantamount to defining the transition rates. This equation can be recast in another form that allows a simple expansion. Interactions among species change the variable \mathbf{n} in steps of size \mathbf{r}_ℓ , where $\ell = 1, 2, \dots, L$ and L is the total number of distinct

possible changes or reactions. For instance, if a species can either decrease or increase its population by k in a given time step, then $L = 2$. Species that change by one or that remain constant in a given time interval are instead described by $L = 3$, and when the total population is conserved, we have $\sum_{\ell=1}^L r_{i,\ell} = 0$. Of all the possible reactions, however, only a few are usually significant, ecologically relevant or meaningful. Moreover, unlike chemical reactions these meaningful reactions can be considered irreversible. Therefore, instead of summing over states \mathbf{n}' , as in eq.(6), we can sum over L different reactions. If $t_\ell[\mathbf{n}]$ is the transition rate of the ℓ -th reaction which involves the jumps \mathbf{r}_ℓ when species' populations are \mathbf{n} , then we can recast eq.(6) in the form

$$\frac{\partial P(\mathbf{n}, t)}{\partial t} = \sum_{\ell=1}^L \{t_\ell[\mathbf{n} - \mathbf{r}_\ell]P(\mathbf{n} - \mathbf{r}_\ell, t) - t_\ell[\mathbf{n}]P(\mathbf{n}, t)\} . \quad (30)$$

Assuming that $t_\ell[\mathbf{n}]$ is only a function of $\mathbf{x} = \mathbf{n}/N$ (this corresponds to eqs.(20), (21) with $N = J_M$ and (24), (25) with $N = J_L$ and $N \gg 1$ in both cases), we can write

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = \sum_{\ell=1}^L \{t_\ell \left[\mathbf{x} - \frac{\mathbf{r}_\ell}{N} \right] P \left(\mathbf{x} - \frac{\mathbf{r}_\ell}{N}, t \right) - t_\ell[\mathbf{x}] P(\mathbf{x}, t)\} , \quad (31)$$

where \mathbf{n} has been replaced by $N\mathbf{x}$. This form of the ME suggests that $|\mathbf{r}_\ell|/N \ll 1$ when $N \gg 1$ and hence, a Taylor expansion around \mathbf{x} of t_ℓ and P is legitimate, at least far from the boundaries. This is known as the Kramers-Moyal expansion (Gardiner, 2009). Truncating to second order in N , we obtain a Fokker-Planck equation of the form

$$\frac{\partial P(\mathbf{x}, \tau)}{\partial \tau} = - \sum_{i=1}^S \frac{\partial}{\partial x_i} [A_i(\mathbf{x})P(\mathbf{x}, \tau)] + \frac{1}{2N} \sum_{i,j=1}^S \frac{\partial^2}{\partial x_i \partial x_j} [B_{ij}(\mathbf{x})P(\mathbf{x}, \tau)] + O(N^{-2}) , \quad (32)$$

where the time has been rescaled ($\tau = t/N$), the functions $A_i(\mathbf{x})$ have been defined as follows

$$A_i(\mathbf{x}) \equiv \sum_{\ell=1}^L r_{i,\ell} t_\ell[\mathbf{x}] \quad (33)$$

and the entries of the matrix $B(\mathbf{x})$ are

$$B_{ij}(\mathbf{x}) \equiv \sum_{\ell=1}^L r_{i,\ell} r_{j,\ell} t_\ell[\mathbf{x}] . \quad (34)$$

The matrix $B(\mathbf{x})$ is a $S \times S$ symmetric matrix that is positive semi-definite (Gardiner, 2009). For one-step jumps, the entries of the matrices $\mathbf{r}_\ell \otimes \mathbf{r}_\ell$ can only be 1, -1 or 0. Unlike

the ME (30), the Fokker-Planck equation (32) governs the evolution of *continuous* stochastic variables that represent the relative species' populations under the so-called diffusion approximation. As the total population extends to infinity, we recover the deterministic evolution given by the Liouville equation

$$\frac{\partial P(\mathbf{x}, \tau)}{\partial \tau} = - \sum_{i=1}^S \frac{\partial}{\partial x_i} [A_i(\mathbf{x})P(\mathbf{x}, \tau)] , \quad (35)$$

whose solution has the simple form $P(\mathbf{x}, \tau) = \delta(\mathbf{x} - \mathbf{x}(\tau))$, where δ is the Dirac delta function. This form of P implies that $\mathbf{x}(\tau)$, i.e. the relative species' populations, are the solutions of the deterministic system $\dot{x}_i(\tau) = A_i(\mathbf{x})$, where $i = 1, 2, \dots, S$ (Gardiner, 2009).

So far we have assumed that the system has a fixed and finite total population, N , and therefore each population cannot exceed N , despite its fluctuations. However, sometimes it is useful to relax this constraint and allow each population to fluctuate within the interval $[0, \infty)$. In this case, a parameter may or may not exist that allows one to recover the deterministic limit. However, if we assume that $t_\ell[\mathbf{n}]$ are smooth functions of \mathbf{n} , and when $\mathbf{n} \rightarrow \mathbf{n} + \mathbf{r}_\ell$, they vary little with respect to a characteristic scale of the model (e.g., the total average population), and then, we can treat \mathbf{n} as continuous variables and expand $t_\ell[\mathbf{n} - \mathbf{r}_\ell]P(\mathbf{n} - \mathbf{r}_\ell, t)$ around \mathbf{n} . Starting from the ME (30) and using \mathbf{x} to indicate the continuous populations, we obtain

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = - \sum_{i=1}^S \frac{\partial}{\partial x_i} [A_i(\mathbf{x})P(\mathbf{x}, t)] + \frac{1}{2} \sum_{i,j=1}^S \frac{\partial^2}{\partial x_i \partial x_j} [B_{ij}(\mathbf{x})P(\mathbf{x}, t)] , \quad (36)$$

where time has not been rescaled and $A_i(\mathbf{x})$ and $B_{ij}(\mathbf{x})$ are as in Eqs. (33), (34). If we are within a neutral framework with one-step jumps, then $S = 1, L = 2$ and $r_1 = 1, r_2 = -1$. Setting $t_1[x] \equiv b(x)$ (birth rate) and $t_2[x] \equiv d(x)$ (death rate), from Eq. (36) one gets

$$\frac{\partial P(x, t)}{\partial t} = \frac{\partial}{\partial x} [d(x) - b(x)]P(x, t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} [d(x) + b(x)]P(x, t) . \quad (37)$$

Starting from the discrete formulation of the birth and death rates that we have seen previously, one can derive the expressions of the corresponding continuous rates. Following our earlier discussion, we can write the following general expansion

$$\frac{b(x)}{x} = b_1 + \frac{b_0}{x} + \dots \quad (38a)$$

$$\frac{d(x)}{x} = d_1 + \frac{d_0}{x} + \dots \quad (38b)$$

where we have dropped higher order terms in $1/x$. Clearly, Eq. (38) can be applied when species have sufficiently large populations. The constants b_0 and d_0 produce a density dependence effect that causes a rare species advantage (disadvantage) when $b_0 > d_0$ ($b_0 < d_0$). This effect has its roots in ecological mechanisms, such as the Janzen-Connell (Connell, 1971; Janzen, 1970) or Chesson-Warner effects (Chesson and Warner, 1981). However, the presence of net immigration or speciation in a local community (Volkov *et al.*, 2003) can also produce such density dependence, which is captured here by a mean field approach. Usually, the skewness of the RSA of various tropical forests and coral reefs indicates a rare species advantage (Volkov *et al.*, 2005), and thus, we will use $b_0 > d_0$. In addition, to simplify the analytical treatment and for parsimony, we choose $b_0 = -d_0$. Substituting the rates in (38) into Eq. (37) and setting $b_0 = -d_0 > 0$, we obtain

$$\frac{\partial P(x, t)}{\partial t} = \frac{\partial}{\partial x} [(\mu x - b)P(x, t)] + D \frac{\partial^2}{\partial x^2} [xP(x, t)] , \quad (39)$$

where $D = (d_1 + b_1)/2$, $\mu = 1/\tau = d_1 - b_1 > 0$ and $b = 2b_0 > 0$. $P(x, t) \equiv P(x, t|x_0, 0)$ is the probability density function of finding x individuals in the community at time t , given that at time $t = 0$ there were x_0 . Therefore $\int_n^{n+\Delta n} P(x, t)dx$ is the fraction of species with a population between n and $n + \Delta n$ at time t . Eq. (39) defines our model for species-rich communities and it governs the time evolution of a community population under the neutral approximation. The deterministic term drives the population to the stationary mean population per species b/μ and therefore, within our model, the population and the number of species can also fluctuate at stationarity with the ratio of the population to the number of species being fixed on average. Thus, this model is more flexible than the original model by Hubbell (Hubbell, 2001) in which ‘zero-sum dynamics’ that fixes the total number of individuals in a community was assumed.

The link established between the FP equation and the ME provides a useful interpretation of the coefficients: $\mu = d_1 - b_1$ is the imbalance between the per capita death and birth rates that inexorably drives the ecosystem to extinction in the absence of immigration or speciation. In this model τ is the characteristic time scale associated with perturbations

to the steady-state. When $D\tau \gg 1$ (as for the tropical forests we have analyzed) we have $d_1 = (2D + \tau^{-1})/2 \simeq D$, and so D can be thought of as an individual death rate. Finally, b plays the role of an effective immigration (or speciation) rate that prevents the community from becoming extinct.

As expected, different choices of b_0 and d_0 lead to very good fits of the RSA of various tropical forests. In particular, it is possible to see that the RSA fits are readily improved for large x when b_0 and d_0 are arbitrary parameters. However, if one introduces a fourth free parameter, the analytical treatment of the dynamics is much more involved. The Fokker-Planck equation with just three parameters is an ideal starting point to understand the dynamics governing species-rich ecosystems in a simplified fashion. Indeed, the agreement between empirical data and the macro-ecological properties predicted by the model is consistent with the simplifying assumption that tropical forests are close to their steady state. In addition, because the model is not only neutral, but also non-interacting, one may speculate that surviving species are those that were able to reach a steady state of coexistence by minimizing interspecies interactions.

Further insights into the nature of stochasticity can be achieved by writing down Eq. (36) in the equivalent Langevin form, which is an equation for the state variables themselves. Within the Itô prescription (Kampen, 2007), Eq. (36) is equivalent to the following Langevin equation

$$\dot{x}_i(t) = A_i(\mathbf{x}) + \sum_{j=1}^S b_{ij}(\mathbf{x})\xi_j(t) , \quad (40)$$

where $\xi_j(t)$ is Gaussian white noise with a zero mean and correlation $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij}\delta(t-t')$, and the matrix $b(\mathbf{x})$ is defined by the equation $B(\mathbf{x}) = b(\mathbf{x})b^T(\mathbf{x})$, where the entries of $B(\mathbf{x})$ are given in Eq. (34). This shows that $b(\mathbf{x})$ may be thought of as the “square root” of $B(\mathbf{x})$. However, because any orthogonal transformation of the noise $\xi_j(t)$ leaves the mean and the correlation unchanged, the matrix $b(\mathbf{x})$ is not uniquely determined.

For the present model, the Fokker-Planck Eq. (39) is equivalent to

$$\dot{x}(t) = b - \frac{x(t)}{\tau} + \sqrt{Dx(t)}\xi(t) \quad (41)$$

where we have instead assumed that the time correlation is $\langle \xi(t)\xi(t') \rangle = 2\delta(t-t')$. Eq. (41) shows how demographic stochasticity accounts for random ecological drift. In fact, given the relatively large number of individuals of any species, one expects that the detailed

nature of the stochastic noise is not important, such that fluctuations are proportional to $\sqrt{x(t)}$ due to the central limit theorem (Kampen, 2007). Thus, the multiplicative noise in the Langevin equation is not introduced *ad hoc* and can be justified on the basis of more general considerations.

B. Boundary conditions

The stochastic dynamics described by Eq. (41) governs the evolution of the population when it is strictly positive. Assuming that $b > 0$ and starting at some positive value, the process cannot reach negative values because the drift brings it to a positive value bdt in the following time interval dt . Therefore, the process is always non-negative with the origin acting as a singular boundary, which must be specified on the basis of ecological considerations, i.e. one should define what happens when the random ecological drift occasionally leads to a vanishing population. The most frequent ecological situations are given by the following boundaries

- We use reflecting boundary conditions at $x = 0$ to describe ecological communities in which a vanishing population is immediately replaced by a new individual (belonging to either a new or old species). For instance, this may happen when an ecosystem is coupled to a meta-community that extends across large spatial scales. With these boundary conditions, one can describe ecosystems in which biodiversity is continuously sustained by the net immigration of new individuals of new/old species. This prevents an ecosystem from becoming extinct and can finally achieve a non-trivial steady-state. In the case of the model defined by Eq.(39), reflecting boundaries are obtained by setting the flux of the probability distribution as zero at $x = 0$, although there are some equations for which such boundaries cannot be arbitrarily set, such as Eq.(39) (Feller, 1951).
- A second possibility arises when a community can lose individuals without any replacement or a net emigration flux of individuals from the ecosystem exists. One can then describe the system by introducing absorbing boundary conditions at $x = 0$. These constraints force ecosystems to march inexorably to extinction and the final steady-state corresponds to the complete extinction of species, i.e. $\lim_{t \rightarrow \infty} P(x, t) = \delta(x)$.

Usually, absorbing boundaries are obtained by setting the probability distribution at $x = 0$ equal to zero, yet some equations do have norm decreasing solutions, such as Eq. (39), (Feller, 1951) which do not vanish at $x = 0$.

It is possible to define other kinds of boundaries, according to different ecological behaviors when the population consists of just a few individuals. However, the reflecting and absorbing boundaries capture the most interesting situations in ecology.

As we have alluded to above, the reflecting and absorbing solutions of the Fokker-Planck equation (39) have rather unusual properties. This is essentially due to the fact that the diffusion term, Dx , vanishes at $x = 0$. Obviously, it is always possible to rewrite a vanishing diffusion term at zero into a non-vanishing one with a suitable change of variables, yet the same change of variables leads to a singular drift term at the boundary. Thus, equations with a vanishing diffusion term or with a singularity of the drift term at the boundary in general share similar unusual properties. In fact, there exist regions in the parameter space of Eq. (39) in which one cannot arbitrarily fix a boundary at zero, and the solution is completely defined once the initial condition is prescribed. This is because the probability of accessing the origin depends on the values of the parameters. The problem can be studied by introducing the probability to reach the origin for the first time. It is possible to show (Masoliver and Perello, 2012) that this first-passage probability to the origin is zero when $b/D > 1$ and therefore, in this case one cannot obtain any solution with absorbing boundaries and the solution can only be reflecting at $x = 0$. However, a (unique) norm decreasing solution exists when $0 < b/D < 1$, which corresponds to the regular absorbing solution but that does not vanish at $x = 0$. Also, in the same region of parameters, a reflecting solution exists with an integrable singularity at $x = 0$. All ecosystems that we have studied so far are well described by b and D within this parameter region.

C. Stationary and time dependent solutions

An ecosystem described by Eq. (39) can reach a non trivial steady state when setting reflecting boundary conditions at $x = 0$. In this case, the normalized stationary solution is the gamma distribution (Kampen, 2007), i.e.

$$P_0(x) = \frac{(D\tau)^{-b/D}}{\Gamma(b/D)} x^{\frac{b}{D}-1} e^{-\frac{x}{D\tau}} \quad (42)$$

where $\Gamma(x)$ is the gamma function (Lebedev, 1972). Note the distribution exists only when $b > 0$, i.e. when the ecosystem has a net immigration rate of species. $P_0(x)$ is the pdf of the relative species abundance at stationarity and it gives the probability that a given arbitrary species present in the ecosystem has x individuals. Fitting the empirical RSAs allows one to estimate two combinations of the parameters, i.e. b/D and $D\tau$, with τ being the characteristic temporal scale to approach stationarity. Interestingly, for $0 < b/D \ll 1$ one obtains the continuum approximation of the celebrated Fisher logseries (Fisher *et al.*, 1943), the RSA distribution of a meta-community.

One can analytically solve Eq. (39) at any time with arbitrary initial conditions, so that the evolution of the population can be traced even far from stationary conditions. The time dependent solution with reflecting boundary conditions ($b > 0$) and initial population x_0 is

$$P_r(x, t|x_0, 0) = \left(\frac{1}{D\tau}\right)^{\frac{b}{D}} x^{\frac{b}{D}-1} e^{-\frac{x}{D\tau}} \frac{\left[\left(\frac{1}{D\tau}\right)^2 x_0 x e^{-t/\tau}\right]^{\frac{1}{2}-\frac{b}{2D}}}{1 - e^{-t/\tau}} \times \exp\left[-\frac{\frac{1}{D\tau}(x+x_0)e^{-t/\tau}}{1 - e^{-t/\tau}}\right] I_{\frac{b}{D}-1}\left[\frac{\frac{2}{D\tau}\sqrt{x_0 x e^{-t/\tau}}}{1 - e^{-t/\tau}}\right]. \quad (43)$$

where $I_\nu(z)$ is the modified Bessel function of the first kind (Lebedev, 1972). Because $I_\nu(z) \simeq (z/2)^\nu/\Gamma(\nu+1)$ when $z \rightarrow 0^+$, for large time intervals Eq. (43) converges to $p_0(x)$ in Eq. (42). By setting $x = Dy^2/4$ ($y > 0$), when $b/D = 1/2$ the process can be mapped into the simpler Ornstein-Uhlenbeck process (Kampen, 2007) with a boundary (reflecting or absorbing) at $x = 0$, and when $b/D = 1$ it can be mapped into the (reflecting) Rayleigh process (Gardiner, 2009).

From the Fokker-Planck Eq. (39), it is easy to derive the evolution of the mean population per species. It is simply

$$\langle x(t) \rangle = b\tau + (x_0 - b\tau)e^{-t/\tau} \quad (44)$$

where $\langle x(0) \rangle = x_0$ is the initial population per species in the community. The mean number of individuals per species converges to $b\tau$ at stationarity, with standard deviation $\tau\sqrt{bD}$. The auto-covariance, $\kappa(t) = \langle x(t)x(0) \rangle - \langle x(t) \rangle \langle x(0) \rangle$, under stationary conditions is given by

$$\kappa(t) = \int_0^\infty [xx_0 - \langle x(t) \rangle \langle x(0) \rangle] P_r(x, t|x_0, 0) P_0(x_0) dx dx_0 \quad (45)$$

where $P_r(x, t|x_0, 0)$ is the reflecting solution in Eq. (43) and $P_0(x_0)$ is the stationary solution in Eq. (42). However, $\kappa(t)$ can be more rapidly obtained from Eq. (41). In fact, since in the Itô prescription $\xi(t)$ and $x(t)$ are uncorrelated, one gets $\langle \sqrt{x(t)}\xi(t) \rangle = \langle \sqrt{x(t)} \rangle \langle \xi(t) \rangle = 0$ because we have assumed that $\langle \xi(t) \rangle = 0$. Therefore the evolution equation for $\langle x(t)x(0) \rangle$ becomes $\partial_t \langle x(t)x(0) \rangle = b \langle x(0) \rangle - \langle x(t)x(0) \rangle / \tau$, whose solution is $\langle x(t)x(0) \rangle = (b\tau)^2 + bD\tau^2 e^{-t/\tau}$. Hence, under stationary conditions, we have $\langle x \rangle = b\tau$, $\langle x^2 \rangle = \langle x \rangle^2 + bD\tau^2$ and then $\kappa(t) = bD\tau^2 e^{-t/\tau}$ for $t > 0$. Therefore, the time correlation function is a simple exponential. Because the equation has only one temporal scale, τ is also the characteristic relaxation time of the process, i.e. the temporal scale upon which the system at stationarity recovers from a small perturbation.

D. The species turnover distribution

According to the NT, the turnover of ecological communities reflects their continuous re-assembly through immigration/emigration and local extinction/speciation. Species' histories overlap by chance due to stochasticity, yet their lifetimes are finite and distributed according to the underlying governing process. Non-trivial stationary communities are reached because old species are continuously replaced by new species, bringing about a turnover of species that can be studied and modeled within our framework.

To measure species turnover one usually considers the population of a given species at different times, then studying the temporal evolution of their ratios. For an ecosystem close to stationarity, one can look at the distribution $\mathcal{P}(\lambda, t)$, i.e. the probability that at time t the ratio $x(t)/x(0)$ is equal to λ , where $x(t)$ and $x(0)$ are the population of a given species at time $t > 0$ and $t = 0$, respectively. Thus, the species turnover distribution (STD) $\mathcal{P}(\lambda, t)$ by definition is

$$\begin{aligned} \mathcal{P}(\lambda, t) &= \langle \delta(\lambda - x/x_0) \rangle \\ &= \int_0^\infty dx_0 \int_0^\infty dx P(x, t|x_0, 0) P_0(x_0) \delta(\lambda - x/x_0). \end{aligned} \tag{46}$$

Here $P(x, t|x_0, 0)$ is either the reflecting or absorbing solution defined in Eqs. (43) and (49), and $P_0(x)$ is given in Eq. (42), and $\lambda > 0$ and $t > 0$.

In Eq. (46), one can use the time-dependent reflecting or absorbing solution according

to different ecological dynamics. We should use the reflecting solution when not concerned with the extinction of the species present at the initial time point and especially, when accounting for any new species introduced through immigration/speciation. The expression for the reflecting STD can be found in Ref. (Azaele *et al.*, 2006).

One can show that it has the following power-law asymptotic behaviour for a fixed t

$$\begin{cases} \mathcal{P}(\lambda, t) \approx \frac{k_1(t)}{\lambda^{\frac{b}{D}+1}} & \text{for } \lambda \gg 1 \\ \mathcal{P}(\lambda, t) \approx k_2(t)\lambda^{\frac{b}{D}-1} & \text{for } \lambda \ll 1 \end{cases} \quad (47)$$

where the functions k_1 and k_2 are independent of λ . Customarily, rather than the random variable λ , ecologists study $r = \log(\lambda)$ that is distributed according to $g(r, t) = e^r \mathcal{P}(e^r, t)$ and that can be compared to empirical data Fig.4. We obtain an estimate of τ , the characteristic time scale for the BCI forest, which is around 3500 ± 1000 yrs (for trees with > 10 cm of stem diameter at breast height (dbh)) and 2900 ± 1100 yrs (for trees with > 1 cm dbh), where the broad uncertainty is due to the fact that the data are sampled over relatively short time intervals.

These fits not only provide direct information about the time scale of evolution but also, they underline the importance of rare species in the STD. b/D is closely tied to the distribution of rare species and in fact, $P_0(x) \sim x^{b/D-1}$ for $x \ll D\tau$. Alternatively, the dependence of the STD on b/D and τ suggests that at any fixed time $t > 0$, rare species are responsible for the shape of the STD.

E. Persistence or lifetime distributions

A theoretical framework to study and analyze persistence or extinctions of species in ecosystems allows one to understand the link between environmental changes (like habitat destruction or climate change (Brown, 1995; Diamond, 1989; May, 2010; Svenning and Condit, 2008; Thomas *et al.*, 2004)) and the increasing number of threatened species.

The persistence or lifetime τ of a species is defined as the time interval between its emergence and its local extinction (see Fig. 5) within a given geographic region (see (Keitt and Stanley, 1998; Pigolotti *et al.*, 2005)). In statistical physics, this is known as the distribution of first passage times to zero of the stochastic processes describing the species

abundance dynamics. According to the neutral theory of biodiversity, species can span very different lifetime intervals and thus, at a local scale, persistence times are largely controlled by ecological processes like random drift, dispersal and immigration.

1. Discrete Population Dynamics

The simplest baseline model to study the persistence time distribution is a random walk in the species abundance \vec{n} , i.e. ME (10) with $b_n = d_n = c$ and absorbing boundary condition in $n = 0$. According to this scheme, local extinction is equivalent to a random walker's first passage to zero and thus, the resulting persistence-time distribution has a power-law decay with exponent $3/2$ (Chandrasekhar, 1943; Pigolotti *et al.*, 2005).

A further step in modeling life-time distributions can be made by taking into account birth, death and speciation (Alonso *et al.*, 2006; Hubbell, 2001; Muneeppeerakul *et al.*, 2008; Volkov *et al.*, 2003) through a mean field scheme of the voter model with speciation (Chave and Leigh, 2002; Durrett and Levin, 1996) (see description in section IV), i.e. ME (6) with birth and death rates given by eqs. (20) and (21) in the large J_M limit, $b(n) = (1 - \nu)n/J_M$ and $d(n) = n/J_M$ for $n \geq 0$. The corresponding persistence or life-time distribution is given by $p_\tau(t) = -\frac{dP(0,t)}{dt}$, where $P(0,t)$ is the probability that a species has a zero population at time t . The asymptotic behavior of the resulting persistence-time distribution (i.e. $p_\tau(t)$) exhibits a power-law scaling modified by an exponential cut-off (Pigolotti *et al.*, 2005):

$$p_\tau(t) \propto t^{-2}e^{-\nu t}, \quad (48)$$

for t greater than some lower cut-off value. Here, t is measured in units of J_M and ν is now a speciation rate rather than a probability (Eq. (48) is derived in appendix B). We note that, in this context, $p_\tau(t)$ has a characteristic timescale $1/\nu$ for local extinctions determined by the speciation/migration rate. The general case when none of the coefficients is zero and are given by Eqs. (38) can also be solved. In this case, p_τ displays a crossover from the $t^{-3/2}$ to the t^{-2} behavior at a certain characteristic time and finally, an exponential decay beyond yet another characteristic time scale (Pigolotti *et al.*, 2005). It has been shown numerically (Bertuzzo *et al.*, 2011; Suweis *et al.*, 2012a) that for the spatial voter model in dimensions $d = 1, 2, 3$, the exponent of the power-law in Eq. (48) changes to $t^{-\alpha}$, with $\alpha < 2$ depending on the topological structure underlying the voter model. In particular, $\alpha = 3/2$ in $d = 1$

and $\alpha = 2$ for any $d \geq 3$, whereas in $d = 2$ one gets $p_\tau(t) \propto t^{-2}e^{-\nu t} \ln t$ as shown in (Pinto and Muñoz, 2011) and references therein.

We have seen in section II that the RSA pattern does not depend on the biological details of the ecosystem under analysis. Thus, one may wonder if the persistence time distribution is also a universal macro-ecological pattern. Indeed, it has been shown (Bertuzzo *et al.*, 2011; Suweis *et al.*, 2012a) that the power-law with an exponential cut-off shape predicted for the persistence time distribution by the NT (see appendix B) is common to very different types of ecosystems (see Figure 6). Other exact formulæ for species ages and species lifetimes have been proposed in neutral (Chisholm and ODwyer, 2014) and non-neutral models (Noble and Fagan, 2015; O’Dwyer and Chisholm, 2014).

Another interesting and related quantity is the *survival distribution* P_{τ_s} defined as the probability that a species randomly sampled from the community at stationarity is still present in the community after a time t . This quantity depends on the initial conditions as $P_{\tau_s}(t) = \int p_\tau(t|n_0)P_0(n_0)dn_0$, where $p_\tau(t|n_0)$ is the lifetime distribution for a species that initially has a population n_0 . Assuming that the stationary distribution of population abundances is the Fisher log series, then $P_{\tau_s}(t) \sim t^{-1}$ when $t \ll t^*$, whereas $P_{\tau_s}(t) \sim e^{-t/t^*}$ when $t \gg t^*$ (Pigolotti *et al.*, 2005; Suweis *et al.*, 2012a). It can be shown that this asymptotic behavior of the survivor distribution is valid regardless of the functional shape of the birth and death rate $b(n)$ and $d(n)$ involved in the ME driving the evolution of $P(n, t)$ (Suweis *et al.*, 2012b).

2. Continuum limit

In the continuum limit, a crucial distribution for the analysis of species’ extinction is the time dependent solution of Eq. (39) with absorbing boundaries at $x = 0$. We refer to (Feller, 1951) for its complete derivation. This probability distribution only exists when $b < D$ and is given by

$$P_a(x, t|x_0, 0) = \frac{(D\tau)^{-1}}{1 - e^{-t/\tau}} \exp \left[-\frac{\frac{1}{D\tau}(x + x_0e^{-t/\tau})}{1 - e^{-t/\tau}} \right] \times \\ \times \left(\frac{x}{x_0} e^{t/\tau} \right)^{\frac{b}{2D} - \frac{1}{2}} I_{1-\frac{b}{D}} \left[\frac{\frac{2}{D\tau} \sqrt{x_0 x e^{t/\tau}}}{e^{t/\tau} - 1} \right]. \quad (49)$$

Note that (49) is finite at $x = 0$ but $\lim_{x \rightarrow 0^+} P_a(x, t|x_0, 0) \neq 0$.

The lifetime distribution can be calculated analytically using Eq. (49), i.e. $p_\tau(t) = -\frac{d}{dt} \int_0^\infty P_a(x, t|1, 0) dx$. It can be shown that the lifetime distribution calculated in this way displays the same asymptotic behavior as (48), although the functional form is different (Azaele *et al.*, 2006). In the continuum limit, one can also obtain the mean extinction time (Azaele *et al.*, 2006):

$$\langle t \rangle = \tau \int_0^1 \frac{x^{-b/D} - 1}{1 - x} dx = -\tau(\gamma + \psi(1 - \beta)) \quad (50)$$

which depends only on $0 < b/D < 1$ and τ . Here, $\gamma = 0.577\dots$ is the Euler constant and $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the logarithmic derivative of the Gamma function (Lebedev, 1972). Eq. (42) can be used to fit the RSA of various tropical forests yielding $\langle t \rangle / \tau = 1.94, 1.67, 0.67, 0.95$ and 1.38 for Yasuni, Lambir, Sinharaja, Korup and Pasoh, respectively (see (Azaele *et al.*, 2006)). These time scales are in accord with the estimates of extinction times presented elsewhere (Pimm *et al.*, 1995) and it is quite interesting that $\langle t \rangle / \tau$ depends on b/D only, which can be calculated from the steady-state RSA without the need for dynamic data. The values of b/D obtained from various tropical forests (Azaele *et al.*, 2006) suggest that $\langle t \rangle \simeq \tau$, although this is not built into the model. In general, if $\tau \gg \langle t \rangle$, extinction would be much faster than recovery and the ecosystem will not reach a steady state. However, if $\tau \ll \langle t \rangle$ the ecosystem would recover from external disturbances very rapidly with respect to the extinction time and therefore, it would be very robust. This would leave little room for the action of evolution. Therefore, the fact that $\langle t \rangle \simeq \tau$ suggests that ecosystems at stationarity might be marginally stable — not so stable that they are frozen in time and not so fragile that they are prone to extinction. From estimates of the model parameters b/D , τ and thus predictions of $\langle t \rangle$, many biological and ecological features of the ecosystem may be understood.

IV. NEUTRAL SPATIAL MODELS AND ENVIRONMENTAL FRAGMENTATION

So far we have considered models, which assume that all individuals, at a given time, experience essentially the same conditions. They live in well-mixed habitats in which environmental heterogeneity and spatial distance are not important. These models are conceptually

and mathematically simple and this is the basic reason why they have been widely used. However, the spatial structure of ecosystems is able to control the shape of many patterns and it is also a critical factor for understanding species' coexistence (Tilman, 1994). In fact, models with well-mixed populations predict that the best competitor is able to displace a pool of species competing for the same limiting resource. There is empirical and theoretical evidence that space plays a crucial role in maintaining species diversity in communities with a single limiting resource, as demonstrated in the grasslands of the Cedar Creek Natural History Area (USA) (Tilman, 1994).

Space is therefore an essential element for understanding the organization of an ecosystem and most empirical observations are spatial. Often, the dynamics and composition of a community cannot be disentangled from its spatial aspects. Unfortunately, it is very difficult to derive analytical predictions for spatial (stochastic) models. The main difficulty is due to the fact that these are out-of-equilibrium models (Grilli *et al.*, 2012a).

Spatial effects can be incorporated into the theoretical framework, with the ME of (6) remaining formally the same by considering the index i in n_i as a composite index $i = (\alpha, \mathbf{r})$, where α identifies the species and \mathbf{r} identifies its spatial location. The set of all spatial locations \mathbf{r} will be denoted by Λ . Now the transition rates should take into account dispersal from nearby locations. To simplify, we will use the notation $n_i(\mathbf{r})$, where i indicates one of the S species and \mathbf{r} a spatial location.

At present, no coherent spatial neutral theory exists but rather, there is a collection of models and techniques that can explain some spatial patterns. In this section, we review some of those approaches.

The relationship between the number of species and the area sampled is probably one of the oldest quantities studied in ecology (Watson, 1835). Schoener (Schoener, 1976) referred to it as “one of community ecology’s few genuine laws”. The Species-Area relationship (SAR) is defined as the average number of species $\langle S(A) \rangle$ sampled in an area A .

Arrhenius (Arrhenius, 1921) postulated a power-law relationship $\langle S(A) \rangle = cA^z$. Empirical curves shows an inverted S -shape (Preston, 1960) (see Fig. 7), with a linear behavior at small and large areas, and a power-law with an exponent z at intermediate scales. This behavior seems to be pervasive and has been reported for distinct ecosystems.

Despite some notable exceptions (Gould, 1979), the value of the exponent z has attracted most attention in studies of SAR. This value of z is far from universal, ranging from 0.1 to

0.5 (Martín and Goldenfeld, 2006) and showing dependence on latitude (Schoener, 1976), body-mass, taxa and general environmental conditions (Martin, 1981; Power, 1972). The exponent z is interpreted as a measure of biodiversity.

Several models tried to reproduce the empirical behavior and a simple but useful assumption involves considering different species as independent realizations of the same process (Coleman, 1981). It is important to note that this assumption is stronger than neutrality, because neutrality does not imply independence. Under these assumptions the SAR is given by Eq. (1).

The Endemic-Area relationship (EAR) is defined as the number of species that are completely contained (i.e. endemic) in a given area (see BOX 1). It is not generally simply related to the SAR (He and Hubbell, 2011). If the species are considered as independent realizations of a unique process, then

$$\langle E(A) \rangle = S_{tot} P_0(A^c) . \quad (51)$$

where S_{tot} is the total number of species in the system, $P_0(A^c)$ is the probability that a species is not present in A^c , the complement of A , and that it is therefore completely contained in A .

The β -diversity is a spatially explicit measure of biodiversity. A simple and useful measure of β -diversity is the similarity index, i.e. the fraction of common species shared between different locations. It can also be defined as the probability $F(r)$ that two individuals at a distance r are conspecific (Chave and Leigh, 2002). This quantity may be related to the two point correlation function $G_{ij}(r)$. Under the assumption of translational invariance, this latter is defined as

$$G_{ij}(r) := \langle n_i(x) n_j(y+r) \rangle = \frac{1}{S} \left[\frac{1}{V} \sum_x \sum_y n_i(x) n_j(y) \delta_{ij} \delta(\|x-y\| - r) \right] , \quad (52)$$

where $n_i(x)$ is the number of individuals of the species i in the location x , $\|x-y\|$ is the distance between x and y , V is the number of site locations and $\delta(\|x-y\| - r)$ selects only pairs at a distance r . $F(r)$ is the probability that two individuals at a distance r are conspecific, i.e. it is the ratio between the number of pairs of individuals belonging to the same species at a distance r and the total number of pairs of individuals at a distance r . Therefore we obtain

$$F(r) = \frac{\sum_i \sum_x \sum_y n_i(x) n_i(y) \delta(\|x-y\| - r)}{\sum_i \sum_j \sum_x \sum_y n_i(x) n_j(y) \delta(\|x-y\| - r)} , \quad (53)$$

that can be rewritten as

$$F(r) = \frac{\sum_i \langle n_i(x)n_i(x+r) \rangle}{\sum_i \sum_j \langle n_i(x)n_j(x+r) \rangle} = \frac{\sum_i G_{ii}(r)}{\sum_i \sum_j G_{ij}(r)}. \quad (54)$$

If the spatial positions of different species are independent, then $G_{ij}(r) = \langle n(x) \rangle^2 := \rho^2$, and by defining $G_{ii}(r) = G_2(r)$, the β -diversity reads

$$F(r) = \frac{G_2(r)}{(S-1)\rho^2 + G_2(r)}. \quad (55)$$

Most theoretical work consists of attempts to relate these ecological quantities with other spatial and non-spatial observable factors. For instance, a typical problem is to calculate the SAR knowing the β -diversity and the RSA over a global scale. One of the future challenges will be to relate and predict spatial patterns on different scales (upscaling and downscaling), having only local information on one or more patterns.

A. Phenomenological models

Phenomenological models do not assume any microscopic dynamics but they are rather based on a given phenomenological distribution of individuals in space.

The simplest assumption is to consider individuals at random positions in space (Coleman, 1981; Coleman *et al.*, 1982).

This null model, usually known as “random placement” model, can be used to obtain predictions for the SAR and EAR having the RSA or the SAD as an input. Even though this assumption is not realistic, the random placement model turns out to be very useful to capture the relevant aspects of the relationship between RSA and SAR, and it also gives reasonable predictions that can be benchmarked against empirical data. In addition, this assumption also allows direct connections between SAR and EAR to be formulated (He and Hubbell, 2011).

Consider $S(A_0)$ species in a region of total area A_0 . Species i has an abundance n_i and the $N = \sum_i n_i$ individuals are uniformly distributed at random in the area A_0 . If we observe a sub-region of area A , the probability of observing a particular individual is simply A/A_0 , while the probability of not observing it is $1 - A/A_0$. The probability of not observing species i will then be $(1 - A/A_0)^{n_i}$, given the fact that the positions of individuals are independent.

We can then obtain the average number of species observed in an area A as

$$\langle S(A) \rangle = S(A_0) - \sum_{i=1}^S \left(1 - \frac{A}{A_0}\right)^{n_i}, \quad (56)$$

where $S(A_0) = S_{tot}$ is the total number of species in the system.

The simple framework of the random placement model also allows the EAR to be calculated. Using (51) we obtain

$$\langle E(A) \rangle = \sum_{i=1}^S \left(\frac{A}{A_0}\right)^{n_i}. \quad (57)$$

Despite the simplicity of the approximation, the EAR evaluated using random placement captures the quantitative behaviour of several observed ecosystems (He and Hubbell, 2011).

Under random placement assumptions, one can obtain the EAR from the SAR and vice versa. To calculate the EAR in (51), we calculated the number of species with zero individuals in the area complementary to that of interest. This number is equal to the difference between the number of species in the whole area and the number of species in the complementary area. The complementary area has a non-trivial shape and under general assumptions, this quantity is not easy to calculate. Under the random placement assumption, the number of species in the complementary area is the SAR of the complementary area. The EAR is therefore,

$$\langle E(A) \rangle = S(A_0) - S(A_0 - A). \quad (58)$$

In (He and Hubbell, 2011), a careful analysis of the reliability of this extrapolation was presented to predict the empirical EAR and it was shown that the random placement approximation describes the empirical data well.

One can obtain a closed form expression for the EAR and SAR by starting with a RSA distribution. In the case of a Fisher log-series (see eq. (18)), the SAR reads

$$\langle S(A) \rangle = S(A_0) - \sum_{n=1}^{\infty} \theta \frac{r^n}{n} \left(1 - \frac{A}{A_0}\right)^n = \theta \log\left(1 + \frac{r}{1-r} \frac{A}{A_0}\right), \quad (59)$$

which follows from the observation that $S(A_0)$ is equal to $-\theta \log(1-r)$ of eq. (18). The same calculation can be performed for the EAR, obtaining

$$\langle E(A) \rangle = -\theta \log\left(1 - r \frac{A}{A_0}\right). \quad (60)$$

In real ecosystems, individuals are not distributed uniformly in space but rather, due to dispersal limitation, individuals of the same species tend to be clustered. This is confirmed

by empirical β -diversity plots (Chave and Leigh, 2002). The phenomenological approach of random placement can be generalized to a non-uniform distribution of individuals in space, taking into account the empirically observed spatial clustering of conspecific individuals (Grilli *et al.*, 2012b). Individuals are distributed in space via a Poisson Cluster Process (PCP). In a PCP, the centers of the cluster of points are uniformly distributed in space. A random number of points are distributed around each center according to a given spatial kernel. The process depends on two distributions: the number of points in each cluster and the spatial kernel. It is possible to show (Grilli *et al.*, 2012b) that these two distributions may be related to the RSA and the β -diversity. An analytical formula for the SAR and the EAR can therefore be obtained given the RSA and the β -diversity. This approach reproduces the S -inverted shape observed in empirical systems, showing that this shape can be explained simply in terms of spatial correlations of conspecific individuals.

B. Spatial Stochastic Processes

There are two possible ways to include space in a neutral stochastic model, either implicitly or explicitly.

1. Spatially implicit model

Spatially implicit models are based on the observation that one can relate the sample area A to the total number of individuals J (Hubbell, 2001), given that $J = \rho A$, where ρ is the density. One can therefore obtain species-area curves in non-spatial models, looking at the scaling of the number of species S with the number of individuals J . Spatial implicit models can thus be thought of as mean-field “well-mixed” models, where one can neglect dispersal limitation.

In a meta-community the number of species $\langle \phi(n) \rangle$ of a population n is a Fisher log-series (see sec. II)

$$\langle \phi(n) \rangle = \theta \frac{r^n}{n}, \quad (61)$$

where r is the ratio of the birth to death rate. The average number of individuals in the meta-community is $J_M = \theta r / (1 - r)$ and the expected number of species can be easily

computed

$$\langle S(J_M) \rangle = \theta \sum_{n=1}^{\infty} \frac{r^n}{n} = -\theta \log(1+r) = \theta \log\left(1 + \frac{J_M}{\theta}\right). \quad (62)$$

This result corresponds exactly to that found using the random placement in (60). At small sample sizes, the number of species $J_M \ll \theta$ is equal to the number of individuals $\langle S(J_M) \rangle = J_M$. In other words when small areas are sampled the individuals belong to different species and the number of species grows along with the number of individuals. With larger sample sizes, the number of species $J_M \gg \theta$, grows logarithmically with the number of individuals. This approach allows one to calculate the SAR directly from the RSA distribution and it is clearly applicable to any RSA distribution.

Real ecosystems are of course spatially explicit, but one might wonder how spatially implicit models or, more generally, models that do not consider space explicitly, are predictive and how their parameters are related to spatially-explicit ones. A way to assess this is to measure the efficacy of non-spatial models in predicting the behavior of spatially explicit models (Etienne and Rosindell, 2011). As expected, non-spatial models have a good predictive power when the dispersal lengths are sufficiently large and they are particularly good in predicting non-spatial patterns such as the RSA.

2. Spatially explicit model

Spatially explicit models are typically defined as birth-death-diffusion processes. A model is fully specified given a ME and can be obtained in several ways, i.e. it is possible to write several different MEs that include space in a neutral model. The ME is not tractable analytically and one has to introduce approximations in order to get analytical results. Spatially explicit models are particularly difficult to solve because of the lack of detailed balance (see sec. II).

The voter model (Holley and Liggett, 1975) was originally introduced to describe opinion formation, whereby voters are located in a network and each one has one opinion among q possibilities. In ecological applications, voters become individuals and their opinion corresponds to the species they belong to (Durrett and Levin, 1996). An individual (voter) is chosen at random and is replaced by a copy of one of its neighbors (see Figure 8). This process has an absorbing state, because when a species disappears there is no way to introduce it again. In the long run, the system is populated by only one species. In order to overcome

this problem, one can introduce the possibility of new species entering the system (Durrett and Levin, 1996; Zillio *et al.*, 2005), a model we will refer to as the multispecies voter model with speciation (MVM).

Consider a lattice of d dimensions (for a typical ecological landscape $d = 2$), where a is the lattice spacing with exactly one individual at each site (a^d is the average volume occupied by one individual), and a total number of N individuals. At each time step, one individual chosen at random is removed and is replaced with a copy of one of its neighbors with a probability $1 - \nu$, or with an individual of a species not already present in the system with a probability ν . The case $\nu = 1$ is trivial, whereas the case $\nu = 0$ has an absorbing state, i.e. a state where a single species is present. The MVM is clearly a neutral model, the microscopic dynamics is the same for all species. It is also a zero-sum process, because the total number of individuals J is conserved. The mean field version of the voter model, where the neighbors of each node are all the other nodes, is governed by the ME (6), with the transition rates given by (20)-(21).

We want to write an equation for $F(r)$ (the probability that two individuals picked at random at a distance r belong to the same species). Assuming translational invariance, one obtains

$$F(\underline{r})^{t+1} = F(\underline{r})^t \left(1 - \frac{2}{N}\right) + \frac{1 - \nu}{dN} \sum_{\mu=1}^d \left(F(\underline{r} + \hat{e}_\mu)^t + F(\underline{r} - \hat{e}_\mu)^t \right), \quad (63)$$

where $F(\underline{r})^t$ is the probability that two individuals separated by \underline{r} at time t belong to the same species. The solution is obtained with the boundary condition $F(\underline{0})^t = 1$.

The stationary solution of (63) can be obtained from a Fourier series in the continuum limit, by taking the limit of $a \rightarrow 0$, $N \rightarrow \infty$ and $\nu \rightarrow 0$, and constraining $\gamma^2 = 2d\nu/a^2$ to be a constant. In this way one obtains (Zillio *et al.*, 2005) the second order differential equation:

$$\nabla^2 F(\underline{r}) - \gamma^2 F(\underline{r}) + c\delta^d(\underline{r}) = 0 \quad (64)$$

where δ^d is the d-dimensional Dirac delta, and whose solution is

$$F(r) = \frac{c\gamma^{d-2}}{(2\pi)^{d/2}} (\gamma r)^{\frac{2-d}{2}} K_{\frac{2-d}{2}}(\gamma r). \quad (65)$$

where $K_z(r)$ is the modified Bessel function and c is a constant such that $\int_{r < a} d^d r F(\underline{r}) = 1$ with a being the average distance between neighboring trees. By using the properties of

$K_z(r)$, one finds that in one dimension the solution is $F(r) = c\xi \exp(-r/\xi)/2$, with $\xi = \gamma^{-1}$. Interestingly, the exact solution in the (1d) discrete lattice is $F_x = \exp(-x/\xi)$ where $x = 0, a, 2a, \dots$ and

$$\frac{\xi}{a} = \left(\log \frac{1 - \nu}{1 - \sqrt{\nu(2 - \nu)}} \right)^{-1}, \quad (66)$$

with a being the lattice spacing. Eq.(66) converges to the equation $\xi = \gamma^{-1}$ in the continuum limit. The correlation length depends only on the speciation rate ν and it grows as $1/\sqrt{\nu}$ when $\nu \rightarrow 0$, a limit that is distinct from the special case $\nu \equiv 0$, in which at stationarity a single species is present across the whole finite lattice (monodominance). The β -diversity obtained in (65) was shown to be in good agreement with empirical data from different forest censa (Chave and Leigh, 2002). The prediction could be improved taking into account the Janzen-Connell effect (Zillio *et al.*, 2005).

It is possible to obtain interesting results on the SAR via extensive numerical simulations. In the case of NT, one can take advantage of the coalescent approach (Kingman, 1982; Rosindell *et al.*, 2008). Instead of simulating the stochastic dynamics directly, one can reconstruct the genealogy of the individuals in the sample area by regressing in time. The main value of this method is that there is no need to wait for any transient state to decay and therefore, this approach is much faster than forward dynamics (Rosindell *et al.*, 2008), as well as allowing infinite landscapes to be simulated.

The coalescent approach has been applied to the MVM with a different dispersal kernel (Rosindell and Cornell, 2007). Instead of a nearest-neighbor diffusion, once an individual is removed it is replaced by the offspring of another individual with a probability that depends on its distance from the individual removed. In an infinite landscape, the SAR shows the characteristic inverted- S shape. The model depends only on two parameters (up to a choice of the functional form of the dispersal kernel): the speciation rate ν and the dispersal length ξ . The SAR scales as

$$\langle S(A, \xi, \nu) \rangle = \xi^r S(A\xi^{-r}, \nu), \quad (67)$$

where the exponent r is independent of ν and very close to 2 (as expected by dimensional analysis). The exponent z of the power-law regime can be calculated as the derivative of the curve evaluated at the inflection point of the SAR (in log-log scale), i.e. the minimum value of $d \log(S)/d \log(A)$. Additional simulations (Pigolotti and Cencini, 2009), obtained

with a wider spectrum of speciation rates, suggests a logarithmic relationship

$$z = \frac{1}{q + m \log(\nu)} , \quad (68)$$

where q and m are two real parameters, which confirms the original prediction of (Durrett and Levin, 1996). This inverse logarithmic trend seems very robust and it has also been observed in other spatial models (Cencini *et al.*, 2012). Indeed, it was shown that power-law dispersal kernels better fit data than other short-ranged kernels (Rosindell and Cornell, 2009), whereby the estimation of the speciation rate corresponding to a given value of the exponent gives much smaller values.

An attempt to connect spatial and temporal patterns can be found (Bertuzzo *et al.*, 2011), where the persistence-time distribution was studied in MVM (see Sec. III.E). It was shown that the empirical persistence-time distribution is consistent with that predicted by MVM and

$$p(t|A) \sim t^{-\alpha} e^{-t/\tau(A)} , \quad (69)$$

where τ is the average time of persistence in an area A . The exponent α is universal and depends only on the dimension of the system, while the time scale τ is a function of the sampled area. Empirically, the time scale τ scales with the area as A^β . In the MVM $\tau = 1/\nu$. Using this fact and that the rate of appearance of a new species is $\lambda = \nu N \sim \nu A$, one can relate these quantities with the SAR. Indeed, the SAR is the product of the rate of appearance of new species and their average persistence time.

$$S(A) = \lambda \langle t \rangle \sim A^{1-\beta(\alpha-1)} . \quad (70)$$

Here we have assumed that $\langle t \rangle \sim \tau^{2-\alpha}$ to obtain a scaling relationship that connects the exponent z and the other exponents: $z = 1 - \beta(\alpha - 1)$. This allows spatial patterns to be connected with temporal ones in a manner consistent with the empirical patterns shown in Figure 6.

C. Environmental Fragmentation and Habitat Loss

Resources are not equally distributed and, even over small scales, their distribution in space is far from uniform. This spatial heterogeneity affects the distribution of individuals and species in space, and has clear implications for the conservation of ecosystems. The

space within which ecosystems are embedded is often fragmented and a species may be present in only part of the landscape. Moreover, different patches are not independent but they are rather connected via immigration.

In the absence of speciation, the voter model (Holley and Liggett, 1975) predicts monodominance in dimension $d \leq 2$. Spatial heterogeneity can be modeled as quenched disorder (Borile *et al.*, 2013) and when two species are considered, it is postulated that different locations on a lattice prefer one species over the other. At each site i , a binary variable σ_i resides that takes values of ± 1 , depending on which species is present. Spatial heterogeneity can be modeled as a quenched external field τ_i , which also takes values ± 1 , and the dynamics are fully specified by the transition rates

$$W[\sigma_i \rightarrow -\sigma_i] = \frac{1 - \epsilon \tau_i \sigma_i}{2z} \sum_{j \in \partial i} (1 - \sigma_i \sigma_j) , \quad (71)$$

where ∂i is the set of nearest neighbors of i , and z is the size of this set. The quantity ϵ measures the strength of the preference. The main result from this model (Borile *et al.*, 2013) was that this randomness enhances the coexistence of species. Indeed, if ϵ is larger than $\epsilon_c = \sqrt{2/(2+N)}$, species coexist in any dimension in the limit of a large system. One can introduce the quantity $\phi = (1/N) \sum_i \sigma_i$ and if coexistence occurs $\langle \phi^2 \rangle < 1$, whereas this quantity tends to one otherwise. It was shown that (Borile *et al.*, 2013), for small ϵ

$$P(\phi) \propto \frac{\exp(-\frac{N}{2} \frac{\epsilon^2}{1-\epsilon^2} \phi^2)}{(1-\epsilon^2)(1-\phi^2) + 2\nu} , \quad (72)$$

where a small mutation rate ($\nu \ll 2/(2+N)$) was introduced to regularize the solution. If $\epsilon > \epsilon_c$, then the average $\langle \phi^2 \rangle < 1$, i.e. coexistence is stable.

NT has also been applied to predict the extinction rate of species after habitat loss (Halley and Iwasa, 2011). Habitat loss corresponds to a reduction of the area available and therefore, to the total number of individuals. When this area is destroyed, the endemic species suddenly disappear and what follows is a delayed series of extinctions due to habitat loss. The community was modeled as a neutral assembly of species and a typical time scale of extinctions was obtained, along with its dependence on the number of species and the habitat destroyed. The predictions obtained in this way reproduced the available data of avifaunal extinctions well (Halley and Iwasa, 2011).

V. BEYOND NEUTRALITY.

A. Reconciling Neutral and Niche Theory

The concept of niche is central in classical ecology (Chase and Leibold, 2003; Chesson, 2000; MacArthur and Wilson, 1967). An ecological niche is “the requirements of a species for existence in a given environment and its impacts on that environment” (Chase and Leibold, 2003) and it describes how an organism or a population respond to changes in resources, competitors and predators. A possible mathematical realization of the concept of niche is the Hutchinsonian niche, which involves a n -dimensional hypervolume, where the axes are environmental conditions or resources. A position in this space represents a set of behaviors and traits characterizing a species or a group of individuals. In niche theory much relevance is given to the specific traits of species and their interdependence. A central concept in niche theory is competitive exclusion, which states that two species cannot occupy the same niche, as two identical, yet distinct species, cannot co-exist for an indefinite time.

The mathematical representation of an ecological community that includes niche aspects typically coincides with the Lotka-Volterra equations. In this case, the focus is on the properties of the fixed points (or other dynamical attractors) of these systems of equations and the typical problem that is analyzed is their stability, in relation to the parameters and the species present in the system.

The main difference between neutral and niche theory therefore depends on which mechanism plays the main role in shaping ecosystems (Jeraldo *et al.*, 2012). Neutral theory assumes that random processes, such as dispersal, demographic stochasticity, speciation and ecological drift, have a stronger impact on many of the observed patterns than niche differences. Niche theory assumes the opposite, that the quantitatively important processes are related to differences in species and their interdependence.

As we might expect in a real ecosystem, both stochasticity and niche differences play a role, and it is natural to try to quantify how neutral behavior emerges from a niche model. In many cases, niche-based and neutral models yield compatible fits of biodiversity patterns (Chave *et al.*, 2002; McGill, 2003; McGill *et al.*, 2006; Mouquet and Loreau, 2003), and it’s impossible to distinguish between the mechanisms by looking at those patterns. As pointed out in (Adler *et al.*, 2007) neutrality emerges when species have the same or very similar

fitness.

Neutrality has often been proposed to emerge under some conditions from models considering niche differences (Gravel *et al.*, 2006; Haegeman and Loreau, 2011; Noble *et al.*, 2011a), and neutrality and niche theory were proposed to be the extremes of a more general model (Gravel *et al.*, 2006; Haegeman and Loreau, 2011). In both cases, a Lotka-Volterra equation is introduced to describe community dynamics, while population dynamics is modeled by also taking into account demographic stochasticity and immigration. By considering different values of parameters, one can move from a scenario where species differences matter a lot and the stable configuration is very close to the solution of the deterministic Lotka-Volterra equation, to a scenario where demographic stochasticity is more important and the community behaves like a neutral community. A slightly different approach has also been considered (Fisher and Mehta, 2014), analyzing a stochastic version of Lotka-Volterra dynamics and quantifying, when stochasticity is varied, the difference between the prediction of the neutral model with the full model. In this case, instead of a continuum of strategies, neutral and niche regimes are two macroscopic phases separated by a phase transition. Other authors have also incorporated neutral and non-neutral features within the same framework (Etienne and Olf, 2004a; Kessler and Shnerb, 2015): species with comparable sizes were considered functionally equivalent, thereby entailing a neutral dynamics; and parameters such as speciation, dispersal and populations of organisms of distinct sizes were obtained from allometric scaling laws. With this approach, Etienne & Olf were able to explain why species richness reaches a maximum at intermediate body size (Etienne and Olf, 2004a).

A different approach was based on effectively considering niche theory as a model where per capita death/birth rates are species dependent (Borile *et al.*, 2012). In this case, one might expect the difference between species abundances to reflect the differences between these parameters. More precisely, in a neutral scenario, all the species fluctuate around a given abundance value, while when niche characteristics play a role, each species fluctuates around its own distinct value of abundance. A third scenario has been proposed (Borile *et al.*, 2012), wherein the per-capita birth/death rates do not have a monotonic effect on species abundance. The symmetry between species, due to the neutrality of the process, can yet be spontaneously broken. In this case, the stable states are not symmetric in terms of species abundance.

So far macro-evolutionary patterns such as phylogenetic trees have not been a major

focus in NT. However, there is evidence that NT is not able to satisfactorily capture phylogenetic diversity (Davies *et al.*, 2011). Multiple patterns of evolutionary history in bacterial communities seem to deviate from the predictions of NT (ODwyer *et al.*, 2015). However, recent models, which incorporate either neutral population dynamics into the cladogenesis (Manceau *et al.*, 2015) or a mild selection into an otherwise neutral model (Rosindell *et al.*, 2015) are able to slowdown the diversification process as well as match phylogenies observed in nature. These represent promising developments, which can improve and broaden the spectrum of predictions of NT.

B. Emergent neutrality

Ecologists have highly criticised NT because of its unrealistic assumptions. The patterns that we have studied so far can in fact be explained without introducing species differences, and this has led some ecologists to oppose NT because it assumes that nature is actually governed by neutral processes, whereas it is not. Clearly, no one believes that nature is truly neutral. The patterns of community ecology are actually generated by a cocktail of processes, and it is both inappropriate and dangerous to consider processes in isolation from a macro-ecological pattern or empirical data set.

However, it is informative to study whether, how and which ecological processes can drive a community towards a state in which demographic stochasticity and immigration play a crucial role in the face of strong species' differences that are dictated by classical competitive exclusion. Such a state should allow similar species to emerge in the niche space with the ability to co-exist for sufficient time. Indeed, when this problem was studied, it was shown that species can evolve into groups of relatively more similar species that co-exist for very long times (Scheffer and van Nes, 2006). First, a large number of species were placed at random along a hypothetical niche axis, which represents a specific trait, assuming that interspecific competition can be calculated through niche overlap. Running a classical Lotka-Volterra competition model and studying evolution, groups of multiple species were evident that aggregated around similar values in the niche axis and they could co-exist for many generations before the majority of them head towards an inexorable extinction. Eventually, only one species survives from each group, producing the expected pattern of single species equally spaced in the niche space.

In other words, the niche similarity of species prevents competitive exclusion from swiftly selecting the best competitor among a group of similar species, allowing their co-existence for very long times even though only the superior species will ultimately persist.

This model may be considered as one of the possible steps towards a reconciliation of niche and neutral theories. Species that are initially ecologically non-equivalent, and that therefore behave in a non-neutral fashion, are driven by community and evolutionary processes towards states in which the dynamics may well be better approximated by neutral models over appropriate spatial and temporal scales. More recently, further support for this approach came from showing that the model can produce multimodal RSAs (Vergnon *et al.*, 2012). Immigration may also be an important component in the neutral-like behaviour of communities (Gravel *et al.*, 2006; Holt, 2006). Parasitoids competing for a common species (Bonsall *et al.*, 2004) have been used to show that clusters of species separated by gaps emerge along the niche axis, confirming — using a quite different approach — that processes exist that can lead community dynamics to be effectively neutral (Purves *et al.*, 2005).

The basic Lotka-Volterra model has been extended to investigate the possibility that some processes decrease the risk of competitive exclusion so that species lumps are not only transient, but ultimately permanent (Scheffer and van Nes, 2006). Density-dependent regulation was introduced that stabilizes the co-existence of species within a group. This approach unfortunately has a drawback that the mechanism introduces a discontinuity in the competition strength of the species (Barabás *et al.*, 2013), which means that unmodeled species differences may be responsible for co-existence in the community. However, it has been shown that more realistic density dependent terms, or even other mechanisms (e.g., migration (Vergnon *et al.*, 2013a)), can eliminate this problem, making the approach more robust (Vergnon *et al.*, 2013b).

C. Maximum Entropy Models

The maximum entropy principle is a useful method to obtain the least biased information from empirical measurements (Jaynes, 2003). This powerful tool, borrowed from statistical mechanics, has a wide range of applications (Banavar *et al.*, 2010), including ecology (Banavar *et al.*, 2010; Harte, 2011). In its ecological application, the Max Ent principle is an inference method (Chayes *et al.*, 1984) used to evaluate the effective strength of interactions

among species based on either species abundance data (Volkov *et al.*, 2009) or simply the presence/absence of the species (Azaele *et al.*, 2010). This methodology provides a way to systematically incorporate the most important species interactions into the development of a theory beyond the purely non-interacting case. In addition, the Max Ent principle was implemented as a method to predict biodiversity patterns across different spatial scales using only the information on local interactions (Adorasio *et al.*, 2014; Harte, 2011). Here we will discuss the extension of Max Ent to the study of spatial biodiversity patterns. Consider an ecosystem in which S species (belonging to the same trophic level, see BOX 2) live within a given area A , divided into N adjacent sites of equal area. Let us assume that there are empirical records of the species contained within each site. From these data one can calculate, for example, the average presence of any species in the ecosystem and the co-occurrence of any pair of species in neighboring sites. When applying Max Ent, we can consider these mean occurrences and co-occurrences as given constraints. We introduce the binary random variable σ_i^α , which records the occurrence of species $\alpha = 1, \dots, S$ at each site $i = 1, \dots, N$. If species α is present in plot i , then $\sigma_i^\alpha = 1$, otherwise $\sigma_i^\alpha = 0$. Therefore the “state” of any species α can be characterized by the random vector $\boldsymbol{\sigma}^\alpha$. There is empirical evidence that species belonging to the same trophic level interact weakly (Azaele *et al.*, 2010; Veech, 2006; Volkov *et al.*, 2009) and therefore, in a first approximation, one may assume that species occur in a given geographical location independently of one another. Due to such independence, the probability of finding the system in the configuration $\boldsymbol{\sigma} = (\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \dots, \boldsymbol{\sigma}^S)$ is thus $P(\boldsymbol{\sigma}) = \prod_{\alpha=1}^S p_\alpha(\boldsymbol{\sigma}^\alpha)$, where p_α gives the probability distribution of finding a species α in the configuration $\boldsymbol{\sigma}^\alpha$. In order to build the Max Ent model, we will maximize the Shannon’s entropy $H = -\sum_{\boldsymbol{\sigma}^1} \dots \sum_{\boldsymbol{\sigma}^S} P(\boldsymbol{\sigma}) \ln P(\boldsymbol{\sigma}) = -\sum_\alpha \sum_{\boldsymbol{\sigma}^\alpha} p_\alpha(\boldsymbol{\sigma}^\alpha) \ln p_\alpha(\boldsymbol{\sigma}^\alpha)$, imposing the average occurrence constraint, i.e. $\langle M \rangle = \sum_\alpha \sum_{\boldsymbol{\sigma}^\alpha} p_\alpha(\boldsymbol{\sigma}^\alpha) M_\alpha(\boldsymbol{\sigma}^\alpha)$ with $M_\alpha(\boldsymbol{\sigma}^\alpha) = \sum_i \sigma_i^\alpha$, and the average co-occurrence constraint, i.e. $\langle E \rangle = \sum_\alpha \sum_{\boldsymbol{\sigma}^\alpha} p_\alpha(\boldsymbol{\sigma}^\alpha) E_\alpha(\boldsymbol{\sigma}^\alpha)$ with $E_\alpha(\boldsymbol{\sigma}^\alpha) = \sum_{(i,j)} \sigma_i^\alpha \sigma_j^\alpha$, where (i, j) indicates two nearest neighbour locations. Both $\langle M \rangle$ and $\langle E \rangle$ are meant to match the corresponding empirical averages, as calculated from the empirical records of species contained in the region. Maximization (Adorasio *et al.*, 2014) provides an explicit expression for the probability, p_α , of finding a species α in the configuration $\boldsymbol{\sigma}^\alpha$. Thus,

$$p_\alpha(\boldsymbol{\sigma}^\alpha) = \frac{e^{J_\alpha E_\alpha(\boldsymbol{\sigma}^\alpha) + h_\alpha M_\alpha(\boldsymbol{\sigma}^\alpha)}}{Z_\alpha}, \quad (73)$$

where $Z_\alpha = Z_\alpha(h_\alpha, J_\alpha)$ is the partition function. From (73) one can characterize the spatial biodiversity patterns of the ecosystem, i.e. calculate the SAR or the EAR (Adorisio *et al.*, 2014). This type of approach may well be suited to infer biodiversity properties of communities over larger spatial scales by upscaling the model results at local scales. However, only a few studies have tackled this problem (see below).

VI. OPEN PROBLEMS AND FUTURE CHALLENGES

In this section, we consider how one might advance theory in terms of adding essential details without making the system unnecessarily complex or having an explosion of species-specific parameters. What is the relative importance of species traits, their interactions, and spatial and environmental effects? Throughout this review, we have treated the total number of species and the total number of individuals as parameters which are provided as input. How would one determine and predict these parameters? A mechanistic explanation is needed for why one region may be more bio-diverse than another, what sustains biodiversity, and how evolutionary pressures sculpt ecological communities.

A. Analytical spatially explicit models.

At the moment, we do not have a spatially explicit theory that can predict analytically the most important ecological patterns. This would be important, because it would allow us to understand what drives biodiversity across spatial and temporal scales. Indeed, it is likely that biological processes are not equally important across scales. As evident in particle physics, we might eventually find that ecosystems will need to be described by different effective theories according to the range of scales in which we are interested. In a pioneering paper, which received great interest, a ME was proposed that seemed to permit analytical calculations of the SAR (O'Dwyer and Green, 2010). Indeed, an analytical expression for the SAR was obtained. However, the promised solution was neither correct and nor was it an approximation of the actual SAR (Grilli *et al.*, 2012a). The main technical difficulty

preventing an analytical solution was that the model was inherently out-of-equilibrium. Detailed balance, which is the condition that makes the calculation of stationary probabilities possible, did not hold (Grilli *et al.*, 2012a). The time is ripe for explorations of this kind. The challenge is to develop new and powerful techniques in non-equilibrium statistical mechanics. Other authors have applied NT to study biodiversity patterns on river networks. By exploiting the network connectivity of such dendritic landscapes (Carrara *et al.*, 2012), they were able to explain several empirical large-scale spatial biodiversity patterns (Carrara *et al.*, 2014; Muneeppeerakul *et al.*, 2008) and show how changes in riverine ecosystems may impact the spread of species and local species richness (Lynch *et al.*, 2011; Mari *et al.*, 2014).

B. Linking different Macro-Ecological Patterns

Over the last few decades, ecologists have come to appreciate the importance of spatial patterns and processes, and the explicit introduction of space has the potential to revolutionize what we know about natural populations and communities (Storch *et al.*, 2007; Tilman and Kareiva, 1997). It has become apparent that key ecological patterns, such as SAR, RSA and spatial patterns of species distributions and turnover, are intimately intertwined and scale dependent (see Fig.1). However, despite a plethora of models that address spatial patterns, only a few practical methods have been proposed to link them across different scales. Specifically, spatial approaches (McGill and Collins, 2003; Plotkin *et al.*, 2000) lack the needed analytical machinery, whereas most theoretical approaches are not spatially explicit or sufficiently flexible (Harte *et al.*, 2008; Volkov *et al.*, 2009). Even neutral theory was conceived to reflect the idealized behavior of natural systems at equilibrium, rather than to reflect non-pristine landscapes produced by environmental change or management (Hubbell, 2001). Therefore, a general methodology is required to predict and link these ecological patterns across scales that is sufficiently robust and flexible to allow its application to a range of natural or managed systems. One possible way to tackle this challenging problem is through a theoretical framework inspired by ideas coming from phenomenological renormalization (Azaele *et al.*, 2015). The fundamental assumption at the core of this theoretical setting is that the functional form of the RSA remains the same across all spatial scales, even though the parameters of the curve are likely to vary. Because of this assumption, the spatial dependence of the abundance distribution can be obtained by making the RSA

parameters suitable functions of scale. Together with the functional shape of the RSA, the other model input is the spatial pair correlation function (PCF), which describes the correlation in species' abundances between pairs of samples as a function of the distance between them (Azaele *et al.*, 2009; Zillio *et al.*, 2005). If populations were randomly distributed in space, distinct communities would on average share the same fraction of species regardless of their spatial separation, and therefore, the PCF would not depend on distance. In contrast, in highly aggregated communities correlations in abundance would fall off steeply with increasing distance. The PCF not only measures the rate of turnover in species composition but it also reflects the variation of population clustering across scales, given that the variance in species abundances at any particular scale can be calculated directly from the PCF (Azaele *et al.*, 2015). Therefore, the PCF is related to the spatial species abundance distribution. Thus, the PCF can link the effects of aggregation, similarity decay, species richness and species abundances across scales. Building on the intrinsic relationship among these patterns, while accounting for spatial correlations with fidelity, is critical for predicting the biodiversity profiles across scales when information on a limited number of fine-scale scattered samples is available (Slik *et al.*, 2015; Ter Steege *et al.*, 2013). The explosion of publicly available large scale biodiversity data for paradigmatic ecosystems such as the Amazonian forest (ter Steege *et al.*, 2015) and ocean plankton (Bork *et al.*, 2015) makes this problem one of the most exciting scientific challenges in this field.

C. Environmental noise

Most neutral (symmetric) models are solely governed by the underlying stochastic birth-death process and do not take into account any environmental stochasticity (Kalyuzhny *et al.*, 2014). Recently it has been shown that environmentally induced variations of the demographic rates dominate the long-term dynamics and have an important impact on some dynamic properties (such as age-size relationships and species extinction time (Kessler *et al.*, 2015)), while not affecting the already good accord of neutral models with ecological static patterns (such as the RSA) (Kalyuzhny *et al.*, 2015). However, analytical results on neutral (symmetric) models for ecological communities subject to correlated environmental noise are mainly missing (although some recent results have been obtained for white noise (Kessler *et al.*, 2015; Melbinger and Vergassola, 2015)), incentivizing physicists to explore

these paths.

D. Multi-Trophic Ecosystems: Ecological Networks

Although we are progressing in understanding the suitability and limits of non-interacting and non-spatial models, most neutral models still assume that species interact randomly with each other. However, a network approach to modelling ecological systems provides a powerful representation of the interactions among species (Bastolla *et al.*, 2009; Krishna *et al.*, 2008; Suweis *et al.*, 2013). Ecological networks may be viewed as a set of different species (nodes) and connections/links (edges) that represent interspecific interactions (e.g., competition, predation, parasitism and mutualism). The architecture of ecological interaction networks has become a bubbling area of research, and it seems to be a critical feature in shaping and regulating community dynamics and structure diversity patterns (Allesina *et al.*, 2015; Allesina and Tang, 2012). An important step relevant for multi-trophic systems will be to obtain a general framework within which a network of preferences/disfavors modifies the birth and death rates of different species and can be superposed on neutral models (like the voter model presented in this review). Recent studies of ecological networks have considered the exciting task of anticipating critical transitions in such systems and to design structures that are less vulnerable to collapse (Scheffer *et al.*, 2012; Suweis and D’Odorico, 2014). Connecting stochastic quasi-neutral models, ecological networks and critical transitions within a unified theoretical framework is an important challenge as it will enhance our capacity to understand and thus manage the crucial interplay between ecological dynamics and species interactions.

VII. CONCLUSIONS

In this paper, we have attempted to describe some of the theoretical frameworks that can be used to understand key issues related to biodiversity and that will serve to address important questions. These frameworks are necessarily elementary and incomplete, yet they have the advantage of being tractable and related to the central issues. Unlike standard approaches to more traditional physics, here the Hamiltonian function or the interactions among the components are completely unknown and even identifying the state variables

may sometimes be a non-trivial task. We have a fleeting picture of what an ecosystem is and it is not necessarily in equilibrium. We have little knowledge of the myriad degrees of freedom and their interactions. The real challenge is to discern the most essential degrees of freedom and to develop a framework to understand and predict the emergent ecological behaviour.

The basic message of this review is that to resolve these challenging problems, ideas and techniques must be recruited from different disciplines. We are still at the beginning of this adventure. Moving forward is not only important but it is also urgent. The pressures of habitat destruction, pollution and climate change are having highly undesirable consequences on the health of ecological communities. To address practical issues related to conservation biology, we need models that can be used across scales in order to extrapolate information on biodiversity from accessible regions to inaccessible yet important scales. There is plenty of room for ideas that matter, and community ecology can greatly benefit from the contribution of other disciplines, including physics.

VIII. ACKNOWLEDGMENTS

We have immensely benefited from discussions with Stephen Hubbell, Stefano Allesina, David Alonso, Enrico Bertuzzo, Ryan Chisholm, Stephen J. Cornell, Paolo D’Odorico, William F. Fagan, Andrea Giometto, William E. Kunin, Egbert Leigh, Alan J. McKane, Miguel A. Muñoz, Andrew Noble, James P. O’Dwyer, Andrea Rinaldo, Ignacio Rodriguez-Iturbe, James L. Rosindell and Nadav Shnerb. S.S. acknowledges the University of Padova Physics and Astronomy Department Senior Grant 129/2013 Prot. 1634.

Appendix A: Density dependence, and assembling of local communities.

In this appendix, we present an alternative method to introduce density dependence and how ecosystems emerge by assembling local communities.

First, one can set $b_k(n) = b \cdot (n + \Upsilon_k)$ (for the k -th species) and $d(n) = dn$, where Υ_k incorporates both the effects of intra-specific interactions, such as those giving rise to density dependence, and the immigration occurring from a meta-community. This approach can be applied to two distinct ecosystems: coral reefs and tropical forests.

The steady-state solution of the ME for $P_k(n)$ yields a negative binomial distribution:

$$P_k(n) = \frac{(1 - r_k)^{\Upsilon_k}}{\Gamma(\Upsilon_k) \frac{r_k^n}{n!} \Gamma(n + \Upsilon_k)} \quad (\text{A1})$$

with a mean $\langle n_k \rangle = r_k \Upsilon_k / (1 - r_k)$, where Γ is the gamma function.

The number of species containing n individuals is given by $\phi_n = \sum_{k=1}^S I_{n,k}$, where $I_{n,k}$ is a random variable that is 1 with a probability $P_k(n)$ and 0 with a probability $1 - P_k(n)$.

Thus, the RSA is given by

$$\langle \phi_n \rangle = \sum_{k=1}^S I_{n,k} = \sum_{k=1}^S P_k(n) = \theta \frac{r^n}{n!} \Gamma(n + \Upsilon), \quad (\text{A2})$$

where $\theta = S / [(1 - r)^{-\Upsilon} - 1] \Gamma[\Upsilon]$ is the biodiversity parameter (Hubbell, 2001), and we have dropped the k dependence because of the symmetric hypothesis. For a small Υ , the RSA for the communities resembles the Fisher log-series and it does not have an interior mode.

Note that a non-trivial k dependence might arise even under the neutral hypothesis. For example, one can set $\Upsilon_k = \tilde{m} p_k$, where \tilde{m} is a measure of the immigration rate from the meta-community, in units of the birth rate b and p_k is the fraction of individuals in the surrounding meta-community belonging to the k -th species. As a result, one can obtain the following RSA for the community of tropical forests.

$$\langle \phi_n \rangle = \theta \frac{x^n}{n!} \int_0^\infty \frac{\Gamma(y + n)}{\Gamma(y + 1)} e^{-\omega y} dy \equiv \theta \frac{x^n}{n!} f(n, \omega), \quad (\text{A3})$$

where $\omega = \frac{\theta}{\tilde{m}} - \ln(1 - x)$. This approach provides a virtually indistinguishable fit to the empirical data as $\langle \phi_n \rangle = \theta \frac{x^n}{n+c}$ considered earlier with the advantage of having ecologically meaningful parameters.

The average number of species observed in the local community is

$$S_{obs} = \langle S \rangle = S - \langle \phi_0 \rangle = \sum_{k=1}^S (1 - r)^{-\Upsilon} = S [1 - (1 - r)^\Upsilon] \quad (\text{A4})$$

If the sample considered has J_L individuals, and thus the community dynamics obeys the zero-sum rule (i.e. it has a fixed total population), then the multivariate probability distribution is

$$P(\mathbf{n} | J_L) = \mathcal{N} \prod_{k=1}^S P_k(n) \delta(J_L - n_1 - n_2 - \dots - n_S) = \binom{J_L + \sum_k \Upsilon_k - 1}{J_L}^{-1} \prod_{k=1}^S \binom{n_k + \Upsilon_k - 1}{n_k} \delta(J_L - n_1 - n_2 \dots) \quad (\text{A5})$$

that is a compound multinomial Dirichlet distribution where \mathcal{N} is the normalization constant.

Meta-community Composition. Now, let us gradually assemble the meta-community of coral reefs by considering it as an assemblage of local communities. Let us start by considering the joint RSA of two local communities, A and B, with n_A and n_B individuals, respectively. The probability that the species has n individuals in the meta-community formed by A and B is (Volkov *et al.*, 2007)

$$P(n = n_A + n_B) = \sum_{n_A + n_B = n} P(n_A)P(n_B) \propto \frac{r^n}{n!} \Gamma(n + 2\Upsilon), \quad (\text{A6})$$

where the actual spatial locations of the local reef communities have been neglected (all local communities are well-mixed in the meta-community, i.e. a mean field approximation). The elegant result of the Υ 's adding to each other follows ecologically from their interpretation as immigration rates. For the meta-community, we can introduce speciation with a rate $\nu \ll 1$ as we have seen it has a crucial role when $n = 0$ (under the assumption of neutrality, the species label of the new species is of no consequence), i.e. $\Upsilon_k = \nu$ and $P(n_k) = \nu r^n / n + O(\nu^2)$.

Extending the calculation of the joint RSA distribution to more and more local communities, it can be shown that the RSA of the meta-community is characterized by an effective immigration parameter $L\Upsilon$, where L is the total number of local communities comprising the meta-community, and it becomes log-normal-like if $L \gg 1$, in agreement with the available data (Volkov *et al.*, 2007).

A Fisher log-series is observed in two limiting cases - in the meta-community in which there are no immigration events and in the very small local community that has a high immigration rate from the meta-community characterized by a Fisher log-series RSA.

Appendix B: Persistence or lifetime distributions

In this appendix, we present the derivation for the persistence (or lifetime) distribution that leads to the empirical choice made in (48). The master equation we want to solve is

$$\frac{\partial P(n, t)}{\partial t} = b(n-1)P(n-1, t) + d(n+1)P(n+1, t) - (b(n) + d(n))P(n, t), \quad (\text{B1})$$

where the birth rate is $b(n) = (1-\nu)n/J_M$ with $b(-1) = 0$, and the death rate is $d(n) = n/J_M$ and $n \geq 0$. We can redefine the time scale $t \rightarrow J_M t$ so that the factor $1/J_M$ disappears from

the birth and death rates. As the initial condition we choose $P(n, t = 0) = \delta_{n, n_0}$ with the initial population as n_0 . Our goal is to calculate the survival probability defined as

$$\mathcal{P}(t|\nu) = \sum_{n \geq 1} P(n, t) = 1 - P(n = 0, t) . \quad (\text{B2})$$

Thus, we introduce the generating function

$$G(z, t) = \sum_{n \geq 0} P(n, t) z^n . \quad (\text{B3})$$

whose radius of convergence is ≥ 1 . Note that $G(z = 1, t) = 1$ and $G(z = 0, t) = P(n = 0, t) = 1 - \mathcal{P}(t|\nu)$. Using Eq. (B1) the time evolution of the generating function is derived immediately

$$\frac{\partial G(z, t)}{\partial t} = [(1 - \nu)z - 1](z - 1) \frac{\partial G(z, t)}{\partial z} . \quad (\text{B4})$$

with the initial condition $G(z, t = 0) = z^{n_0}$. The previous equation is a linear partial differential equation and it can be solved by standard methods. One introduces a function $Z(\tau)$ satisfying the time evolution equation

$$\frac{dZ(\tau)}{d\tau} = -[(1 - \nu)Z(\tau) - 1](Z(\tau) - 1) , \quad (\text{B5})$$

with a "final" condition $Z(\tau = t) = z$. Using Eq. (B4), one is led to

$$\frac{dG(Z(\tau), t)}{d\tau} = 0 , \forall \tau , \quad (\text{B6})$$

implying that

$$G(z, t) = G(Z(t), t) = G(Z(0), 0) = Z(0)^{n_0} . \quad (\text{B7})$$

The solution of eq. (B5), with the chosen "final" condition gives

$$Z(\tau = 0) = \frac{1 - A(z, t)}{1 - (1 - \nu)A(z, t)} \quad \text{with} \quad A(z, t) = \frac{1 - z}{1 - (1 - \nu)z} e^{-\nu t} \quad (\text{B8})$$

and thus,

$$G(z, t) = \left(\frac{1 - A(z, t)}{1 - (1 - \nu)A(z, t)} \right)^{n_0} . \quad (\text{B9})$$

Finally, we get the survival probability

$$\mathcal{P}(t|\nu) = 1 - G(z = 0, t) = 1 - \left(1 + \nu (e^{\nu t} - 1)^{-1} \right)^{-n_0} . \quad (\text{B10})$$

In the scaling, if we consider the limit of fixed νt as t becomes large, we get the scaling form

$$\mathcal{P}(t|\nu) = \frac{1}{t} F(\nu t) , \quad (\text{B11})$$

with

$$F(x) = n_0 \frac{x}{e^x - 1} . \quad (\text{B12})$$

The lifetime distribution is simply given by

$$p_\tau(t) = -d\mathcal{P}(t|\nu)/dt = \frac{1}{t^2} f(\nu t) , \quad (\text{B13})$$

where the last equality holds in the scaling limit as

$$f(x) = n_0 e^x \left(\frac{x}{e^x - 1} \right)^2 . \quad (\text{B14})$$

which for $x \rightarrow 0$ tends to a non-zero constant. As $x \rightarrow \infty$, $f(x)$ decays exponentially, leading to the empirical scaling form of the kind given by (48).

Appendix C: The species turnover distribution with absorbing boundary conditions

In this appendix we present another temporal pattern predicted by the model defined by (39) or (41). For an ecosystem at or near stationarity, the model can provide the exact expression for the STD when using the time-dependent absorbing solution of (39): the so-called the species turnover distribution with absorbing boundary conditions. Unlike the reflecting species turnover distribution, this distribution corresponds to a new kind of measure that only accounts for the species present at the initial time, and it does not take into account any new species introduced by immigration/speciation or any old species that reappear after their apparent extinction until $t > 0$. This amounts to the selection of a particular sample and the study of the temporal behavior of those selected individuals. The species turnover distribution with absorbing boundary conditions, $\mathcal{P}^{abs}(\lambda, t)$, can be obtained through formula (46) where the conditional probability must now be $p_a(x, t|x_0, 0)$ defined as in (49). The final expression is

$$\begin{aligned} \mathcal{P}^{abs}(\lambda, t) &= \frac{\sin(\pi b/D)}{\pi(1-b/D)} \frac{e^{-t/\tau} (e^{t/\tau} - 1)^{b/D}}{(\lambda + 1)^2} \times \\ &\times {}_2F_1 \left(1, \frac{3}{2}, 2 - b/D; \frac{4\lambda e^{-t/\tau}}{(\lambda + 1)^2} \right) , \end{aligned} \quad (\text{C1})$$

where $b/D < 1$. Note that $\mathcal{P}^{abs}(\lambda, t)$ decays exponentially to zero at a rate $(1 - \beta)/\tau$, regardless of λ .

It is noteworthy that the STD's with absorbing and reflecting boundaries are indistinguishable whenever $t \ll \tau$ and λ are not too small (for $b/D < 1$). Since the BCI data are sampled over relatively short times intervals (at most 10 years), the distributions are almost the same within the interval $1/2 < \lambda < 3/2$.

REFERENCES

- Adler, P. B., J. HilleRisLambers, and J. M. Levine (2007), *Ecology letters* **10** (2), 95.
- Adorisio, M., J. Grilli, S. Suweis, S. Azaele, J. R. Banavar, and A. Maritan (2014), arXiv preprint arXiv:1407.2425.
- Allen, A. P., and V. M. Savage (2007), *Ecology letters* **10** (7), 637.
- Allesina, S., J. Grilli, G. Barabás, S. Tang, J. Aljadeff, and A. Maritan (2015), *Nature communications* **6**, 7842.
- Allesina, S., and S. Tang (2012), *Nature* 10.1038/nature10832.
- Alonso, D., R. S. Etienne, and A. J. McKane (2006), *Trends in Ecology & Evolution* **21** (8), 451.
- Alonso, D., and A. J. McKane (2004), *Ecology Letters* **7** (10), 901, 853XA Times Cited:70 Cited References Count:50.
- Alonso, D., A. Ostling, and R. S. Etienne (2008), *Ecology Letters* **11** (2), 93, 249DI Times Cited:27 Cited References Count:56.
- Arrhenius, O. (1921), *Journal of Ecology* **9** (1), 95.
- Azaele, S., A. Maritan, S. J. Cornell, S. Suweis, J. R. Banavar, D. Gabriel, and W. E. Kunin (2015), *Methods in Ecology and Evolution*.
- Azaele, S., R. Munepeerakul, A. Maritan, A. Rinaldo, and I. Rodriguez-Iturbe (2009), *Proceedings of the National Academy of Sciences* **106** (17), 7058.
- Azaele, S., R. Munepeerakul, A. Rinaldo, and I. Rodriguez-Iturbe (2010), *Journal of theoretical biology* **262** (2), 323.
- Azaele, S., S. Pigolotti, J. R. Banavar, and A. Maritan (2006), *Nature* **444** (7121), 926.
- Babak, P. (2006), *Physical Review E* **74** (2), 021902.
- Babak, P., and F. He (2009), *Theoretical Population Biology* **75** (1), 76.
- Banavar, J. R., A. Maritan, and I. Volkov (2010), *Journal of Physics: Condensed Matter* **22** (6), 063101.

- Barabás, G., R. DAndrea, R. Rael, G. Meszéna, and A. Ostling (2013), *Oikos*.
- Bastolla, U., M. a. Fortuna, A. Pascual-García, A. Ferrera, B. Luque, and J. Bascompte (2009), *Nature* **458** (7241), 1018.
- Bell, G. (2000), *The American Naturalist* **155** (5), 606.
- Bertuzzo, E., S. Suweis, L. Mari, A. Maritan, I. Rodríguez-Iturbe, and A. Rinaldo (2011), *Proceedings of the National Academy of Sciences* **108** (11), 4346.
- Black, A. J., and A. J. McKane (2012), *Trends in Ecology & Evolution* **27** (6), 337.
- Blythe, R. A., and A. J. McKane (2007), *Journal of Statistical Mechanics: Theory and Experiment* **2007** (07), P07018.
- Bonsall, M. B., V. A. A. Jansen, and M. P. Hassell (2004), *Science* **306** (5693), 111, 859YJ Times Cited:58 Cited References Count:32.
- Borile, C., A. Maritan, and M. A. Muñoz (2013), *Journal of Statistical Mechanics: Theory and Experiment* **2013** (04), P04032.
- Borile, C., M. A. Muñoz, S. Azaele, J. R. Banavar, and A. Maritan (2012), *Physical Review Letters* **109** (3), 038102.
- Bork, P., C. Bowler, C. de Vargas, G. Gorsky, E. Karsenti, and P. Wincker (2015), *Science* **348** (6237), 873.
- Brown, J. H. (1995), *Macroecology* (The University of Chicago Press, Chicago).
- Butler, T., and N. Goldenfeld (2009), *Physical Review E* **80** (3), 030902.
- Carrara, F., F. Altermatt, I. Rodríguez-Iturbe, and A. Rinaldo (2012), *Proceedings of the National Academy of Sciences* **109** (15), 5761.
- Carrara, F., A. Rinaldo, A. Giometto, and F. Altermatt (2014), *The American Naturalist* **183** (1), 13.
- Cencini, M., S. Pigolotti, and M. A. Muñoz (2012), *PloS one* **7** (6), e38232.
- Chandrasekhar, S. (1943), *Review of Modern Physics* **15** (1), 1.
- Chase, J. M., and M. A. Leibold (2003), *Ecological niches: linking classical and contemporary approaches* (University of Chicago Press).
- Chave, J. (2004), *Ecology Letters* **7** (3), 241.
- Chave, J., D. Alonso, and R. S. Etienne (2006), *Nature* **441** (7089), E1.
- Chave, J., and E. G. Leigh (2002), *Theoretical Population Biology* **62** (2), 153.
- Chave, J., H. C. Muller-Landau, and S. A. Levin (2002), *The American Naturalist* **159** (1), 1.

- Chayes, J. T., L. Chayes, and E. H. Lieb (1984), *Communications in Mathematical Physics* **93** (1), 57.
- Chesson, P. (2000), *Annual review of Ecology and Systematics* , 343.
- Chesson, P. L., and R. R. Warner (1981), *American Naturalist* , 923.
- Chisholm, R. A. (2011), *Theoretical Population Biology* **80** (2), 71.
- Chisholm, R. A., R. Condit, K. A. Rahman, P. J. Baker, S. Bunyavejchewin, Y.-Y. Chen, G. Chuyong, H. Dattaraja, S. Davies, C. E. Ewango, *et al.* (2014), *Ecology letters* **17** (7), 855.
- Chisholm, R. A., and J. P. ODwyer (2014), *Theoretical population biology* **93**, 85.
- Clark, J. S. (2009), *Trends in Ecology & Evolution* **24** (1), 8.
- Coleman, B. D. (1981), *Mathematical Biosciences* **54** (3), 191.
- Coleman, B. D., M. A. Mares, and Y.-H. Hsieh (1982), *Ecology* **63** (4), 1121.
- Connell, J. H. (1971), *Dynamics of populations* **298**, 312.
- Darwin, C. R. (1909), *The Origin of Species*, Vol. Vol. 11 (The Harvard Classics. New York).
- Davies, T. J., A. P. Allen, L. Borda-de Águas, J. Regetz, and C. J. Melián (2011), *Evolution* **65** (7), 1841.
- Diamond, J. M. (1989), *Philosophical Transactions of the Royal Society B* **325** (1228), 469.
- Du, X., S. Zhou, and R. S. Etienne (2011), *Journal of theoretical biology* **278** (1), 127.
- Durrett, R., and S. Levin (1996), *Journal of Theoretical Biology* **179** (2), 119.
- Etienne, R. S. (2005), *Ecology letters* **8** (3), 253.
- Etienne, R. S., and D. Alonso (2005), *Ecology Letters* **8** (11), 1147.
- Etienne, R. S., D. Alonso, and A. J. McKane (2007), *Journal of Theoretical Biology* **248** (3), 522.
- Etienne, R. S., and H. Olf (2004a), *The American Naturalist* **163** (1), 69.
- Etienne, R. S., and H. Olf (2004b), *Ecology Letters* **7** (3), 170.
- Etienne, R. S., and J. Rosindell (2011), *Plos One* **6** (3).
- Feller, W. (1951), *Annals of Mathematics* **54**, 173.
- Fisher, A., R. Corbet, and C. Williams (1943), *Journal of Animal Ecology* (12), 42.
- Fisher, C. K., and P. Mehta (2014), *Proceedings of the National Academy of Sciences* **111** (36), 13111.
- Gardiner, C. (2009), *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer).
- Gause, G. F. (1934), *The struggle for existence* (Williams & Wilkins Company, Baltimore).

Goldenfeld, N., and C. Woese (2007), *Nature* **445** (7126), 369.

Goldenfeld, N., and C. Woese (2010), arXiv preprint arXiv:1011.4125.

Gould, S. J. (1979), *American Naturalist* , 335.

Gravel, D., C. D. Canham, M. Beaudet, and C. Messier (2006), *Ecology Letters* **9** (4), 399.

Grilli, J., S. Azaele, J. R. Banavar, and A. Maritan (2012a), *EPL (Europhysics Letters)* **100** (3), 38002.

Grilli, J., S. Azaele, J. R. Banavar, and A. Maritan (2012b), *Journal of Theoretical Biology*.

Haegeman, B., and R. S. Etienne (2008), *Journal of Theoretical Biology* **252** (2), 288.

Haegeman, B., and R. S. Etienne (2009), *Bulletin of mathematical biology* **71** (6), 1507.

Haegeman, B., and R. S. Etienne (2010), *Physical Review E* **81** (3), 031911.

Haegeman, B., and M. Loreau (2011), *Journal of Theoretical Biology* **269** (1), 150.

Halley, J. M., and Y. Iwasa (2011), *Proceedings of the National Academy of Sciences* **108** (6), 2316.

Harte, J. (2011), *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics* (Oxford University Press).

Harte, J., T. Zillio, E. Conlisk, and A. Smith (2008), *Ecology* **89** (10), 2700.

Hashitsume, N., M. Toda, R. Kubo, and N. Saitō (1991), *Statistical physics II: nonequilibrium statistical mechanics*, Vol. 30 (Springer Science & Business Media).

He, F., and S. P. Hubbell (2011), *Nature* **473** (7347), 368.

Holley, R. A., and T. M. Liggett (1975), *The annals of probability* , 643.

Holt, R. D. (2006), *Trends in Ecology & Evolution* **21** (10), 531.

Hu, X. S., F. L. He, and S. P. Hubbell (2006), *Oikos* **113** (3), 548.

Huang, K. (2001), *Introduction to statistical physics* (CRC Press).

Hubbell, S. (2001), *The Unified Theory of Biodiversity and Biogeography* (Princeton University Press).

Hubbell, S. P. (1979), *Science* **203** (4387), 1299.

Hubbell, S. P. (2005), *Functional ecology* **19** (1), 166.

Hubbell, S. P. (2006), *Ecology* **87** (6), 1387.

Hutchinson, G. E. (1957), in *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. 22 (Cold Spring Harbor Laboratory Press) pp. 415–427.

Hutchinson, G. E. (1959), *American naturalist* , 145.

- Janzen, D. H. (1970), *American naturalist* , 501.
- Jaynes, E. T. (2003), *Probability Theory: The Logic of Science* (Cambridge University Press).
- Jeraldo, P., M. Sipos, N. Chia, J. M. Brulc, A. S. Dhillon, M. E. Konkel, C. L. Larson, K. E. Nelson, A. Qu, L. B. Schook, *et al.* (2012), *Proceedings of the National Academy of Sciences* **109** (25), 9692.
- J.Harte, (2003), *Nature* **424**, 1007.
- Kalyuzhny, M., R. Kadmon, and N. M. Shnerb (2015), *Ecology letters* **18** (6), 572.
- Kalyuzhny, M., E. Seri, R. Chocron, C. H. Flather, R. Kadmon, and N. M. Shnerb (2014), *The American Naturalist* **184** (4), 439.
- Kampen, N. V. (2007), *Stochastic Processes in Physics and Chemistry, Third Edition (North-Holland Personal Library)* (North Holland).
- Keil, P., T. Herben, J. Rosindell, and D. Storch (2010), *Journal of Theoretical Biology* **265** (1), 78.
- Keitt, T. H., and H. E. Stanley (1998), *Nature* **393** (6682), 257.
- Kessler, D. A., and N. M. Shnerb (2015), *Physical Review E* **91** (4), 042705.
- Kessler, D. A., S. Suweis, M. Formentin, and N. M. Shnerb (2015), *Physical Review E* **92** (1), 022722.
- Kimura, M. (1985), *The Neutral Theory of Molecular Evolution*.
- Kingman, J. F. C. (1982), *Stochastic processes and their applications* **13** (3), 235.
- Krishna, A., P. R. Guimara, P. Jordano, and J. Bascompte (2008), *Oikos* **117** (May), 16091617.
- Lande, R., S. Engen, and B.-E. Saether (2003), *Stochastic population dynamics in ecology and conservation* (Oxford University Press Oxford).
- Lebedev, N. N. (1972), *Special functions and their applications* (Courier Dover Publications).
- Lehngig, S. H. (1993), *The generalized Feller equations and related topics* (Longman).
- Leigh, E. G. (2007), *Journal of evolutionary biology* **20** (6), 2075.
- Lynch, H. J., E. H. Campbell Grant, R. Muneeppeerakul, M. Arunachalam, I. Rodriguez-Iturbe, and W. F. Fagan (2011), *Water Resources Research* **47** (5).
- MacArthur, R. (1970), *Theoretical population biology* **1** (1), 1.
- MacArthur, R. H., and E. O. Wilson (1967), *The Theory of Island Biogeography* (Princeton University Press).
- Manceau, M., A. Lambert, and H. Morlon (2015), *Ecology letters* **18** (4), 347.

- Mari, L., R. Casagrandi, E. Bertuzzo, A. Rinaldo, and M. Gatto (2014), *Ecology letters* **17** (4), 426.
- Martín, H. G., and N. Goldenfeld (2006), *Proceedings of the National Academy of Sciences* **103** (27), 10310.
- Martin, T. E. (1981), *American Naturalist* **118** (6), 823.
- Masoliver, J., and J. Perello (2012), *Physical Review E* **86** (4).
- May, R. M. (2001), *Stability and complexity in model ecosystems*, Vol. 6 (Princeton University Press).
- May, R. M. (2010), *Philosophical Transactions of the Royal Society B* **365** (1537), 41.
- May, R. M., and R. H. Mac Arthur (1972), *Proceedings of the National Academy of Sciences* **69** (5), 1109.
- McGill, B. (2003), *Oikos* **102** (3), 679.
- McGill, B., and C. Collins (2003), *Evolutionary Ecology Research* **5** (4), 469.
- McGill, B. J., B. A. Maurer, and M. D. Weiser (2006), *Ecology* **87** (6), 1411.
- McKane, A., D. Alonso, and R. V. Solé (2000), *PRE* **62** (6), 8466.
- McKane, A. J., D. Alonso, and R. V. Sol (2004), *Theoretical Population Biology* **65** (1), 67.
- Melbinger, A., and M. Vergassola (2015), *Scientific reports* **5**.
- Morlon, H., G. Chuyong, R. Condit, S. Hubbell, D. Kenfack, D. Thomas, R. Valencia, and J. L. Green (2008), *Ecology Letters* **11** (9), 904.
- Mouquet, N., and M. Loreau (2003), *The american naturalist* **162** (5), 544.
- Muneepeerakul, R., E. Bertuzzo, H. J. Lynch, W. F. Fagan, A. Rinaldo, and I. Rodriguez-Iturbe (2008), *Nature* **453** (7192), 220.
- Mutshinda, C. M., R. B. OHara, and I. P. Woiwod (2008), *Functional ecology* **22** (2), 340.
- Nee, S. (2005), *Functional Ecology* **19**, 173.
- Noble, A. E., and W. F. Fagan (2015), *Theoretical Ecology* **8** (1), 149.
- Noble, A. E., A. Hastings, and W. F. Fagan (2011a), *Physical review letters* **107** (22), 228101.
- Noble, A. E., N. M. Temme, W. F. Fagan, and T. H. Keitt (2011b), *Journal of theoretical biology* **273** (1), 1.
- O'Dwyer, J. P., and R. Chisholm (2014), *Ecology letters* **17** (8), 961.
- O'Dwyer, J. P., and J. L. Green (2010), *Ecology Letters* **13** (1), 8795.
- ODwyer, J. P., S. W. Kembel, and T. J. Sharpton (2015), *Proceedings of the National Academy*

- of Sciences **112** (27), 8356.
- Pigolotti, S., and M. Cencini (2009), *Journal of theoretical biology* **260** (1), 83.
- Pigolotti, S., A. Flammini, and A. Maritan (2004), *Physical Review E* **70** (1), 11916.
- Pigolotti, S., A. Flammini, M. Marsili, and A. Maritan (2005), *Proceedings of the National Academy of Sciences of the United States of America* **102** (44), 15747.
- Pimm, S. L., G. J. Russell, J. L. Gittleman, and T. M. Brooks (1995), *Science-AAAS-Weekly Paper Edition* **269** (5222), 347.
- Pinto, O. A., and M. A. Muñoz (2011), *PloS one* **6** (7), e21946.
- Plotkin, J. B., M. D. Potts, N. Leslie, N. Manokaran, J. LaFrankie, and P. S. Ashton (2000), *Journal of Theoretical Biology* **207** (1), 81.
- Power, D. M. (1972), *Evolution*, 451.
- Preston, F. W. (1948), *Ecology* **29**, 254.
- Preston, F. W. (1960), *Ecology* **41** (4), 611.
- Pueyo, S. (2006), *Oikos* **112** (2), 392.
- Purves, D. W., S. W. Pacala, D. Burslem, M. Pinard, S. Hartley, *et al.* (2005), *Biotic interactions in the tropics: their role in the maintenance of species diversity*, 107.
- Purves, D. W., and L. A. Turnbull (2010), *Journal of Animal Ecology* **79** (6), 1215.
- Ricklefs, R. E. (2003), *Oikos* **100** (1), 185.
- Ricklefs, R. E. (2006), *Ecology* **87** (6), 1424.
- Ricklefs, R. E., and S. S. Renner (2012), *Science* **335** (6067), 464.
- Rosindell, J., and S. J. Cornell (2007), *Ecology Letters* **10** (7), 586.
- Rosindell, J., and S. J. Cornell (2009), *Ecology* **90** (7), 1743.
- Rosindell, J., S. J. Cornell, S. P. Hubbell, and R. S. Etienne (2010), *Ecology Letters* **13** (6), 716.
- Rosindell, J., L. J. Harmon, and R. S. Etienne (2015), *Ecology letters* **18** (5), 472.
- Rosindell, J., S. P. Hubbell, and R. S. Etienne (2011), *Trends in Ecology & Evolution* **26** (7), 340.
- Rosindell, J., S. P. Hubbell, F. L. He, L. J. Harmon, and R. S. Etienne (2012), *Trends in Ecology & Evolution* **27** (4), 203.
- Rosindell, J., Y. Wong, and R. S. Etienne (2008), *Ecological Informatics* **3** (3), 259.
- Roughgarden, J., *et al.* (1989), *Perspectives in ecological theory* **203**, 226.
- Scheffer, M., S. R. Carpenter, T. M. Lenton, J. Bascompte, W. Brock, V. Dakos, J. Van De Koppel, I. A. Van De Leemput, S. A. Levin, E. H. Van Nes, *et al.* (2012), *science* **338** (6105), 344.

- Scheffer, M., and E. H. van Nes (2006), *Proceedings of the National Academy of Sciences of the United States of America* **103** (16), 6230.
- Schoener, T. W. (1976), in *Proceedings of the XVI International Ornithological Congress*, Vol. 6 (Australian Academy of Science: Canberra, ACT) pp. 629–642.
- Simpson, E. H. (1949), *Nature*.
- Slik, J. F., V. Arroyo-Rodríguez, S.-I. Aiba, P. Alvarez-Loayza, L. F. Alves, P. Ashton, P. Balvanera, M. L. Bastian, P. J. Bellingham, E. Van Den Berg, *et al.* (2015), *Proceedings of the National Academy of Sciences* **112** (24), 7472.
- ter Steege, H., N. C. Pitman, T. J. Killeen, W. F. Laurance, C. A. Peres, J. E. Guevara, R. P. Salomão, C. V. Castilho, I. L. Amaral, F. D. de Almeida Matos, *et al.* (2015), *Science Advances* **1** (10), e1500936.
- Storch, D., P. A. Marquet, J. H. Brown, *et al.* (2007), *Scaling biodiversity* (Cambridge University Press Cambridge, UK).
- Suweis, S., E. Bertuzzo, L. Mari, I. Rodriguez-Iturbe, A. Maritan, and A. Rinaldo (2012a), *Journal of theoretical biology* **303**, 15.
- Suweis, S., and P. D’Odorico (2014), *PLoS ONE* **9** (7).
- Suweis, S., A. Rinaldo, and A. Maritan (2012b), *Journal of Statistical Mechanics: Theory and Experiment* **2012** (07), P07017.
- Suweis, S., F. Simini, J. R. Banavar, and A. Maritan (2013), *Nature* **500** (7463), 449.
- Svenning, J.-C., and R. Condit (2008), *Science* **322** (5899), 206.
- Ter Steege, H., N. C. Pitman, D. Sabatier, C. Baraloto, R. P. Salomão, J. E. Guevara, O. L. Phillips, C. V. Castilho, W. E. Magnusson, J.-F. Molino, *et al.* (2013), *Science* **342** (6156), 1243092.
- Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. F. de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. van Jaarsveld, G. F. Midgley, L. Miles, M. A. Ortega-Huerta, A. T. Peterson, O. L. Phillips, and S. E. Williams (2004), *Nature* **427** (6970), 145.
- Tilman, D. (1994), *Ecology* **75** (1), 2.
- Tilman, D., and P. M. Kareiva (1997), *Spatial ecology: the role of space in population dynamics and interspecific interactions*, Vol. 30 (Princeton University Press).
- Ulrich, W., M. Ollik, and K. I. Ugland (2010), *Oikos* **119** (7), 1149.

- Vallade, M., and B. Houchmandzadeh (2003), *Physical Review E* **68** (6), 061902.
- Vallade, M., and B. Houchmandzadeh (2006), *Physical Review E* **74** (5, Part 1), 10.1103/PhysRevE.74.051914.
- Vanpeteghem, D., O. Zemb, and B. Haegeman (2008), *Mathematical biosciences* **212** (1), 88.
- Veech, J. A. (2006), *Journal of Biogeography* **33** (12), 2145.
- Vergnon, R., R. Leijts, E. H. van Nes, and M. Scheffer (2013a), *The American Naturalist* **182** (1), 67.
- Vergnon, R., E. H. van Nes, and M. Scheffer (2012), *Nature communications* **3**, 663.
- Vergnon, R., E. H. van Nes, and M. Scheffer (2013b), *Oikos*.
- Volkov, I., J. Banavar, F. He, S. Hubbell, and A. Maritan (2005), *Nature* **438** (7068), 658.
- Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan (2003), *Nature* **424** (6952), 1035.
- Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan (2007), *Nature* **450** (7166), 45.
- Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan (2009), *Proceedings of the National Academy of Sciences* **106** (33), 13854.
- Volkov, I., J. R. Banavar, and A. Maritan (2006), *Journal of Statistical Physics* **123** (1), 167.
- Walker, S. C. (2007), *Theoretical Population Biology* **71** (3), 318.
- Watson, H. C. (1835), *Remarks on the Geographical Distribution of British Plants;: Chiefly in Connection with Latitude, Elevation, and Climate* (Longman, Rees, Orme, Brown, Green, and Longman, Paternoster-Row.).
- Zillio, T., J. R. Banavar, J. L. Green, J. Harte, and A. Maritan (2008), *Proceedings of the National Academy of Sciences of the United States of America* **105** (48), 18714.
- Zillio, T., I. Volkov, J. R. Banavar, S. P. Hubbell, and A. Maritan (2005), *Physical Review Letters* **95** (9), 098101.

FIGURES

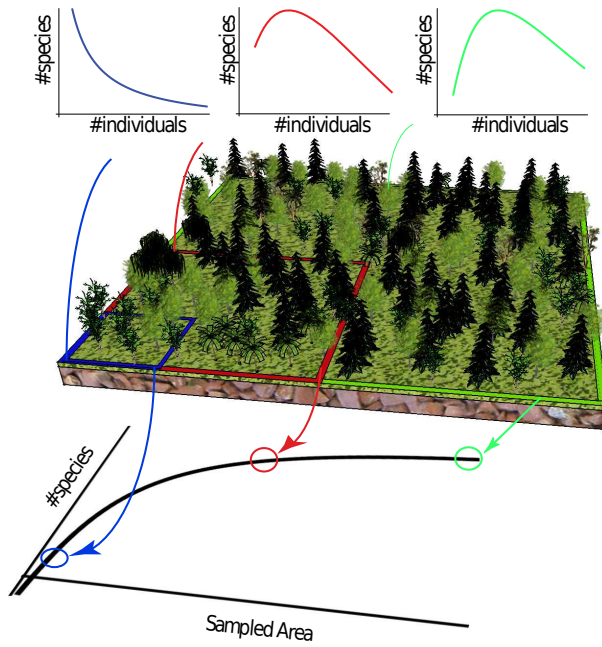


FIG. 1 [From (Azaele *et al.*, 2015)]. Visual Scheme of two important macro ecological patterns (see BOX 1): RSA and SAR. The functional shape of the RSA depends on the spatial scale considered, while the SAR generally displays a tri-phasic behavior (see Section IV). There is a growing appreciation that the various descriptors of biodiversity are intrinsically inter-related, and substantial efforts have been devoted to understand the links between them (Azaele *et al.*, 2015).

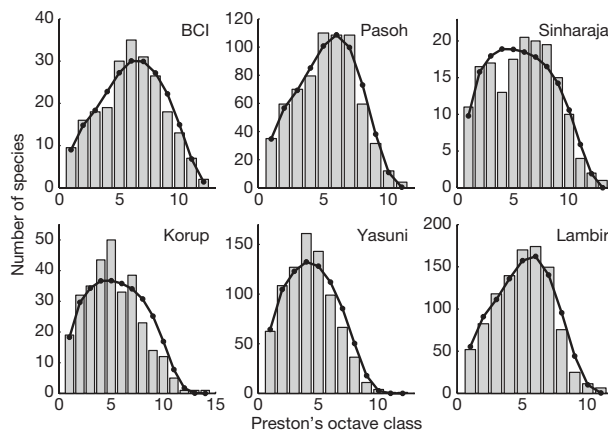


FIG. 2 [From (Volkov *et al.*, 2007)]. Tree relative species abundance data from the BCI, Yasuni, Pasoh, Lambir, Korup, and Sinharaja plots, for trees that are 10 cm in stem diameter at breast height. The frequency distributions are plotted using Preston’s binning method as described in (Volkov *et al.*, 2003) and the bars are the observed number of species binned into $\log(2)$ abundance categories. The first histogram bar represents $\frac{\langle\phi_1\rangle}{2}$, the second bar $\frac{\langle\phi_1\rangle}{2} + \frac{\langle\phi_2\rangle}{2}$, the third bar $\frac{\langle\phi_2\rangle}{2} + \langle\phi_3\rangle + \frac{\langle\phi_4\rangle}{2}$, the fourth bar $\frac{\langle\phi_4\rangle}{2} + \langle\phi_5\rangle + \langle\phi_6\rangle + \langle\phi_7\rangle + \frac{\langle\phi_8\rangle}{2}$ and so on. Here $\langle\phi_n\rangle$ is the number of species with an abundance n . As examples, we show the fits of a density-dependent symmetric model (black line), which will be studied in greater detail in sec. II.

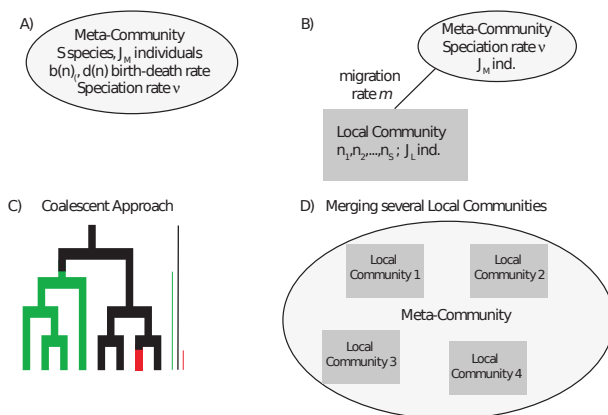


FIG. 3 Different neutral models of community ecology. In all these models, n_i represents the abundance of the species i in the community, and J_L and J_M the total abundance in the local and meta-communities, respectively. (A) *Hubbell’s zero – sum neutral model* (Hubbell, 2001). In the local community, each death is immediately followed by a birth or an immigration event. Speciation (or, equivalently, immigration) enables diversity to be maintained in the meta-community. (B) *Local community with immigrants from a meta-community* (Vallade and Houchmandzadeh, 2003). The local community now interacts with the meta-community through a migration process (m). (C) *Coalescent-type approach* (Etienne and Olf, 2004b), where community members are traced back to the ancestors that once immigrated into the community (D) *Joint RSA of many local communities* (Volkov *et al.*, 2007). The whole meta-community RSA distribution is built by considering the joint RSA distributions of multiple local communities

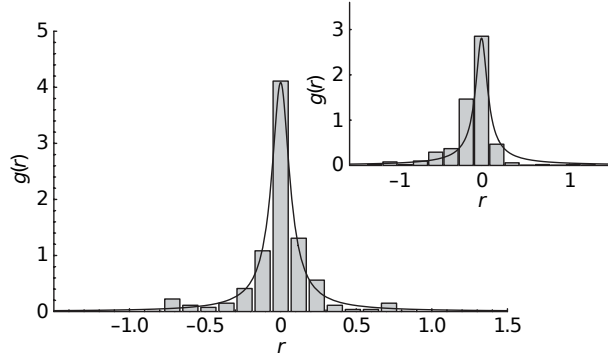


FIG. 4 [From (Azaele *et al.*, 2006)]. STD for the interval 1990-95 in the BCI forest. The main panel shows the results for individuals of more than 10 cm d.b.h., and the inset the results for individuals of more than 1 cm d.b.h. (Center for Tropical Forest Science website). The black line represents the analytical solution given by Eq. (46).

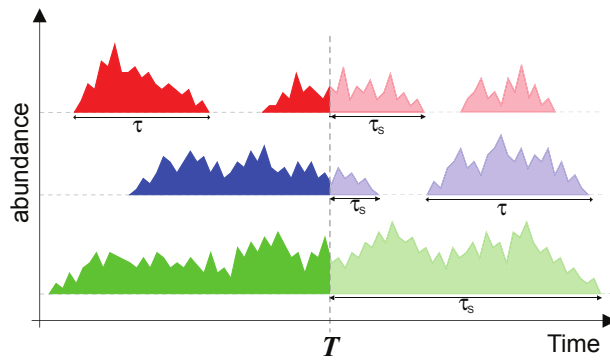


FIG. 5 [From (Suweis *et al.*, 2012b)]. Schematic representation of persistence time (or lifetime) of a species τ and survival times τ_s , defined as the time to local extinction of a species randomly sampled among the observed assemblages at a certain time T .

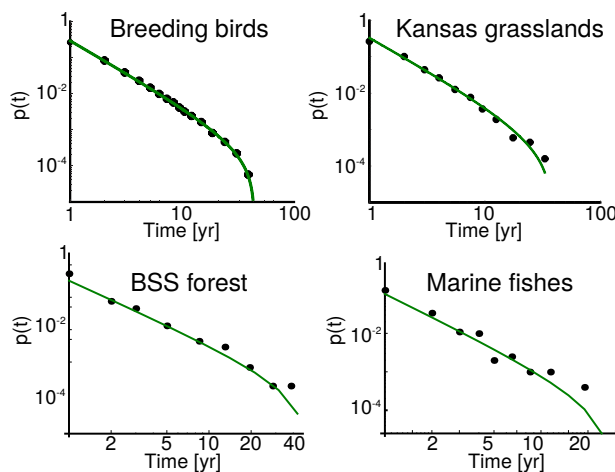


FIG. 6 [Modified *from* (Suweis *et al.*, 2012a)]. Comparison between persistence empirical distributions for (a) North American Breeding birds, (b) Kansas grasslands, (c) New Jersey BSS forest, (d) an estuarine fish community and the corresponding theoretical species persistence times pdfs. The circles and solid lines show the observational distributions and fit, respectively. The finiteness of the time window ΔT_w imposes a cut-off in the maximum observable persistence time and thus, only lifetimes where $\tau < \Delta T_w$ have been considered and the theoretical predictions have been adjusted appropriately (appendix B).

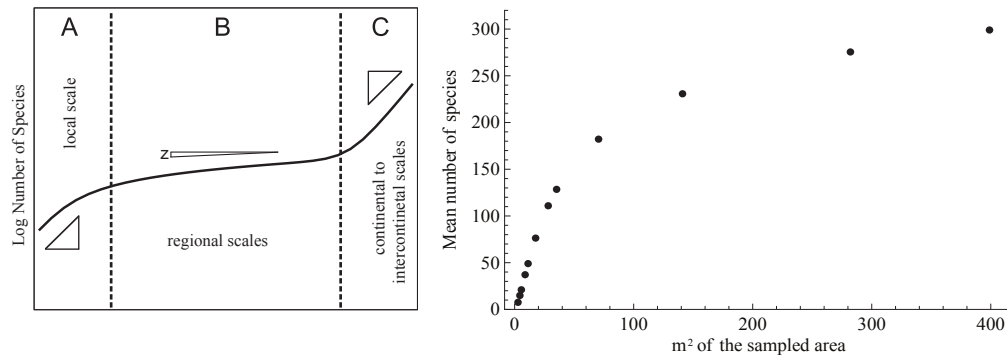


FIG. 7 Tri-phasic shape of the Species-Area relationship. The left panel shows the three behaviors on different scales. At a local scale the relationship is linear, becoming a power-law relationship at the regional scale and returning to linear at very large intercontinental scales. The right panel shows empirical data of species diversity at the regional scale (from BCI forest) (Azaele *et al.*, 2015).

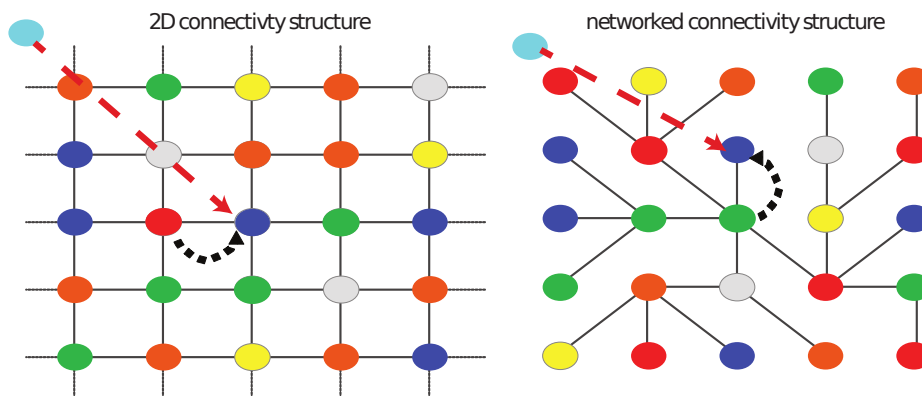


FIG. 8 Microscopic moves of the Multispecies Voter Model in a 2d lattice and with a non-regular network (Carrara *et al.*, 2012). In the 2d structure, each site is occupied by one and only one individual, whose color represents the species it belongs to. At each time step, one random individual is replaced by a daughter of one of its neighbors with probability $1 - \nu$ (black line). The probability that a speciation event occurs is ν , wherein the individual is replaced by an individual of a new species (red line). In the case of a non-regular network (right panel), the number of neighbors depends on the site considered.