



Statistical Mechanics of Ecosystem Assembly

José A. Capitán,¹ José A. Cuesta,¹ and Jordi Bascompte²

¹*Grupo Interdisciplinar de Sistemas Complejos (GISC), Departamento de Matemáticas, Universidad Carlos III de Madrid, E 28911 Leganés, Madrid, Spain*

²*Integrative Ecology Group, Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, c/ Americo Vespucio s/n, E 41092 Sevilla, Spain*

We introduce a toy model of ecosystem assembly for which we are able to map out all assembly pathways generated by external invasions. The model allows us to display the whole phase space in the form of an assembly graph whose nodes are communities of species and whose directed links are transitions between them induced by invasions. We characterize the process as a finite Markov chain and prove that it exhibits a unique set of recurrent states (the end state of the process), which is therefore resistant to invasions. This also shows that the end state is independent of the assembly history. The model shares all features with standard assembly models reported in the literature, with the advantage that all observables can be computed in an exact manner.

DOI: [10.1103/PhysRevLett.103.168101](https://doi.org/10.1103/PhysRevLett.103.168101)

PACS numbers: 87.23.Cc, 64.60.De, 87.23.Kg, 89.75.Fb

Understanding why ecosystems are both stable and complex is still one of the open questions in ecology [1]. At present, it is widely accepted that the dynamic process by which ecological communities are assembled is key to solving this puzzle. It also has profound implications for conservation—e.g., it can shed light on how biodiversity may recover after major crisis, or the influence it has on community stability. Although ecosystem assembly has been studied experimentally [2,3], the bulk of the studies are theoretical [4–10]. These assembly models are but idealizations of the complex processes taking place in real ecosystem assembly, which nevertheless implement the same mechanisms acting in the formation of real ecosystems [11]. In this respect these models are close in spirit to the general approach of statistical physics of devising oversimplified paradigms which provide insights into the real phenomena.

Previous work on ecosystem assembly made stochastic realizations of sequential invasions based on a finite set of possible invaders (known as “regional species pool” [7]) whose trophic interactions are determined in advance. These studies conclude that (i) at the end of the process, a final end state resistant to invasions is reached, which can be either a single community or a set involving more than one community between which the system fluctuates [8], (ii) the average resistance of an ecosystem to be colonized increases in time, and (iii) species richness also increases in time. Then the assembly process favors increasing species richness as well as stability, understood as resistance to invasions.

Successful as these models may be to provide insights into the formation of ecosystems, we still lack a global picture of the process. The reason is at least twofold. First, these models are defined as very complex stochastic processes not amenable to analytical treatment, so that one can only hope to simulate a limited set of realizations of the

process and take averages on them. This leaves some questions open such as, for instance, whether the end state depends on the assembly history. Although most evidence points to the uniqueness of this end state [11], it is still a matter of discussion [12]. Second, the size of the typical species pool is always small, so that it has been argued [13] that the exhaustion of good invaders might justify the increasing resistance to invasion.

Our aim in this Letter is to propose a toy version of an assembly model which, despite the stylized communities it deals with, still contains all the ingredients of standard assembly models (something like the “Ising model” of ecosystem assembly) and lacks some of its limitations (like the finiteness of the regional species pool). Its major advantage over them is that it allows us to map out the whole “phase space” of the system, as well as to analyze *all* assembly pathways, something that permits us to obtain exact results, to explore the set of parameters characterizing different regimes, and to answer questions out of reach of standard assembly models. The model recovers all the results mentioned above and provides new ones—the independence of the end state from the assembly history being the most prominent among them.

In what follows *ecosystem* will refer to the system as a whole, whereas *community* will stand for any particular collection of *species*, i.e., any realization of the ecosystem.

There are two ingredients that assembly models build upon: a deterministic population dynamics for the species forming the community, and a stochastic invasion process. For the former we essentially adopt a mean-field-like Lotka-Volterra model recently proposed to study coexistence in competing communities and trophic level organization in food webs [14–16] (these kind of models are called *neutral* in ecology [17], although neutral models assume a stochastic population dynamics, unlike the one defined here). In this model, communities are arranged in

trophic levels, and species at level ℓ are assumed to feed only on species at level $\ell - 1$, and on all of them alike (mean field). Although omnivory (i.e., feeding from lower levels) can be accounted for, it is still a matter of debate how common it is [18], so we disregard it. Accordingly, if n_i^ℓ denotes the population of species i at level $1 \leq \ell \leq L$,

$$\begin{aligned}\frac{\dot{n}_i^\ell}{n_i^\ell} &= -\alpha + \gamma_+ N^{\ell-1} - (1-\rho)n_i^\ell - \rho N^\ell - \gamma_- N^{\ell+1}, \\ \frac{\dot{n}_0}{n_0} &= R - n_0 - \gamma_- N^1,\end{aligned}\quad (1)$$

where $N^\ell \equiv \sum_{i=1}^{s_\ell} n_i^\ell$, s_ℓ being the number of species at level ℓ , γ_+ controls the amount of energy available to reproduction for each predation event, and γ_- takes into account the mean damage caused by predation to prey reproduction. Following the standard rule of efficiency on upwards energy transport to the next trophic level [19], we assume that these two constants are related by $\gamma_+ = 0.1\gamma_-$. Direct interspecific competition is measured by $\rho < 1$, while intraspecific competition is set to one to fix the population scale [16]. We regard all these species as consumers, and so they have a death rate $\alpha > 0$. For simplicity all parameters are assumed to be the same for all species (this assumption is harmless because the system has been shown to be structurally stable against variations of the constants [15]). The second of Eqs. (1) states that all species at the first level (basal species) predate on a single resource whose ‘‘amount’’ is given by n_0 . In the absence of basal species this resource reaches a steady value of $n_0 = R$; thus, this constant determines the maximum amount of resource available to consumers. Finally, as real populations cannot be arbitrarily small, it is required that extant species have a population $n_i^\ell > n_c$, the *extinction threshold*. If a population falls below this value it is considered extinct and removed from the community (see below). Low populations are vulnerable against external variations or adverse mutations [20], and this stochastic effect is somehow accounted for in the deterministic dynamics (1) by the introduction of this viability condition. Calculations have been carried out with $\alpha = 1$, $n_c = 1$, $\gamma_- = 5$, $\rho = 0.3$ and $0 < R \leq 1700$. We have checked that variations of these parameters do not change the qualitative picture.

Equations (1) have several equilibria, the main one being the *interior equilibrium*. In it all species of level ℓ have the same population n_ℓ^* . These are the solution to the system

$$\begin{aligned}-\alpha &= \gamma_- N_{\ell+1}^* + [1 + (s_\ell - 1)\rho]N_\ell^*/s_\ell - \gamma_+ N_{\ell-1}^*, \\ R &= N_0^* + \gamma_- N_1^*,\end{aligned}\quad (2)$$

where $\ell = 1, \dots, L$, $N_\ell^* = s_\ell n_\ell^*$, and we have the constraints $s_0 = 1$, $s_{L+1} = 0$. The remaining equilibria are obtained by setting to zero any subset of the populations and solving the linear system resulting from eliminating those variables. Therefore one only needs the solutions of the linear systems (2) for all possible choices of s_ℓ in order to characterize the dynamical equilibrium of this model.

Interior equilibria are globally stable if, and only if, $n_\ell^* > 0$ for all $\ell = 0, \dots, L$, because

$$V(\{n_i^\ell\}) = \sum_{\ell=0}^L \left(\frac{\gamma_-}{\gamma_+}\right)^\ell \sum_{i=1}^{s_\ell} (n_i^\ell - n_\ell^* \log n_i^\ell) \quad (3)$$

is a Lyapunov function [21] for (1). Therefore, given the set of species numbers $\{s_\ell\}_{\ell=1}^L$, the corresponding community is viable and stable if, and only if, $n_\ell^* > n_c$ for all $\ell = 1, \dots, L$. Thus by solving (2) for all choices of species numbers we can determine all viable and stable communities. Let \mathcal{F} denote the set of these communities. Although the total size of a trophic level is not explicitly limited in any way, the extinction threshold n_c imposes a constraint and therefore \mathcal{F} is finite.

Equilibrium communities will undergo invasions which change their composition. We shall assume that they take place at a uniform rate ξ . Two hypotheses underlying this model, as well as most assembly models, are (i) the typical dynamical time is much smaller than ξ^{-1} , so that communities are never invaded during transients, and (ii) the population of invaders is small [22]. Obviously, if the invasion rate is too high [23,24] or if invaders arrive with high populations [25], the assembly dynamics—and, consequently, the final end states—will be seriously altered. These situations are outside the scope of this model.

The invasion process goes as follows. Consider a community $E \in \mathcal{F}$, with L trophic levels, at its equilibrium point. Invasions can occur at any level $\ell = 1, \dots, L + 1$, so we randomly select ℓ and add a new species, with initial population n_c , at level ℓ of the community E . Because of the global stability of our model, the extended community evolves to the equilibrium given by (2) with s_ℓ replaced by $s_\ell + 1$. If this equilibrium is viable, then we will have a new community E' consisting of E plus the invader, and a transition will have occurred from E to E' . If, on the contrary, the population of the species of at least one level falls below n_c , then extinctions will occur. It is unrealistic to extinguish the whole level, even though all its species have the same population. In reality, if many species are threatened, by chance one of them will be the first to become extinct. This fact may help the remaining species to survive. Accordingly we shall extinguish species in an inviable level as follows. As we are monitoring the whole trajectory of the system, we can detect the moment when the first trophic level crossed n_c . At that point we remove one species from that level and restart the evolution from that point. We keep on removing one species at a time and restarting until the new resulting equilibrium becomes viable. Two things can thus happen: Either the first level to fall below n_c is the invaded level, in which case the invader is simply rejected and no transition occurs, or it is another level that falls below n_c . In this case we end up with a new community E'' , and a transition will have occurred from E to E'' .

Starting from the empty community, \emptyset , and considering for every community all possible invasions, we construct a

directed graph, the *assembly graph*, \mathcal{G} , whose nodes are elements of \mathcal{F} and whose links are the transitions obtained by the invasion process just described. Figure 1 represents a typical assembly graph. From the viewpoint of statistical mechanics, \mathcal{G} is the phase space of our system.

The assembly process becomes a Markov chain if to every pair of communities E and E' we assign the transition probability $P_{EE'} = \delta_{EE'} + \xi Q_{EE'}$, where $Q_{EE'}$ is the fraction of the $L + 1$ different invasions of E that lead to E' ($Q_{EE'} \neq 0$ provided \mathcal{G} contains the link $E \rightarrow E'$). P is the transition matrix of a finite, aperiodic, Markov chain, so its states are either transient or recurrent. There can be one or several subsets of recurrent states, the chain being ergodic in each of them [26]. Every recurrent subset is a different end state of the assembly process. The assembly will depend on the history only if there are at least two such recurrent subsets. Each recurrent subset has a stationary probability distribution on its states.

It is worth noticing that the computation of P is exact. This means that we have a *complete and exact* description of the evolution of the assembly process on the phase space. In particular this means that we can compute the evolution of any observable in an exact—albeit numeric—manner, without resorting to taking averages over realizations of the process. This is the most important difference of this model with respect to all assembly models considered so far.

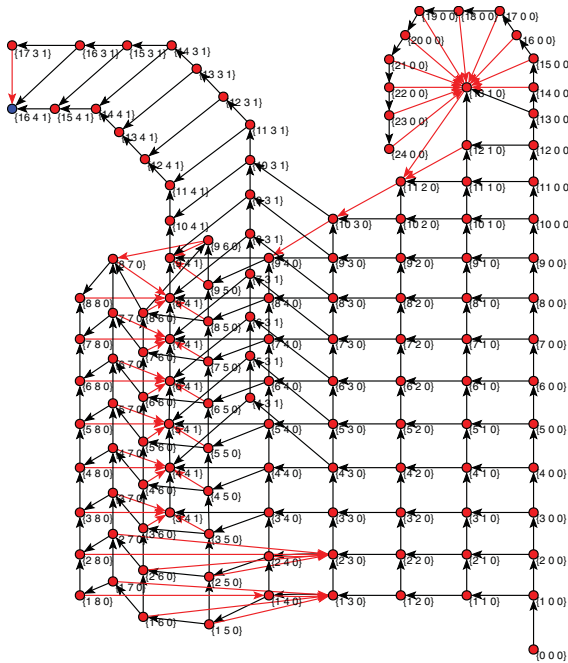


FIG. 1 (color online). Assembly graph obtained for $R = 140$. It displays 130 communities with up to three trophic levels. The numbers indicate how many species are in each level. Black arrows represent accepted invasions; red arrows represent transitions inducing a species loss. The only node in blue (corresponding to the community $\{16, 4, 1\}$) represents the end state of the assembly process.

Let us now describe the results. First of all, the process has a unique set of recurrent states for all $0 < R \leq 1700$. This means that for this model we are able to prove that the end state does not depend on the assembly history [5]. This result agrees with previous evidence found in other assembly models [8,11] as well as in experiments [3], where the same kind of assumptions about the invasion rate are made. For some values of R this end state contains a single community but typically it contains more than one, often very many. Figure 2 illustrates one of these sets along with the transitions between its states. At any time the ecosystem is realized by one of the communities of the set. Invasions may induce transitions between communities of this set but never lead outside it; hence, the set as a whole is resistant to invasions. The frequency with which a community is visited defines a stationary probability on the set. Notice that only a few communities (rather similar to each other) are visited with a high probability, so the net result is as if the community were “fluctuating.” Figure 3 shows the number of communities in the end state vs R . One can see that as R is increased towards the onset of a new level appearance the number of communities increases considerably, to drop down to just one once the new level has established.

Inset in Fig. 3 represents the mean number of species in the end state vs R . The dependence is basically linear, except for some dips near the onset of a new level followed by a discontinuous jump once the new level is established. This behavior reflects a top-predator effect: Top predators control the populations of their preys, avoiding their ex-

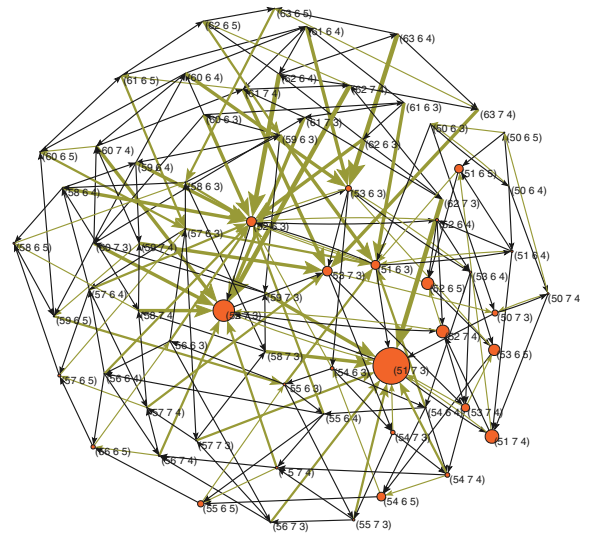


FIG. 2 (color online). End state of the assembly process obtained for $R = 430$. It contains 68 communities (the assembly graph contains 3060 communities). Black arrows represent accepted invasions, whereas yellow ones represent transitions with species loss (width is proportional to the number of extinctions). The diameter of the nodes is proportional to the stationary probability of its community. Labels provide the number of species at each level.

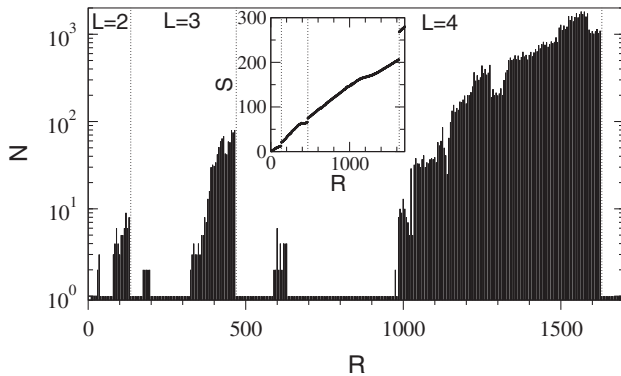


FIG. 3. Number of recurrent states of the Markov chain as a function of the resource saturation R , obtained using the algorithm of Ref. [28]. Inset: Mean number of species in the stationary state vs R .

tion by overconsumption of their resources [27]. As a matter of fact we have checked that when a new level appears it contains a single top predator and the number of species at lower levels rises.

Other interesting observables are the probabilities of accepting the invader, $P_i(t)$, and of undergoing a reconfiguration after the invasion, $P_a(t)$. Both are obtained as $P_{i/a}(t) = \sum_E (\sum_{E'} P_{EE'}) (P^t)_{E\emptyset}$, where the inner sum runs over transitions from E to $E' \neq E$ in which the invader is accepted, or in which there is a reconfiguration of the community, respectively. Both functions show a monotonic increase of the resistance to invasion as the community assembles. Species richness can be computed similarly. It grows monotonically up to saturation in the steady state. Both features, increasing resistance to invasion and increasing species richness, are common results of standard assembly models [7,8].

In summary, this minimalistic assembly model exhibits the same behavior as those reported in the literature, indicating that this behavior is very robust and probably shared also by real ecosystems. Our model, however, provides a complete and exact description of both, the set of microstates and the dynamical pathways of the assembly process. So we are not limited, as in standard assembly models, to compute averages over a small set of realizations of the process. To give a hint about what this means, we have calculated, for $R = 300$ (a case with an end state made of a single community of three trophic levels and 50 species), that there are $\sim 10^{10}$ different minimum-length pathways leading from the empty to the end state. This number is nothing that a simulation can come close to. This model allows us to prove rigorously that its end state does not depend on the assembly history. Whether this is a feature that real ecosystems exhibit will, of course, depend on how well they fulfill the assumptions about the invasion rate underlying this and other assembly models. But a caveat should be made: As species of each trophic level are indistinguishable, uniqueness (and hence independence on history) refers only to the number of species at each level.

We thank U. Bastolla, R. Law, S.C. Manrubia, A. Arenas, and J. Camacho for useful discussions. This work is funded by projects MOSAICO (Ministerio de Educación y Ciencia) and MOSSNOHO-CM (Comunidad de Madrid), by the European Heads of Research Councils, the European Science Foundation, and the EC Sixth Framework Programme through an EURYI grant (J.B.), and by a contract from Comunidad de Madrid and Fondo Social Europeo (J. A. Capitán).

- [1] K. S. McCann, *Nature (London)* **405**, 228 (2000).
- [2] R. Law, A. J. Weatherby, and P. H. Warren, *Oikos* **88**, 319 (2000).
- [3] P. H. Warren, R. Law, and A. J. Weatherby, *Ecology* **84**, 1001 (2003).
- [4] W. M. Post and S. L. Pimm, *Math. Biosci.* **64**, 169 (1983).
- [5] J. A. Drake, *J. Theor. Biol.* **147**, 213 (1990).
- [6] T. J. Case, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9610 (1990).
- [7] R. Law and R. D. Morton, *Ecology* **77**, 762 (1996).
- [8] R. D. Morton and R. Law, *J. Theor. Biol.* **187**, 321 (1997).
- [9] B. Drossel, P. G. Higgs, and A. J. McKane, *J. Theor. Biol.* **208**, 91 (2001).
- [10] N. Loeuille and M. Loreau, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5761 (2005).
- [11] R. Law, in *Advanced Ecological Theory*, edited by J. McGlade (Blackwell Science, Oxford, 1999).
- [12] T. Fukami and P. J. Morin, *Nature (London)* **424**, 423 (2003).
- [13] T. J. Case, *Biological J. Linnean Soc.* **42**, 239 (1991).
- [14] M. Lässig, U. Bastolla, S. C. Manrubia, and A. Valleriani, *Phys. Rev. Lett.* **86**, 4418 (2001).
- [15] U. Bastolla, M. Lässig, S. C. Manrubia, and A. Valleriani, *J. Theor. Biol.* **235**, 521 (2005).
- [16] U. Bastolla, M. A. Fortuna, A. Pascual García, A. Ferrera, B. Luque, and J. Bascompte, *Nature (London)* **458**, 1018 (2009).
- [17] R. S. Etienne and D. Alonso, *J. Stat. Phys.* **128**, 485 (2007).
- [18] J. Bascompte and C. J. Melián, *Ecology* **86**, 2868 (2005).
- [19] S. L. Pimm, *Food Webs* (Chapman & Hall, London, 1982).
- [20] S. L. Pimm, *The Balance of Nature: Ecological Issues in the Conservation of Species and Communities* (University of Chicago Press, Chicago, 1991).
- [21] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, England, 1998).
- [22] *Species Invasions: Insights into Ecology, Evolution, and Biogeography*, edited by D. F. Sax, J. J. Stachowicz, and S. D. Gaines (Sinauer Associates Inc., Sunderland, Massachusetts, 2005).
- [23] T. Fukami, *Popul. Ecol.* **46**, 137 (2004).
- [24] U. Bastolla, M. Lässig, S. C. Manrubia, and A. Valleriani, *J. Theor. Biol.* **235**, 531 (2005).
- [25] C. L. Hewitt and G. R. Huxel, *Biol. Invasions* **4**, 263 (2002).
- [26] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes* (Academic Press, New York, 1975).
- [27] K. R. Crooks and M. E. Soulé, *Nature (London)* **400**, 563 (1999).
- [28] A. Xie and P. A. Beerel, *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **17**, 1334 (1998).