On the limits to invasion prediction using coexistence outcomes

Jie Deng[1*], Washington Taylor[2], Simon A. Levin[3,4], Serguei Saavedra[1]

[1]Department of Civil and Environmental Engineering, MIT,
77 Massachusetts Av., 02139 Cambridge, MA, USA

[2]Center for Theoretical Physics, MIT,
77 Massachusetts Av., 02139 Cambridge, MA, USA

[3]Department of Ecology & Evolutionary Biology, Princeton University,
Princeton, NJ 08544 USA

[4]High Meadows Environmental Institute, Princeton University,
Princeton, NJ 08544 USA

* To whom correspondence should be addressed. E-mail: jiedeng@mit.edu

Keywords: geometry | invasion | exclusion | probability | interactions | predictability

Competing financial interests The authors declare no competing financial interests.

Code accessibility The code supporting the results will be archived on Zenodo upon publication. This code is available to reviewers on Github: `https://github.com/MITEcology/Deng_WLS_2023`

Preprint number: MIT-CTP/5544 (internal use only)

Abstract

The dynamics of ecological communities in nature are typically characterized by probabilistic, sequential assembly processes (i.e., invasion dynamics). Because of technical challenges, however, the majority of theoretical and experimental studies have focused on the outcomes derived from simultaneous assembly processes (i.e., coexistence dynamics). Therefore, it has become central to understand the extent to which the outcomes from such simultaneous processes can be used to predict analogous outcomes from sequential processes relevant for systems in nature. Here, we study the limits to this predictability under a geometric and probabilistic Lotka-Volterra framework. We show that while survival probability in simultaneous assembly can be fairly closely translated into colonization probability in sequential assembly, the translation is less precise between community persistence and community augmentation, and worse between exclusion probability and replacement probability. These results provide a guiding and testable theoretical framework regarding the (lack of) translatability of outcomes between simultaneous and sequential processes when communities are represented by Lotka-Volterra dynamics under environmental uncertainty.

Introduction

The assembly of ecological communities depends on multiple factors associated with eco-evolutionary dynamics and randomness (Chang et al., 2021, Drake, 1991, Hubbell, 1997, Odum and Barrett, 2005, Schreiber and Rittenhouse, 2004, Song et al., 2021, Vellend, 2016, Warren and Spencer, 1996). In nature, these factors are typically expressed under probabilistic, sequential, assembly processes (i.e., adding new species across time via invasion under changing environmental conditions) (Deng et al., 2021, Fukami, 2015, Hill et al., 2004, Jones et al., 2022, Odum, 1969, Williamson and Fittera, 1996). Because of technical challenges, however, the majority of theoretical and experimental studies have focused on the outcomes derived from simultaneous assembly processes (i.e., having all species included at once under fixed environmental conditions) (Case, 2000, Clark et al., 2021, Friedman et al., 2017, Gause, 1932, MacArthur and Levins, 1967, Maynard et al., 2020, Vandermeer, 1975). For example, it has proved extremely difficult to maintain a stable initial community for a sufficient time to then do sequential assembly experiments (Friedman et al., 2017, Gilpin et al., 1986, Gould et al., 2018, Park, 1954). Similarly, in the field, it becomes easier to record coexistence (or co-occurrence) outcomes rather than assembly outcomes (Barner et al., 2018). Therefore, it has become central to understand the extent to which the outcomes (e.g., individual survival, community persistence, competitor exclusion) from such simultaneous processes can be used to predict analogous outcomes (e.g., invader colonization, invader-resident augmentation, resident replacement) expected from sequential processes (Case, 1990, Fukami, 2015, Hubbell, 1997, Vagnon et al., 2022).

A key connection between coexistence (simultaneous assembly) and invasion (sequential assembly) dynamics is centered on the concept of *invasion growth rate* (Barabás et al., 2018, Case, 1990, Chesson and Kuang, 2008, Grainger and Gilbert, 2019, Law and Morton, 1996). The invasion growth rate quantifies the instantaneous rate at which an invader will grow (positive values) or decline (negative values) right after its introduction into a community of resident species at equilibrium. The invasion growth rate assumes that an invader's abundance (or density) is sufficiently small such that it does not affect the growth rate of resident species, but sufficiently large such that demographic stochasticity can be ignored (Barabás et al., 2018, Hofbauer and Schreiber, 2022). Therefore, positive invasion growth rates guarantee the successful short-term introduction of an invader species (Barabás et al., 2018). Under these assumptions and under specific model parameterizations, coexistence outcomes can be derived from information about invasion outcomes and vice versa (Arnoldi et al., 2022, Hofbauer and Schreiber, 2022, Logofet, 1993, Serván and Allesina, 2021). While these criteria have provided key insights regarding the translatability between simultaneous and sequential assembly, these results can be difficult to validate in natural communities where outcomes and conditions are rarely the same (Angulo et al., 2021, Miller and Allesina, 2021, Song et al., 2021, Zhao et al., 2021). This has revealed the importance of extending and developing testable theory to test the translatability between invasion and coexistence dynamics under a probabilistic framework (Deng et al., 2021, Jones et al., 2022, Williamson and Fittera, 1996).

Here, to study the limits to invasion prediction using coexistence outcomes, we develop a theoretical integration of invasion and coexistence under a geometric and probabilistic framework following generalized Lotka-Volterra dynamics. We illustrate an application of this integration by providing a theoretical expectation of the translatability between analogous probabilistic outcomes from simultaneous and sequential processes. We show, geometrically and using an *in silico* experiment, that while survival probability can be translated into colonization probability, there can be a large loss of translation between community persistence and community augmentation, as well as between exclusion probability and replacement probability. These results provide a guiding and testable theoretical framework regarding the (lack of) translatability of outcomes between simultaneous and sequential processes when communities are represented by Lotka-Volterra dynamics under environmental uncertainty.

Theoretical framework

Our theoretical framework follows the central assumption that ecological dynamics can be described by the generalized Lotka-Volterra (gLV) model of the form

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i \left(\theta_i + \sum_{j \in \mathcal{S}} a_{ij} N_j \right). \tag{1}$$

The vector $\mathbf{N} = (N_1, \cdots, N_{|\mathcal{S}|})^\intercal$ ($|\cdot|$ denotes the cardinality of a set) represents the densities of all species $i \in \{1, \cdots, |\mathcal{S}|\} \equiv \mathcal{S}$ in the community (system) $\mathcal{S}$. The matrix $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}$ corresponds to the pairwise interaction matrix, where the term a_{ij} represents the per-capita effect of species j on an individual of species i. In particular, $a_{ij} < 0$ and $a_{ij} > 0$ translate into a negative and positive interaction, respectively. The effective growth rate $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{|\mathcal{S}|})^\intercal \in \mathbb{R}^{|\mathcal{S}|}$ consists of the phenomenological joint effect of the internal (e.g., intrinsic growth rate), abiotic (e.g., temperature), and biotic factors (i.e., species not considered explicitly) acting on the per-capita growth rate of each particular species (Arnoldi et al., 2022, Deng et al., 2022). Note that if we multiply $\boldsymbol{\theta}$ by any positive scalar, it will not change the qualitative solution of the system. Thus, the choice of the norm (e.g., unit sphere) does not change the results (Rohr et al., 2016).

Without any prior information about $\Theta^{|\mathcal{S}|}$, we assume that all directions of $\boldsymbol{\theta}$ are possible and equally likely, making the parameter space $\Theta^{|\mathcal{S}|}$ of $\boldsymbol{\theta}$ a closed unit sphere in $|\mathcal{S}|$-dimensional space with uniform probability measure. In line with an internalist (also known as structuralist) perspective (Alberch, 1989, Rohr et al., 2014, Saavedra et al., 2017), this model description implies that pairwise species interactions are invariant and represent the structural constraints of biological systems. We simplify our study (and correspondingly restrict interaction matrices $\mathbf{A}$) to cases of globally stable gLV dynamics (Takeuchi, 1996), such that results do not depend on the initial conditions of the community, only on the interaction matrix and effective growth rates (AlAdwani and Saavedra, 2020, Saavedra et al., 2020) (non-globally stable matrices produce qualitatively similar results, Supplementary Section S3).

This assumption further implies that any positive equilibrium of Eq. (1) guarantees the long-term persistence (also permanence) of species (Hofbauer and Sigmund, 1998). In most of the analysis in this paper we thus assume a fixed interaction matrix $\mathbf{A}$, but that $\boldsymbol{\theta}$ is unknown and can take any value from a uniform distribution over the unit sphere.

Looking at invasion growth rate from a probabilistic perspective

Assuming that an invader's abundance is sufficiently small, the invasion growth rate can accurately predict the invader's behavior over a short term. Under gLV dynamics Eq. (1), the invasion growth rate of an invader i within a community $\mathcal{S}$ formed by a pool of $|\mathcal{S}|$ species can be formalized as

$$g(i) = \theta_i + \sum_{j:j\in\mathcal{S},\ j\neq i} a_{ij} N_j^*(\bar{i}), \tag{2}$$

where $(N_1^*(\bar{i}), \cdots, N_{i-1}^*(\bar{i}), N_{i+1}^*(\bar{i}), \cdots, N_{|\mathcal{S}|}^*(\bar{i}))^\intercal$ are the equilibrium abundances of the $(|\mathcal{S}|-1)$ resident species without the presence of species (invader) i. Note that this definition assumes the existence of steady-state positive abundances $N_j^*(\bar{i}) > 0$ for all resident species $j \neq i$ in the absence of species i. Ecologically, the invasion growth rate (Eq. (2)) specifies the short-term population change of the invader species i when the resident species are at equilibrium. Per definition, the calculation of the invasion growth rate requires not only the interspecific effects of the residents on the invader species a_{ij}, but also knowledge of the effective growth rate of the invader species.

In the absence of prior information about θ_i, the invasion growth rate can be represented geometrically and probabilistically. To show this, first we define a geometric quantity, that we call the *growth indicator*, and then we associate this quantity with the invasion growth rate defined in Eq. (2). For an arbitrary vector of effective growth rates $\boldsymbol{\theta}$ associated with the species forming a community $\mathcal{S}$, the growth indicator of an invader i can be defined as

$$GI(i,\theta) = \frac{\boldsymbol{\theta}\cdot\mathbf{n}(i)}{\mathbf{e}_i\cdot\mathbf{n}(i)}, \tag{3}$$

where $\mathbf{e}_i$ is the unit vector whose i^{th} entry is 1 and 0 elsewhere; $\mathbf{n}(i) = (n_1, n_2, \cdots, n_{|\mathcal{S}|})^\intercal \in \mathbb{R}^{|\mathcal{S}|}$ can be any normal vector that is orthogonal to the *facet* $F(i)$ in $|\mathcal{S}|$-dimensional space. Specifically, $F(i)$ is spanned by the line segments from the origin to the points corresponding to the $(|\mathcal{S}|-1)$ column vectors excluding the ith column vector of the interaction matrix $\mathbf{A}$ (i.e., $F(i)$ is spanned by the line segments from the origin to the points corresponding to the $1^{\text{st}}, \cdots, (i-1)^{\text{th}}, (i+1)^{\text{th}}, \cdots, |\mathcal{S}|^{\text{th}}$ column vectors, so that the normal vector $\mathbf{n}(i)$ is specific to the invader i). By definition, any normal vector $\mathbf{n}(i)$ that is orthogonal to the facet $F(i)$ has zero inner products with the vectors on the facet $F(i)$. In addition, for vectors on the same side of the facet $F(i)$, their inner products with the normal vector $\mathbf{n}(i)$ to the facet $F(i)$ have the same sign. The sign of $GI(i,\theta)$ reveals whether the vector of effective growth rates $\boldsymbol{\theta}$ is on the same side of $F(i)$ as the unit vector $\mathbf{e}_i$ relative to the facet $F(i)$. It can be

proved that $GI(i,\theta)$ is equivalent to the invasion growth rate $g(i)$ (Supplementary Section S1). More specifically, if $\boldsymbol{\theta}$ is on the same (resp. opposite) side as $\mathbf{e}_i$ (i.e., $GI(i,\theta) > 0$; resp. $GI(i,\theta) < 0$), then the invasion growth rate is $g(i) > 0$ (resp. $g(i) < 0$). Otherwise, if $\boldsymbol{\theta}$ is on the facet $F(i)$ (i.e., $GI(i,\theta) = 0$), then the invasion growth rate is $g(i) = 0$. That is, the area spanned by $GI(i,\theta) > 0$ corresponds to the range of conditions leading to a positive invasion growth rate.

Figure 1A illustrates geometrically the regions in the parameter space of $\boldsymbol{\theta}$ corresponding to positive and negative invasion growth rates for a hypothetical community with three species. In this representation, we assume species 3 is the invader and species $\{1,2\}$ are the residents. The interaction matrices are arbitrary and the column vectors correspond to the plane $F(i)$ determined by the residents where the invasion growth rate $g(i) = 0$. Focusing on the feasibility domain of the residents in isolation (bounded by the green dotted lines), if the vector of effective growth rates $\boldsymbol{\theta}$ is located inside the gray region $GI(i,\theta) > 0$ (resp. white region $GI(i,\theta) < 0$), the invasion growth rate is positive (resp. negative), meaning that the invader's population with a small initial abundance will grow (resp. decline) in the short term. Otherwise, if the vector of effective growth rates $\boldsymbol{\theta}$ is located at the border (i.e., facet $F(i)$) dividing the gray and white regions, the invasion growth rate is zero, meaning that the invader's population with a small initial abundance will remain unchanged over a short term. In the absence of prior information, we can assume that all values of $\boldsymbol{\theta}$ are equally likely. Then, the probability that species i enters into a resident community with a positive invasion growth rate ($g(i) > 0$) can be measured by the area of $\boldsymbol{\theta}$ compatible with $GI(i,\theta) > 0$ (gray region in Fig. 1A) relative to the area of $\boldsymbol{\theta}$ compatible with the persistence of all resident species (dotted region in Fig. 1A). Mathematically, in $|\mathcal{S}|$-dimensional space, the set of effective growth rates $\boldsymbol{\theta}$ allowing the persistence of a resident community $\mathcal{S} \setminus \{i\}$ in isolation is

$$D_F(\bar{i}) = \left\{ \boldsymbol{\theta} \in \mathbb{R}^{|\mathcal{S}|} : \theta_i \in \mathbb{R}, (\theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_{|\mathcal{S}|}) \in D_F(\mathcal{C} = \mathcal{S} \setminus \{i\}, \mathcal{S} \setminus \{i\}) \right\}, \tag{4}$$

where the feasibility domain $D_F(\mathcal{C}, \mathcal{S})$ of a species collection $\mathcal{C}$ ($\mathcal{C} \subseteq \mathcal{S},\ \mathcal{C} \neq \emptyset$) in a multispeices community $\mathcal{S}$ is defined in Eq. (7). Without prior information, we assume that the different directions of $\boldsymbol{\theta}$ in $D_F(\bar{i})$ are equally likely, and consider these directions as a sub-region of the closed unit sphere in dimension $|\mathcal{S}|$. Thus, the feasibility domain of residents in isolation $\boldsymbol{\theta}^*(\bar{i})$ can be represented by

$$\boldsymbol{\theta}^*(\bar{i}) = D_F(\bar{i}) \cap \Theta^{|\mathcal{S}|}, \tag{5}$$

where $\Theta^{|\mathcal{S}|}$ is the $(|\mathcal{S}| - 1)$-dimensional closed unit sphere in dimension $|\mathcal{S}|$. Formally, the probability of positive invasion growth rate of species i in community $\mathcal{S}$ can be defined as

$$\frac{\mathsf{vol}\left((GI(i,\theta) > 0) \cap \boldsymbol{\theta}^*(\bar{i})\right)}{\mathsf{vol}\left(\boldsymbol{\theta}^*(\bar{i})\right)}. \tag{6}$$

Looking at coexistence from a probabilistic perspective

As previously discussed, the invasion growth rate determines the invasion dynamics over the short term. Yet, the establishment of an invader species and its effects on residents require the analysis of species persistence over a longer time scale. Under global equilibrium dynamics (Takeuchi, 1996), feasibility (positive solutions to Eq. (1)) constitutes the necessary and sufficient condition for species persistence (and permanence) (Hofbauer and Sigmund, 1998). Following gLV dynamics, the set of vectors $\boldsymbol{\theta}$ leading to feasible solutions has been called the feasibility domain (Saavedra et al., 2020). Formally, the feasibility domain $D_F(\mathcal{C},\mathcal{S})$ of a species collection $\mathcal{C}$ ($\mathcal{C} \subseteq \mathcal{S},\ \mathcal{C} \neq \emptyset$) in a multispecies community $\mathcal{S}$ characterized by an interaction matrix $\mathbf{A}$, can be represented by (Deng et al., 2022)

$$D_F(\mathcal{C},\mathcal{S}) = \left\{ \left(\sum_{j\in\mathcal{C}} N_j^* \mathbf{a}_j - \sum_{k\in\mathcal{S}\setminus\mathcal{C}} N_k^* \mathbf{e}_k \right) \in \mathbb{R}^{|\mathcal{S}|} : \mathbf{A} = -\left(\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_{|\mathcal{S}|}\right), N_{j,k}^* \in \mathbb{R}_{>0} \right\}, \tag{7}$$

where $\mathbf{e}_k$ is a unit vector whose k^{th} entry is 1 and 0 elsewhere. Figure 1B illustrates geometrically the feasibility domains of different sub-communities $\mathcal{C}$ (which equivalently indicate the exclusion of species j not contained in $\mathcal{C}$) within the same arbitrary community formed by three species as shown in Figure 1A. The column vectors of each interaction matrix determine the structure of the geometric partition (i.e., the feasibility domains). Specifically, when the vector of effective growth rates $\boldsymbol{\theta}$ is located in the pink region labeled with $\mathcal{C} = \{1,2,3\}$, all three species are able to persist in the long term. Similarly, when the vector of effective growth rates $\boldsymbol{\theta}$ is located in the blue region labeled with $\mathcal{C} = \{1,2\}$, species 1 and 2 can persist and species 3 will be excluded in the long term.

Having no prior information about the values of $\boldsymbol{\theta}$ (i.e., environmental uncertainty), we can again assume that all vector directions are equally likely and the feasibility domain $D_F(\mathcal{C},\mathcal{S})$ can be normalized by the volume of the entire parameter space (Saavedra et al., 2016). Then, the *persistence probability* over the long term of a composition $\mathcal{C}$ within a multispecies community $\mathcal{S}$ can be analytically defined under environmental uncertainty as (Deng et al., 2022)

$$F(\mathcal{C},\mathcal{S}) = \frac{\mathsf{vol}\left(D_F(\mathcal{C},\mathcal{S}) \cap \Theta^{|\mathcal{S}|}\right)}{\mathsf{vol}\left(\Theta^{|\mathcal{S}|}\right)}. \tag{8}$$

When $\mathcal{C} = \mathcal{S}$, Eq. (8) corresponds to the persistence probability of the entire community $\mathcal{S}$. Otherwise, when $\mathcal{C} \subset \mathcal{S}$, Eq. (8) corresponds to the persistence probability of a sub-community $\mathcal{C}$ only. Ecologically, the persistence probability indicates the fraction of feasible and compatible environmental contexts ($\boldsymbol{\theta}$) for each sub-community $\mathcal{C}$ within a specific community $\mathcal{S}$ relative to all equally possible contexts.

Integration of theories from a probabilistic perspective

To integrate invasion and coexistence theories, it is necessary to link short-term and long-term dynamics (Arnoldi et al., 2022, Hofbauer and Schreiber, 2022). From the invader's perspective (species i), a positive invasion growth rate (i.e., $g(i) > 0$, measured as a function of the effect of residents on the invader, Eq. (2)) guarantees that its population will increase immediately after its introduction. For invader i, the probability of a positive invasion growth rate is given by Eq. (6). However, the effect of a successful invader on the resident species demands knowledge of how residents will affect each other as well as how the invader will affect the growth rate of residents over the long run (Arnoldi et al., 2022, Hofbauer and Schreiber, 2022, Song et al., 2021).

Following our geometric representation in the parameter space of $\boldsymbol{\theta}$ (Fig. 1), we can illustrate the possible long-term effects of a successful invader on the resident species by the colored regions in Figure 1C for the hypothetical community (Figs. 1A-1B). The pink region, where the entire community (i.e., residents 1-2 and invader 3) can persist over the long term, corresponds to the domain of $\boldsymbol{\theta}$ compatible with *community augmentation*. The two blue regions, where either resident 1 or resident 2 are *excluded*, correspond to the domain compatible with the *replacement* of one resident. The yellow region, where both residents are excluded by the invader, corresponds to the domain compatible with the replacement of two residents. The area constrained by the dotted lines corresponds to the environmental contexts compatible with the persistence of residents in the absence of invader i—the possible starting contexts for invasion.

Similar to the persistence probability (Eq. (8)) in coexistence theory, we can define probabilities to describe the chances of obtaining different outcomes in an invasion process under environmental uncertainty. Formally, after invader i enters a resident community $\mathcal{S} \setminus \{i\}$, the probability of observing a species composition $\mathcal{C}$ ($\mathcal{C} \subseteq \mathcal{S}$) can be defined as

$$I_i(\mathcal{C}, \mathcal{S}) = \frac{\mathsf{vol}\left(D_F(\mathcal{C}, \mathcal{S}) \cap \boldsymbol{\theta}^*(\bar{i})\right)}{\mathsf{vol}\left(\boldsymbol{\theta}^*(\bar{i})\right)}, \tag{9}$$

where $\boldsymbol{\theta}^*(\bar{i})$ is the feasibility domain of residents in isolation (without invader i).

Translating simultaneous to sequential probabilistic outcomes

As an application of our integration of coexistence and invasion theories, we examine the extent to which we can use results from simultaneous assembly processes to infer results from sequential assembly processes. For example, under simultaneous assembly processes, our theoretical framework can quantify the probabilities of survival of a focal species, community persistence, and the exclusion of competitor species. Our theoretical framework can also describe analogous outcomes under sequential assembly processes, which correspond to the probabilities of invader colonization, invader-resident augmentation, and resident replacement, respectively. For illustrative purposes, we focus

on corresponding pairs of the aforementioned outcomes, but this study can be generalized to other outcomes.

Focusing on simultaneous assembly processes (Fig. 2, left column), the *probability of survival* of a focal species i ($i \in \mathcal{S}$) can be calculated by the sum of feasibility domains (Eq. (7)) of all sub-communities containing species i. Mathematically, this probability can be expressed as

$$P_s(i, \mathcal{S}) = \sum_{\mathcal{C}: i \in \mathcal{C}} F(\mathcal{C}, \mathcal{S}). \tag{10}$$

For example, considering a three-species community ($\mathcal{S} = \{1, 2, 3\}$), the probability of survival of the focal species 3 is the summation of four feasibility domains: $P_s(3, \{1, 2, 3\}) = F(\{3\}, \{1, 2, 3\}) + F(\{1, 3\}, \{1, 2, 3\}) + F(\{2, 3\}, \{1, 2, 3\}) + F(\{1, 2, 3\}, \{1, 2, 3\})$ (Fig. 2A, left). Next, the *probability of community persistence* can be calculated by the feasibility domain where all species in the community $\mathcal{S}$ can persist together. Formally, it can be written as

$$P_p(\mathcal{S}) = F(\mathcal{C} = \mathcal{S}, \mathcal{S}). \tag{11}$$

In this case, the probability of persistence of a three-species community ($\mathcal{S} = \{1, 2, 3\}$) is given by just one feasibility domain: $P_p(\{1, 2, 3\}) = F(\{1, 2, 3\}, \{1, 2, 3\})$ (Fig. 2B, left). Lastly, the *probability of exclusion* can be calculated by the difference between the probability of survival and the probability of community persistence (i.e., at least one non-focal species goes extinct, while the focal species survives). Mathematically, the probability of exclusion can be written as

$$P_e(i, \mathcal{S}) = \sum_{\mathcal{C}: i \in \mathcal{C}, \mathcal{C} \subset \mathcal{S}} F(\mathcal{C}, \mathcal{S}) = P_s(i, \mathcal{S}) - P_p(i, \mathcal{S}). \tag{12}$$

For example, considering a three-species community ($\mathcal{S} = \{1, 2, 3\}$) with focal species 3, the probability of excluding at least one non-focal species is given by three feasibility domains: $P_e(\{1, 2, 3\}) = F(\{3\}, \{1, 2, 3\}) + F(\{1, 3\}, \{1, 2, 3\}) + F(\{2, 3\}, \{1, 2, 3\})$ (Fig. 2C, left).

Shifting our focus to sequential assembly processes (Fig. 2, right column), the *probability of colonization* of an invader i can be calculated by the fraction of the feasibility domain $\boldsymbol{\theta}^*(\bar{i})$ of the residents within which the invasion growth rate is positive. This probabilistic outcome is analogous to the probability of survival (Eq. (10)) under simultaneous assembly processes. Formally, it can be expressed as

$$P_c(i, \mathcal{S}) = \frac{\mathsf{vol}\left((GI(i, \theta) > 0) \cap \boldsymbol{\theta}^*(\bar{i})\right)}{\mathsf{vol}\left(\boldsymbol{\theta}^*(\bar{i})\right)} = \sum_{\mathcal{C}: i \in \mathcal{C}} I_i(\mathcal{C}, \mathcal{S}), \tag{13}$$

where the second relation holds under globally stable gLV dynamics (Eq. (1)). Results are qualitatively similar for locally stable communities (Supplementary Section S3). For example, considering species 3 as an invader and species $\{1, 2\}$ as the resident community, the probability that the invader can successfully colonize is given by $P_c(3, \{1, 2, 3\}) = I_3(\{3\}, \{1, 2, 3\}) + I_3(\{1, 3\}, \{1, 2, 3\}) +$

$I_3(\{2,3\},\{1,2,3\}) + I_3(\{1,2,3\},\{1,2,3\})$ (Fig. 2A, right). Next, the *probability of community augmentation* can be calculated by the fraction of the feasibility domain of the residents where all species in the community $\mathcal{S}$ can persist together and the invasion growth rate for species 3 is positive. This probabilistic outcome is analogous to the probability of community persistence (Eq. (11)) under simultaneous assembly processes. This probability can be written as

$$P_a(\mathcal{S}) = I_i(\mathcal{C} = \mathcal{S}, \mathcal{S}). \tag{14}$$

For example, the probability of community augmentation, where invader 3 can persist with the resident community $\{1,2\}$, is $P_a(\{1,2,3\}) = I_3(\{1,2,3\},\{1,2,3\})$ (Fig. 2B, right). Lastly, the *probability of replacement* can be calculated by the sum of feasibility domains under which an invader i can replace at least one resident species. This probabilistic outcome is analogous to the probability of exclusion (Eq. (12)) under simultaneous assembly processes. Mathematically, the probability that an invader i replaces any resident is specified by

$$P_r(i,\mathcal{S}) = \sum_{\mathcal{C}:i\in\mathcal{C},\mathcal{C}\subset\mathcal{S}} I_i(\mathcal{C},\mathcal{S}) = P_c(i,\mathcal{S}) - P_a(i,\mathcal{S}). \tag{15}$$

For instance, the probability that invader 3 can replace at least one resident in the community $\{1,2\}$ is $P_r = I_3(\{3\},\{1,2,3\}) + I_3(\{1,3\},\{1,2,3\}) + I_3(\{2,3\},\{1,2,3\})$.

Importantly, our geometric framework can be used to gain insights about the extent to which probabilistic outcomes derived from simultaneous assembly process can be used to obtain information regarding analogous outcomes under sequential assembly processes. Figure 2 shows the geometric illustration of the three aforementioned pairs of analogous probabilities for hypothetical three-species communities ($\mathcal{S} = \{1,2,3\}$). A similar geometric intuition can be applied for higher dimensions. The two spheres in each row correspond to the same three-species community $\mathcal{S}$ (i.e., the same interaction matrix $\mathbf{A}$), with specific but fairly arbitrary values of the interactions a_{ij}, illustrating the possibility of translation between the relevant quantities. In each panel of Figure 2, we represent an original and changed community at the top and bottom, respectively—where in each case the change corresponds to a modification of certain elements of the interaction matrix $\mathbf{A}$.

Figure 2A illustrates the pair formed by survival probability (Eq. (10)) and colonization probability (Eq. (13)). In this hypothetical case, an increase in the colonization probability is generally associated with an increase in the survival probability (Fig. 2A, change shown by red arrows). For example, lower competitive effects of residents on an invader, which move the lower boundaries of the relevant regions downward in the lower spheres indicating the changed systems, increase the region of positive invasion growth rate and are associated with an increase in both probabilities.

Hence, we can expect a strong translatability between this pair of analogous probabilities. Figure 2B shows the pair formed by persistence probability (Eq. (11)) and augmentation probability (Eq.

(14)). In this second case, if the persistence probability increases, the augmentation probability does not necessarily increase (Fig. 2B, change shown by red arrows). For example, higher competitive effects of an invader on residents (as shown by lower spheres in Fig. 2B) increase the region of community persistence, but not the region of positive invasion growth rate. This difference leads to an expected weaker translatability between this pair of analogous probabilities. Fig. 2C illustrates the pair formed by exclusion probability (Eq. (12)) and replacement probability (Eq. (15)). In this third case, the replacement probability can be significantly smaller than the exclusion probability. For example, higher competitive effects of an invader on residents (as shown in lower spheres in Fig. 2C) decrease the exclusion probability, but can eliminate the possibility of replacement altogether (change shown by red arrows). Furthermore, while this quantity is simply the difference of the other two quantities, the strongest correlative effect associated with changes in the effect of residents on the invader largely cancels. Taken together these effects suggest an expected weaker translatability between replacement and exclusion probabilities. In sum, Figure 2 indicates a translatability between survival and colonization probabilities, but some loss of translation between persistence and augmented probabilities and even less correlation between exclusion and replacement probabilities. Note that our notion of translatability can be expressed in terms of predictability as correlations of computations of specific quantities given some known information about the system (e.g., this can be specified by the coefficient of determination, R^2).

In silico experiments

Next, we provide an *in silico* experiment as a way to illustrate the inputs and outputs required to corroborate our theoretical integration. Specifically, we perform numerical tests using simulations to confirm the loss of translatability between simultaneous and sequential assembly processes established by our geometric framework. To this end, we generate 100 *in silico* communities of three species, whose interspecific effects (off-diagonal entries of $\mathbf{A}$) are randomly drawn from a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 0.5$. The intraspecific effects are set to $a_{ii} = -1$. Thus, all three-species communities are globally stable under the gLV dynamics (Eq. (1)), making the results independent from the species' initial abundances. These values are in line with the geometric illustration in Figure 2, but the qualitative results are robust to this choice (Supplementary Section S2).

All simultaneous and sequential outcomes from these *in silico* communities can be calculated either analytically or numerically using their corresponding definitions (frequencies) established in Eqs. (8-10). The analytic calculation, however, is valid only in the idealized theoretical limit of an infinite number of experimental trials and time steps (Deng et al., 2022). Hence, we use a numerical approximation using simulations to make our analysis more comparable to the finite nature of lab/field experiments or observations (Deng et al., 2022). For this purpose, for each community $\mathbf{A}$, we randomly and uniformly sample 1000 vectors of effective growth rates $\boldsymbol{\theta}$ inside the feasibility domain of residents in isolation $\boldsymbol{\theta}^*(\bar{i})$. Each one of these vectors mimics an experimental trial initiated under different environmental

conditions where it is known that residents can all coexist. Under global stability, species with a positive equilibrium abundance ($\mathbf{N}^* = -\mathbf{A}^{-1}\boldsymbol{\theta} > 0$) can be classified as persistent (Takeuchi, 1996). This classification can also be reached via simulations after a given period of time assuming a given threshold for extinction (Deng et al., 2022). We simulate the gLV model (Eq. (1)) with total time 200, step size 10^{-2}, and an extinction threshold of 10^{-6}. Simulations are conducted using the Runge-Kutta method. This threshold may also correspond to the resolution level at which species can be detected in the lab or in the field (Deng et al., 2022). Then, the probability $I_i(\mathcal{C}, \mathcal{S})$ of a composition $\mathcal{C}$ within a multispecies community $\mathcal{S}$ can be approximated by the frequency with which all species in the composition $\mathcal{C}$ persist together.

We use the coefficient of determination R^2 to measure the translatability of analogous outcomes between simultaneous and sequential assembly processes. Panels A-C in Fig. 3 show, respectively, the relationship between survival probability and colonization probability ($R^2 = 0.98$), persistence probability and augmentation probability ($R^2 = 0.74$), and exclusion probability and replacement probability ($R^2 = 0.58$). Importantly, this figure confirms the translatability between analogous outcomes expected from the geometric analysis (Fig. 2). That is, if one aims to use the outcomes from simultaneous assembly processes to predict the outcomes of sequential assembly processes (or vice versa), the only highly reliable case would be the translation between survival and colonization probabilities.

Conclusions

Linking simultaneous and sequential assembly processes has been a fundamental scientific problem, underpinning our ability to predict the rise and fall of populations forming part of natural ecological communities (Fukami, 2015, Hastings, 2004, Odum, 1969, Vellend, 2016). Recent efforts to integrate these two types of assembly processes have relied on the assumption that it is possible to fully parameterize ecological models (Arnoldi et al., 2022, Hofbauer and Schreiber, 2022), the assumption that environmental conditions remain the same across the entire assembly process (Arnoldi et al., 2022, Hofbauer and Schreiber, 2022, Logofet, 1993, Serván and Allesina, 2021), or the assumption that one has sufficient information about coexistence or invasion outcomes (Friedman et al., 2017, Grainger and Gilbert, 2019, Maynard et al., 2020). While these studies have provided key insights regarding the association of simultaneous and sequential assembly, these assumptions are seldom met in nature (Fukami, 2015, Hubbell, 1997, Miller and Allesina, 2021, Song et al., 2021). Therefore, it has become central to understand the extent to which the probabilistic outcomes from such simultaneous processes are representative of analogous outcomes expected from more complex sequential processes. This translation, however, requires the non-trivial integration of coexistence and invasion theories under a probabilistic perspective.

To study the limits to invasion prediction using coexistence outcomes, we have integrated coexistence and invasion outcomes using a geometric and probabilistic framework (Fig. 1). We have derived this

theoretical framework from generalized Lotka-Volterra dynamics assuming that per capita growth rates are a function of the environment, considering a one-to-one mapping between the environment and per capita growth rates, considering no a priori information about the environment (i.e., environmental uncertainty), and assuming that species interactions constrain the set of per capita growth rates compatible with coexistence and invasion outcomes (following an structuralist perspective). As an application, we have used this framework to study the translatability from simultaneous assembly processes into sequential assembly processes. We have provided probabilistic measures (Fig. 2) of simultaneous assembly outcomes (i.e., individual survival, community persistence, competitor exclusion) and sequential assembly outcomes (i.e., invader colonization, invader-resident augmentation, resident replacement). We have shown that the translatability between the probability of analogous outcomes can be valid between survival and colonization. However, this translatability becomes weaker between persistence and augmentation, as well as between exclusion and replacement (Fig. 3). This suggests that outcomes from sequential processes involving more than one species should be treated differently from their simultaneous counterparts. These results further support empirical observations showing a clear separation between simultaneous and sequential assembly processes (Angulo et al., 2021, Deng et al., 2021). Importantly, our integration provides a guiding and testable theoretical framework regarding the (lack of) translatability of outcomes between simultaneous and sequential processes—which is relevant for a better understanding of systems in nature.

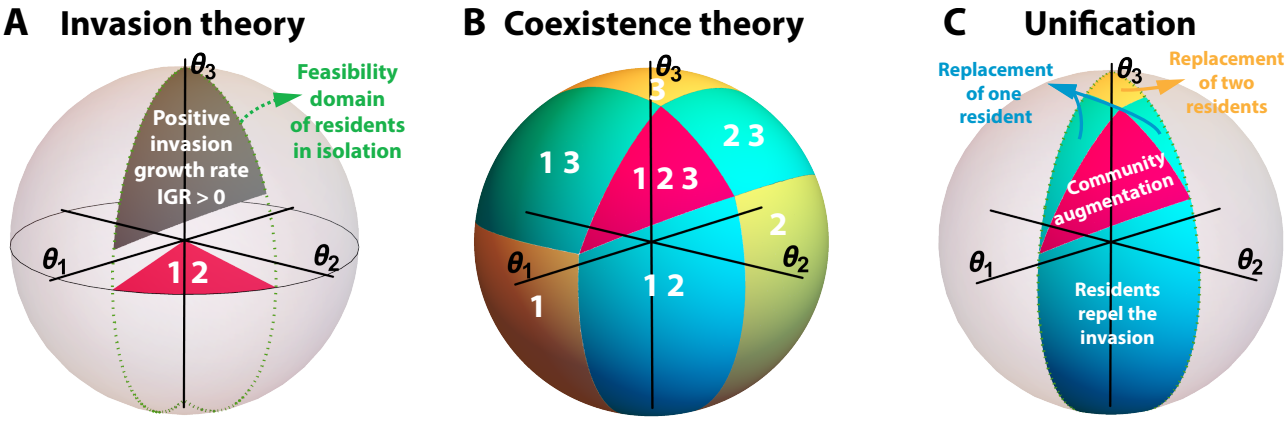


Figure 1: **Geometric and probabilistic framework under generalized Lotka-Volterra dynamics.** The figure illustrates our framework for a hypothetical three-species community. The sphere represents the parameter space of effective growth rates $\boldsymbol{\theta}$, in which we color-code the regions compatible with feasible solutions ($N_i^* > 0, \forall i$) for different sets of species under the generalized Lotka-Volterra dynamics (Eq. (1)). In this illustration, residents are labeled $\{1, 2\}$ and the invader species $\{3\}$. Panel A shows the feasibility domain (pink area) of $\boldsymbol{\theta}^*(\bar{3})$ for residents in the absence of the invader. Note that this region is on the plane (θ_1, θ_2). The area spanned by the dotted green lines represents the region of the three-dimensional sphere that projects into this 2D feasibility region. Then, the gray area corresponds to the region compatible with a positive invasion growth rate for the invader (Eqs. (2)-(6)) assuming that the pre-existing residents are at a feasible equilibrium (area constrained inside the green dotted lines). Panel B illustrates the hypothetical feasibility domains compatible with the long-term persistence of different collections of species (labeled on top of each region) assuming that all directions of $\boldsymbol{\theta}$ are possible (Eqs. (7)-(8)). Panel C shows the possible long-term persistence outcomes under a successful or unsuccessful short-term invasion (combining previous panels): community augmentation, replacement of one or two residents, and residents repelling the invasion (Eq. (9)). The possible outcomes are restricted to the area spanned by the feasibility of residents in isolation.

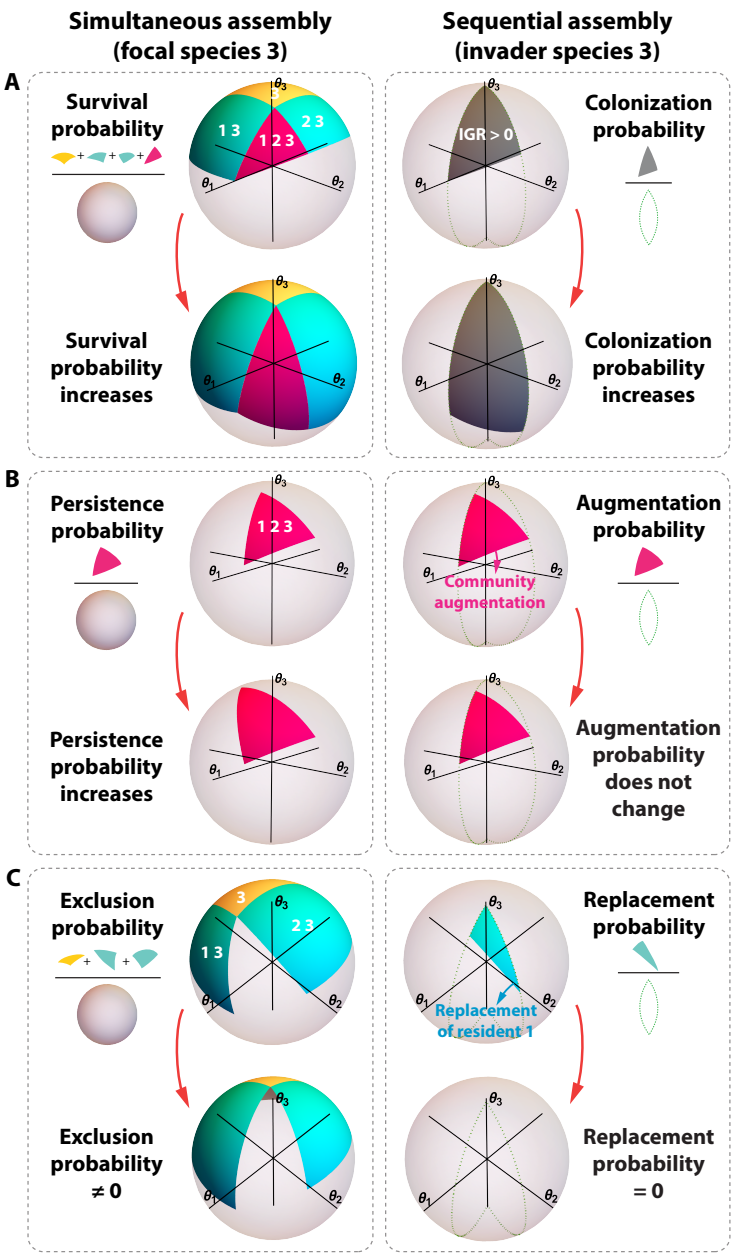


Figure 2: **Geometric intuition on the translatability of analogous outcomes between simultaneous and sequential assembly processes.** For a hypothetical three-species community, each panel (row) illustrates the relationship of analogous probabilistic outcomes between simultaneous (left column, Eqs. (10)-(12)) and sequential (right column, Eqs. (13)-(15)) assembly processes. The two spheres in each row correspond to the same community $\mathcal{S}$ (i.e., the same interaction matrix $\mathbf{A}$). Inside each panel, we represent an original and changed community (i.e., a different interaction matrix $\mathbf{A}$) at the top and bottom, respectively. Panel A shows the survival probability of focal species 3 (left, Eq. (10)) and the colonization probability of invader species 3 (right, Eq. (13)). In this example, both colonization and survival probabilities increase when residents decrease their competitive effect on the invader (change shown by red arrows). Panel B shows the persistence probability of the community (left, Eq. (11)) and the augmentation probability of the community (right, Eq. (14)). In this example, the persistence probability increases and the augmentation probability does not change when the competitive effects of the invader on residents increase (change shown by red arrows). Panel C shows the exclusion probability that focal species 3 excludes non-focal species {1,2} (left, Eq. (12)) and the replacement probability that invader species 3 replaces resident species {1,2} (right, Eq. (15)). By increasing the competitive effect of the invader on residents, this panel shows that the replacement probability is geometrically a small fraction of the exclusion probability (top) and can become zero when the exclusion probability is significantly greater than zero (bottom). Based on the geometric intuition illustrated here, there is an expected strong translatability between survival and colonization probabilities, but an expected weaker translatability between persistence and augmentation probabilities as well as between exclusion and replacement probabilities.

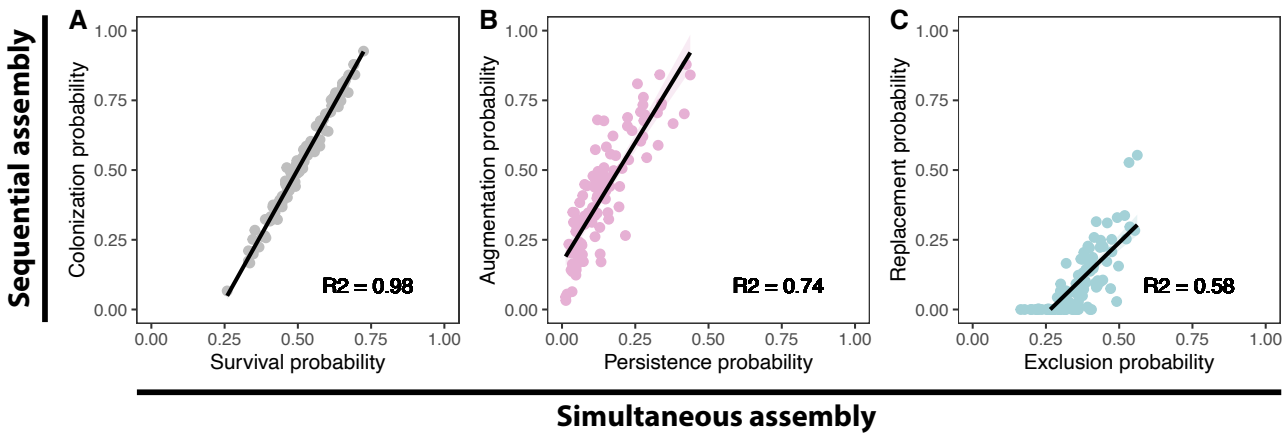


Figure 3: *In silico* **experiment corroborating the (lack of) translatability between simultaneous and sequential assembly processes.** The figure illustrates the numerical results using gLV simulations of three pairs of analogous probabilistic outcomes using 100 randomly generated three-species communities (see main text for details). The x-axes and y-axes correspond to the outcomes under simultaneous (Eqs. (10)-(12)) and sequential (Eqs. (13)-(15)) assembly processes, respectively. Each point corresponds to a three-species community, whose interaction matrix $\mathbf{A}$ is randomly generated using a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 0.5$ (diagonal entries $a_{ii} = -1$). We also tested other parameters (e.g., community sizes, standard deviations) and obtained qualitatively the same results (Supplementary Section S2). These numerical tests confirm the expected strong translatability between survival and colonization probabilities (Panel A: $R^2 = 0.98$), and the expected weaker translatability between persistence and augmentation probabilities (Panel B: $R^2 = 0.74$) as well as between exclusion and replacement probabilities (Panel C: $R^2 = 0.58$).

References

AlAdwani, M. and Saavedra, S. 2020. Ecological models: higher complexity in, higher feasibility out. J. of the Roy. Soc. Interface 17:20200607.

Alberch, P. 1989. The logic of monsters: Evidence for internal constraint in development and evolution. Geobios 22:21–57.

Angulo, M. T., Kelley, A., Montejano, L., Song, C., and Saavedra, S. 2021. Coexistence holes characterize the assembly and disassembly of multispecies systems. Nature Ecology & Evolution 5:1091–1101.

Arnoldi, J. F., Barbier, M., Kelly, R., Barabás, G., and Jackson, A. L. 2022. Invasions of ecological communities: Hints of impacts in the invader's growth rate. Methods in Eco. and Evo. 13:167–182.

Barabás, G., D'Andrea, R., and Stump, S. M. 2018. Chesson's coexistence theory. Ecological Monographs 88:277–303.

Barner, A. K., Coblentz, K. E., Hacker, S. D., and Menge, B. A. 2018. Fundamental contradictions among observational and experimental estimates of non-trophic species interactions. Ecology 99:557–566.

Case, T. J. 1990. Invasion resistance arises in strongly interacting species-rich model competition communities. Proc. Natl. Acad. Sci. U.S.A. 87:9610–9614.

Case, T. J., 2000. An Illustrated Guide to Theoretical Ecology. Oxford Univ. Press, Oxford.

Chang, C. Y., Vila, J. C. C., Bender, M., Li, R., Mankowski, M. C., Bassette, M., Borden, J., Golfier, S., Sanchez, P. G. L., Waymack, R., Zhu, X., Diaz-Colunga, J., Estrela, S., Rebolleda-Gomez, M., and Sanchez, A. 2021. Engineering complex communities by directed evolution. Nature Eco & Evo 5:1011–1023.

Chesson, P. and Kuang, J. J. 2008. The interaction between predation and competition. Nature 456:235–238.

Clark, R. L., Connors, B. M., Stevenson, D. M., Hromada, S. E., Hamilton, J. J., Amador-Noguez, D., and Venturelli, O. S. 2021. Design of synthetic human gut microbiome assembly and butyrate production. Nature Communications 12:1–16.

Deng, J., Angulo, M. T., and Saavedra, S. 2021. Generalizing game-changing species across microbial communities. ISME Communications 1:1–8.

Deng, J., Taylor, W., and Saavedra, S. 2022. Understanding the impact of third-party species on pairwise coexistence. PLoS Comput. Biol. 18:e1010630.

Drake, J. A. 1991. Community assembly mechanics and the structure of an experimental species ensemble. Am. Nat. 137:1–26.

Friedman, J., Higgins, L. M., and Gore, J. 2017. Community structure follows simple assembly rules in microbial microcosms. Nature Ecology & Evolution 1:0109.

Fukami, T. 2015. Historical contingency in community assembly: integrating niches, species pools, and priority effects. Annual Review of Ecology, Evolution, and Systematics 46.

Gause, G. F. 1932. Experimental studies on the struggle for existence. Journal of Experimental Biology 9:389–402.

Gilpin, M. E., Carpenter, M. P., and Pomerantz, M. J., 1986. The assembly of a laboratory community: multispecies competition in Drosophila, Pages 23–40. Harper and Row, New York.

Gould, A. L., Zhang, V., Lamberti, L., Jones, E. W., Obadia, B., Korasidis, N., Gavryushkin, A., Carlson, J. M., Beerenwinkel, N., and Ludington, W. B. 2018. Microbiome interactions shape host fitness. Proceedings of the National Academy of Sciences 115:E11951–E11960.

Grainger, T. N. and Gilbert, J. M. L. B. 2019. The invasion criterion: A common currency for ecological research. Trends in Ecology and Evolution 34:925–935.

Hastings, A. 2004. Transients: the key to long-term ecological understanding? TREE 19:39–45.

Hill, M. F., Witman, J. D., and Caswell, H. 2004. Markov chain analysis of succession in a rockysubtidal community. The Am. Nat. 164:E46–E61.

Hofbauer, J. and Schreiber, S. J. 2022. Permanence via invasion graphs: Incorporating community assembly into modern coexistence theory. Journal of Mathematical Biology 85:54.

Hofbauer, J. and Sigmund, K., 1998. Evolutionary games and population dynamics. Cambridge University Press.

Hubbell, S. P. 1997. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. Coral reefs 16:S9–S21.

Jones, E. W., Carlson, J. M., Sivak, D. A., and Ludington, W. B. 2022. Stochastic microbiome assembly depends on context. Proc. nat. Acad. Sci. USA 119:e2115877119.

Law, R. and Morton, R. D. 1996. Permanence and the assembly of ecological communities. Ecology 77:762–775.

Logofet, D. O., 1993. Matrices and Graphs: Stability Problems in Mathematical Ecology. CRC Press.

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

Maynard, D. S., Miller, Z. R., and Allesina, S. 2020. Predicting coexistence in experimental ecological communities. Nature Ecology & Evolution 4:91–100.

Miller, Z. R. and Allesina, S. 2021. Metapopulations with habitat modification. PNAS 118:e2109896118.

Odum, E. P. 1969. The strategy of ecosystem development. Science 164:262–270.

Odum, E. P. and Barrett, G. W., 2005. Fundamentals of Ecology. Thomson, Brooks-Cole, Fifth Edition.

Park, T. 1954. Experimental studies of interspecies competition ii. temperature, humidity, and competition in two species of tribolium. Physiological Zoology 27:177–238.

Rohr, R. P., Saavedra, S., and Bascompte, J. 2014. On the structural stability of mutualistic systems. Science 345:1253497.

Rohr, R. P., Saavedra, S., Peralta, G., Frost, C. M., Bersier, L.-F., Bascompte, J., and Tylianakis, J. M. 2016. Persist or produce: a community trade-off tuned by species evenness. Am. Nat. 188:411–422.

Saavedra, S., Medeiros, L. P., and AlAdwani, M. 2020. Structural forecasting of species persistence under changing environments. Ecology Letters 23:1511–1521.

Saavedra, S., Rohr, R. P., Bascompte, J., Godoy, O., Kraft, N. J., and Levine, J. M. 2017. A structural approach for understanding multispecies coexistence. Ecological Monographs 87:470–486.

Saavedra, S., Rohr, R. P., Olesen, J. M., and Bascompte, J. 2016. Nested species interactions promote feasibility over stability during the assembly of a pollinator community. Ecology and Evolution 6:997–1007.

Schreiber, S. J. and Rittenhouse, S. 2004. From simple rules to cycling in community assembly. Oikos 105:349–358.

Serván, C. A. and Allesina, S. 2021. Tractable models of ecological assembly. Ecol Letts. 24:1029–1037.

Song, C., Fukami, T., and Saavedra, S. 2021. Untangling the complexity of priority effects in multi-species communities. Ecology Letters 24:2301–2313.

Takeuchi, Y., 1996. Global dynamical properties of Lotka-Volterra systems. World Scientific.

Vagnon, C., Rohr, R. P., Bersier, L.-F., Cattanéo, F., Guillard, J., and Frossard, V. 2022. Combining food web theory and population dynamics to assess the impact of invasive species. Frontiers in Ecology and Evolution page 679.

Vandermeer, J. H. 1975. Interspecific competition: a new approach to the classical theory. Science 188:253–255.

Vellend, M., 2016. The theory of ecological communities. Princeton University Press, NJ.

Warren, P. H. and Spencer, M. 1996. Community and food-web responses to the manipulation of energy input and disturbance in small ponds. Oikos Pages 407–418.

Williamson, M. and Fittera, A. 1996. The varying success of invaders. Ecology 77:1661–1666.

Zhao, N., Saavedra, S., and Liu, Y. Y. 2021. The impact of colonization history on the composition of ecological systems. Phys. Rev. E 108:052403.

Online Appendix for

Integrating coexistence and invasion theories under a probabilistic framework

Jie Deng[1], Washington Taylor[2], Simon A. Levin[3,4], Serguei Saavedra[1]

[1]Department of Civil and Environmental Engineering, MIT,
77 Massachusetts Av., 02139 Cambridge, MA, USA
[2]Center for Theoretical Physics, MIT,
77 Massachusetts Av., 02139 Cambridge, MA, USA
[3]Department of Ecology & Evolutionary Biology, Princeton University,
Princeton, NJ 08544 USA
[4]High Meadows Environmental Institute, Princeton University,
Princeton, NJ 08544 USA

Contents

S1 Growth indicator **S2**

S2 Robustness of results **S5**

S3 Locally stable communities **S9**

S1 Growth indicator

Writing the generalized Lotka-Volterra model (Eq. (1)) in the matrix form $\frac{\mathrm{d}N}{\mathrm{d}t} = N(\boldsymbol{\theta} + \mathbf{A}N)$, the interaction matrix $\mathbf{A}$ is an $|\mathcal{S}| \times |\mathcal{S}|$ matrix, i.e.,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1|\mathcal{S}|} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2|\mathcal{S}|} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3|\mathcal{S}|} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{|\mathcal{S}|1} & a_{|\mathcal{S}|2} & a_{|\mathcal{S}|3} & \cdots & a_{|\mathcal{S}||\mathcal{S}|} \end{pmatrix} = (\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_{|\mathcal{S}|}).$$

We assume, without loss of generality, species 1 as the invader and the other $(|\mathcal{S}| - 1)$ species $\{2, \cdots, |\mathcal{S}|\}$ as the residents. Let $\mathbf{n}$ denote one of the normal vectors that is orthogonal to the facet $F(i)$ (also called hyperface) spanned by the negative of the 2nd to $|\mathcal{S}|^{\text{th}}$ column vectors of the interaction matrix $\mathbf{A}$ (i.e., $-\mathbf{a}_2, -\mathbf{a}_3, \cdots, -\mathbf{a}_{|\mathcal{S}|}$). Note that the facet and its normal vector in the main text are denoted as $F(i)$ and $\mathbf{n}(i)$, respectively (Eq. (3)) because the facet $F(i)$ and its normal vector $\mathbf{n}(i)$ can change with invader i. Here, for the convenience of reading, we simplify the facet and normal vector corresponding to invader 1 as F and $\mathbf{n}$. Specifically, any normal vector $\mathbf{n}$ that is orthogonal to the facet F has zero inner products with the vectors on the facet F. In addition, for vectors on the same side of the facet F, their inner products with the normal vector $\mathbf{n}$ to the facet F have the same sign. Moreover, to represent the facet F, we use the negative of the column vectors of the interaction matrix $\mathbf{A}$ which are also the spanning vectors of the feasibility domain $D_F(\mathcal{C}, \mathcal{S})$ for easier geometric reference (see Eq. (7)).

Thus by definition, the following inner products are zero

$$\begin{cases} -\mathbf{a}_2 \cdot \mathbf{n} = 0, \\ -\mathbf{a}_3 \cdot \mathbf{n} = 0, \\ \quad \cdots \\ -\mathbf{a}_{|\mathcal{S}|} \cdot \mathbf{n} = 0, \end{cases} \tag{S1}$$

which is equivalent to the form removing the negative sign

$$\begin{cases} \mathbf{a}_2 \cdot \mathbf{n} = 0, \\ \mathbf{a}_3 \cdot \mathbf{n} = 0, \\ \quad \cdots \\ \mathbf{a}_{|\mathcal{S}|} \cdot \mathbf{n} = 0. \end{cases} \tag{S2}$$

Let us denote $\mathbf{n} = (n_1, n_2, \cdots, n_{|\mathcal{S}|})^\intercal$, then the matrix form of Eq. (S2) can be written as follows:

$$\begin{pmatrix} \mathbf{a}_2^\intercal \\ \mathbf{a}_3^\intercal \\ \vdots \\ \mathbf{a}_{|\mathcal{S}|}^\intercal \end{pmatrix} \mathbf{n} = \begin{pmatrix} \mathbf{a}_2^\intercal \\ \mathbf{a}_3^\intercal \\ \vdots \\ \mathbf{a}_{|\mathcal{S}|}^\intercal \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_{|\mathcal{S}|} \end{pmatrix} = \begin{pmatrix} a_{12} & a_{22} & a_{32} & \cdots & a_{|\mathcal{S}|2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{|\mathcal{S}|3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1|\mathcal{S}|} & a_{2|\mathcal{S}|} & a_{3|\mathcal{S}|} & \cdots & a_{|\mathcal{S}||\mathcal{S}|} \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_{|\mathcal{S}|} \end{pmatrix} = \mathbf{0}. \tag{S3}$$

Next, we are going to simplify the matrix form Eq. (S3). First, let us set $\mathbf{a}_{inv} = (a_{12},\ a_{13},\ \cdots,\ a_{1|\mathcal{S}|})$, $\mathbf{A}_{res} = \begin{pmatrix} a_{22} & a_{23} & \cdots & a_{2|\mathcal{S}|} \\ a_{32} & a_{33} & \cdots & a_{3|\mathcal{S}|} \\ \vdots & \vdots & \ddots & \vdots \\ a_{|\mathcal{S}|2} & a_{|\mathcal{S}|3} & \cdots & a_{|\mathcal{S}||\mathcal{S}|} \end{pmatrix}$, $\mathbf{n}_{res} = (n_2,\ \cdots,\ n_{|\mathcal{S}|})^\intercal$, then Eq. (S3) can be written as

$$\begin{pmatrix} \mathbf{a}_{inv}^\intercal & \mathbf{A}_{res}^\intercal \end{pmatrix} \begin{pmatrix} n_1 \\ \mathbf{n}_{res} \end{pmatrix} = n_1 \mathbf{a}_{inv}^\intercal + \mathbf{A}_{res}^\intercal \mathbf{n}_{res} = \mathbf{0}. \tag{S4}$$

Hence,

$$\mathbf{n}_{res} = (\mathbf{A}_{res}^\intercal)^{-1}(-n_1 \mathbf{a}_{inv}^\intercal) = -n_1 (\mathbf{A}_{res}^\intercal)^{-1} \mathbf{a}_{inv}^\intercal = -n_1 (\mathbf{A}_{res}^{-1})^\intercal \mathbf{a}_{inv}^\intercal = -n_1 (\mathbf{a}_{inv} \mathbf{A}_{res}^{-1})^\intercal. \tag{S5}$$

For an arbitrary effective growth rate $\boldsymbol{\theta}$, we define the *growth indicator* as

$$GI(1, \theta) = \frac{\boldsymbol{\theta} \cdot \mathbf{n}}{\mathbf{e}_1 \cdot \mathbf{n}}, \tag{S6}$$

where $\mathbf{e}_1 = (1, 0, \ldots, 0)^\intercal$. Apparently, a positive $GI(1, \theta)$ means that the inner products of the effective growth rate $\boldsymbol{\theta}$ and the unit vector $\mathbf{e}_1$ with the normal vector $\mathbf{n}$ to the facet F have the same sign. This can be geometrically interpreted as the effective growth rate $\boldsymbol{\theta}$ and the unit vector $\mathbf{e}_1$ are on the same side of the facet F.

Obviously, $\mathbf{e}_1 \cdot \mathbf{n} = n_1$. Thus the growth indicator can be written as

$$GI(1, \theta) = \frac{\theta_1 n_1 + \boldsymbol{\theta}_{res} \cdot \mathbf{n}_{res}}{n_1} = \theta_1 + \frac{\mathbf{n}_{res}^\intercal \boldsymbol{\theta}_{res}}{n_1}, \tag{S7}$$

where $\boldsymbol{\theta}_{res} = (\theta_2,\ \cdots,\ \theta_{|\mathcal{S}|})^\intercal$.

Substitute $\mathbf{n}_{res}$ by Eq. (S5), then

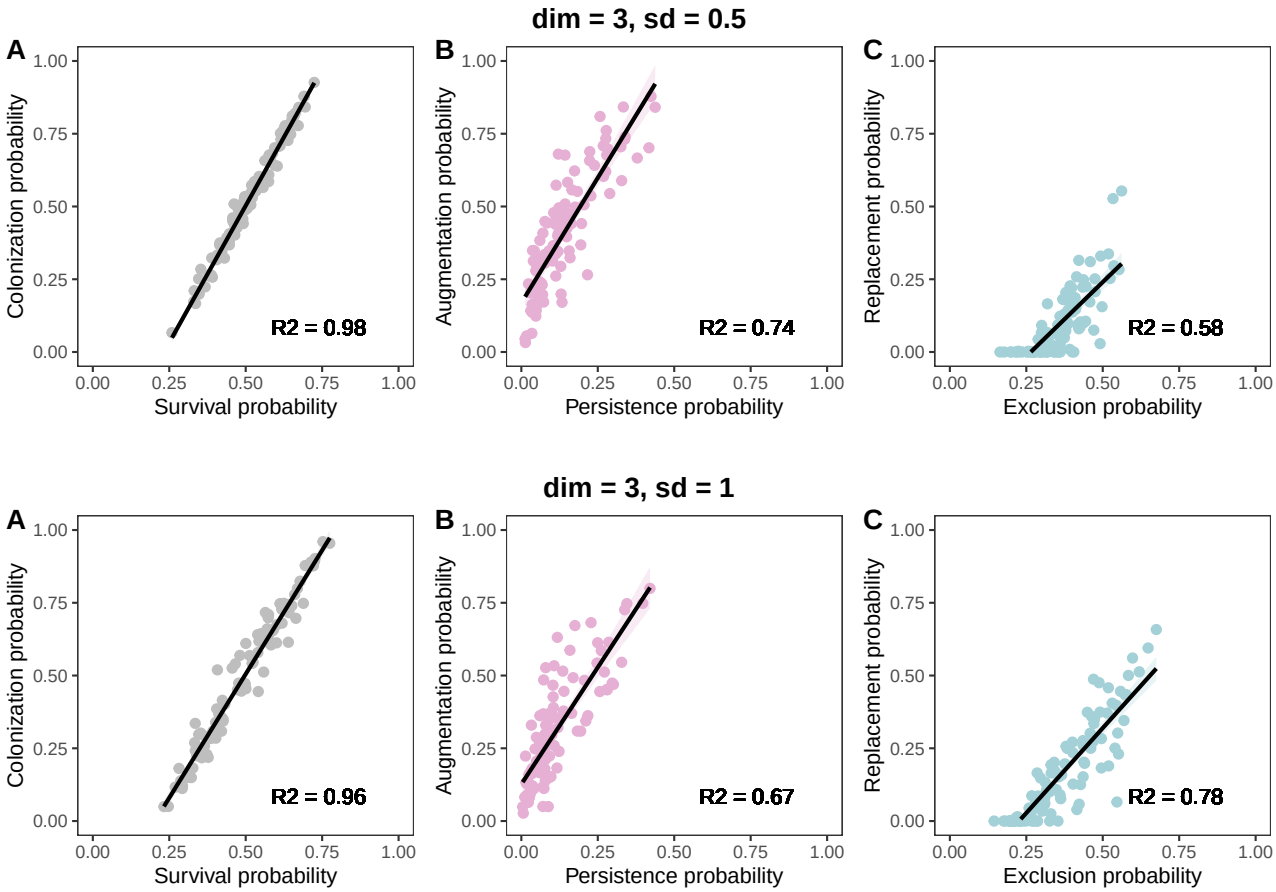
$$GI(1, \theta) = \theta_1 + \frac{(-n_1 (\mathbf{a}_{inv} \mathbf{A}_{res}^{-1})^\intercal)^\intercal \boldsymbol{\theta}_{res}}{n_1} = \theta_1 - \mathbf{a}_{inv} \mathbf{A}_{res}^{-1} \boldsymbol{\theta}_{res} = \theta_1 + \mathbf{a}_{inv} \mathbf{N}_{res}^*, \tag{S8}$$

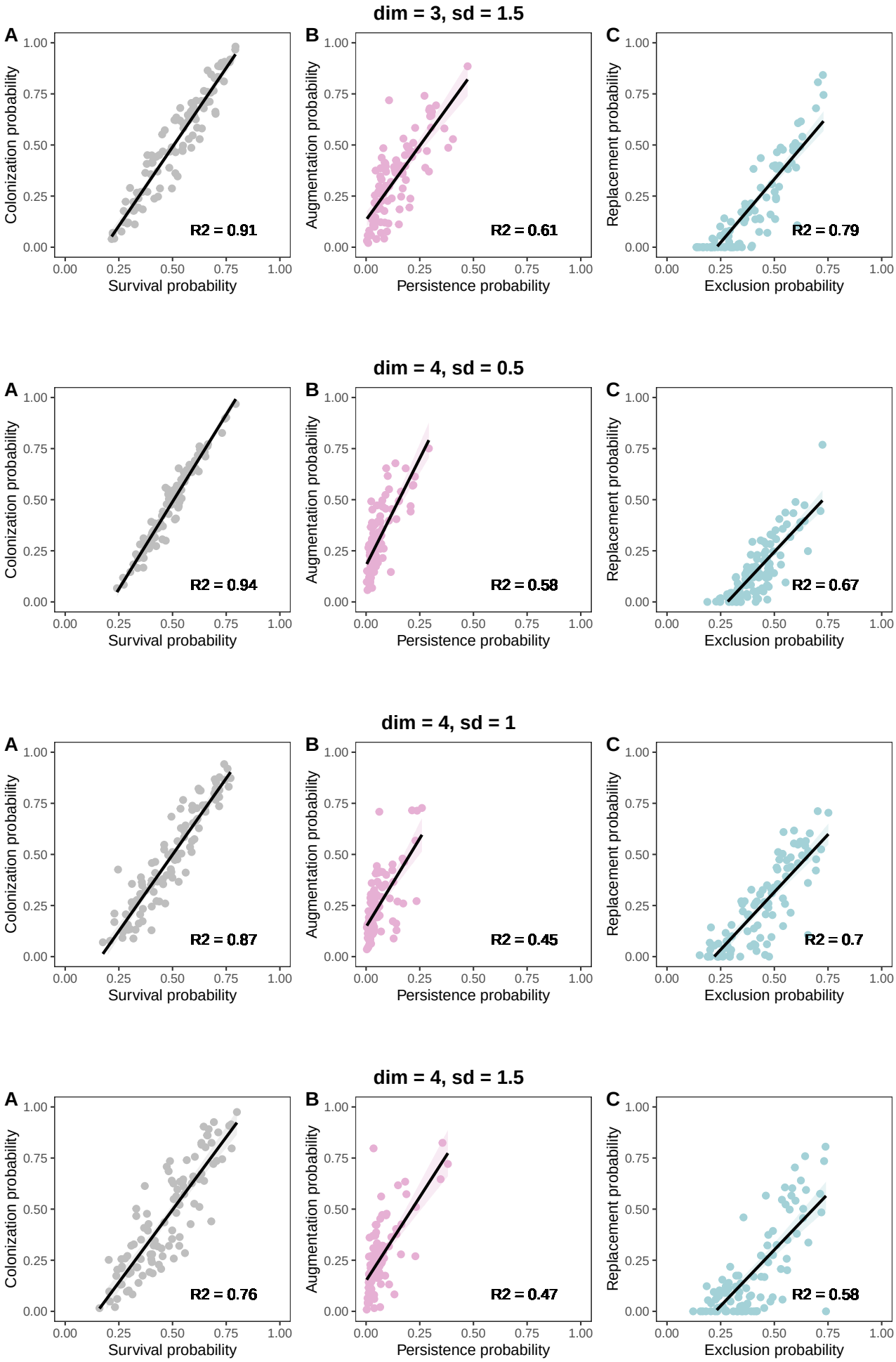
where $\mathbf{N}_{res}^* = -\mathbf{A}_{res}^{-1} \boldsymbol{\theta}_{res} = (N_2^*,\ N_3^*,\ \cdots,\ N_{|\mathcal{S}|}^*)^\intercal$ is the interior equilibrium of the resident community

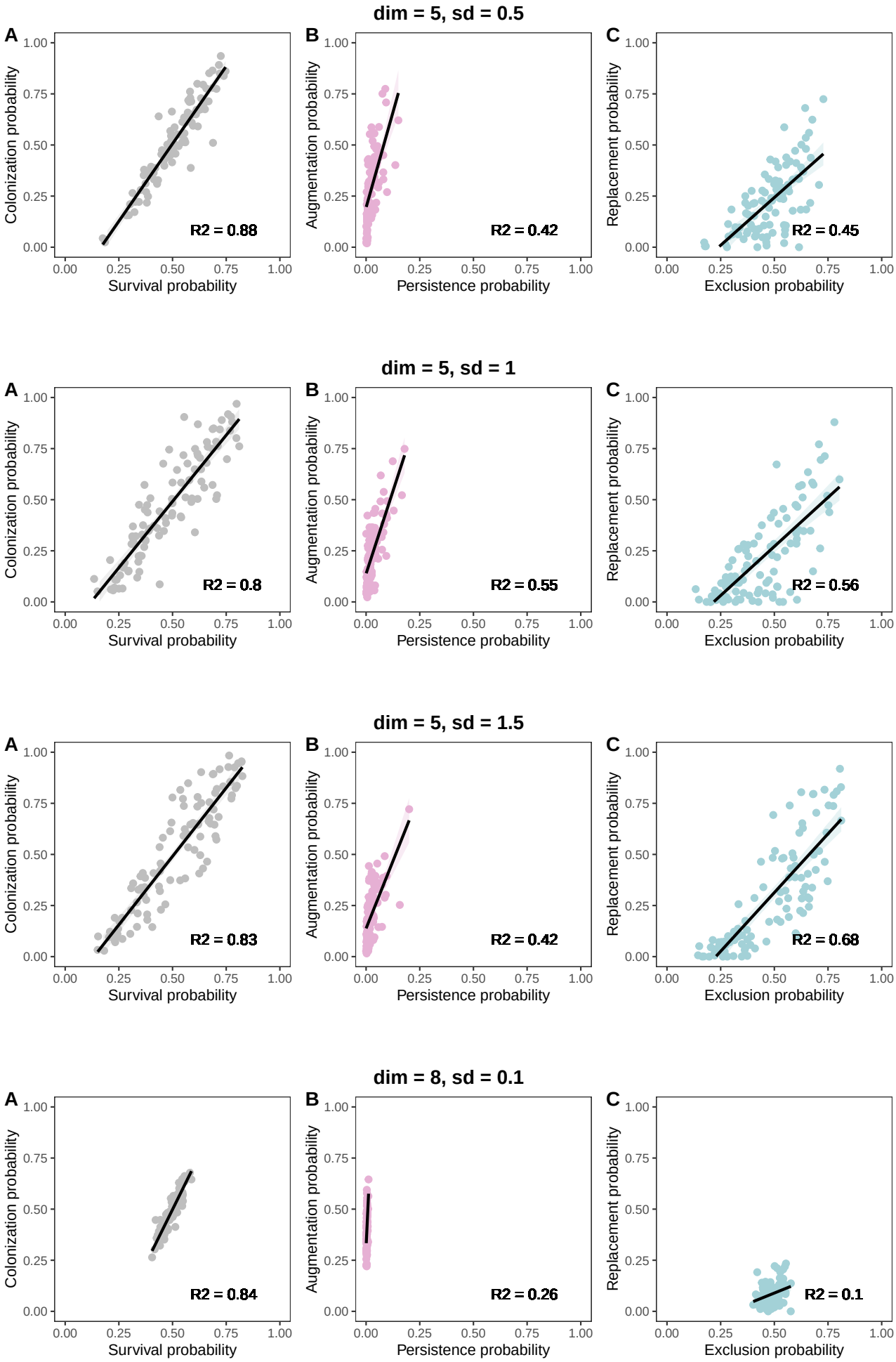
$\{2,\ 3,\ \cdots,\ |\mathcal{S}|\}$ in isolation. Note that the right side of Eq. (S8) is exactly the invasion growth rate $g(1)$. Hence we prove that the growth indicator $GI(1,\theta)$ is equivalent to the invasion growth rate $g(1)$. Since this proof also applies to a general invader i in any dimension, it is safe to conclude that the growth indicator $GI(i,\theta)$ is equivalent to the invasion growth rate $g(i)$. Therefore, the geometric representation of the region with positive invasion growth rate (i.e., $g(i) > 0$) is on the same side as the unit vector $\mathbf{e}_i$ (whose i^{th} entry is 1 and 0 elsewhere) relative to the facet $F(i)$ spanned by the $(|\mathcal{S}| - 1)$ column vectors (except for the ith column vector) of the interaction matrix $\mathbf{A}$.

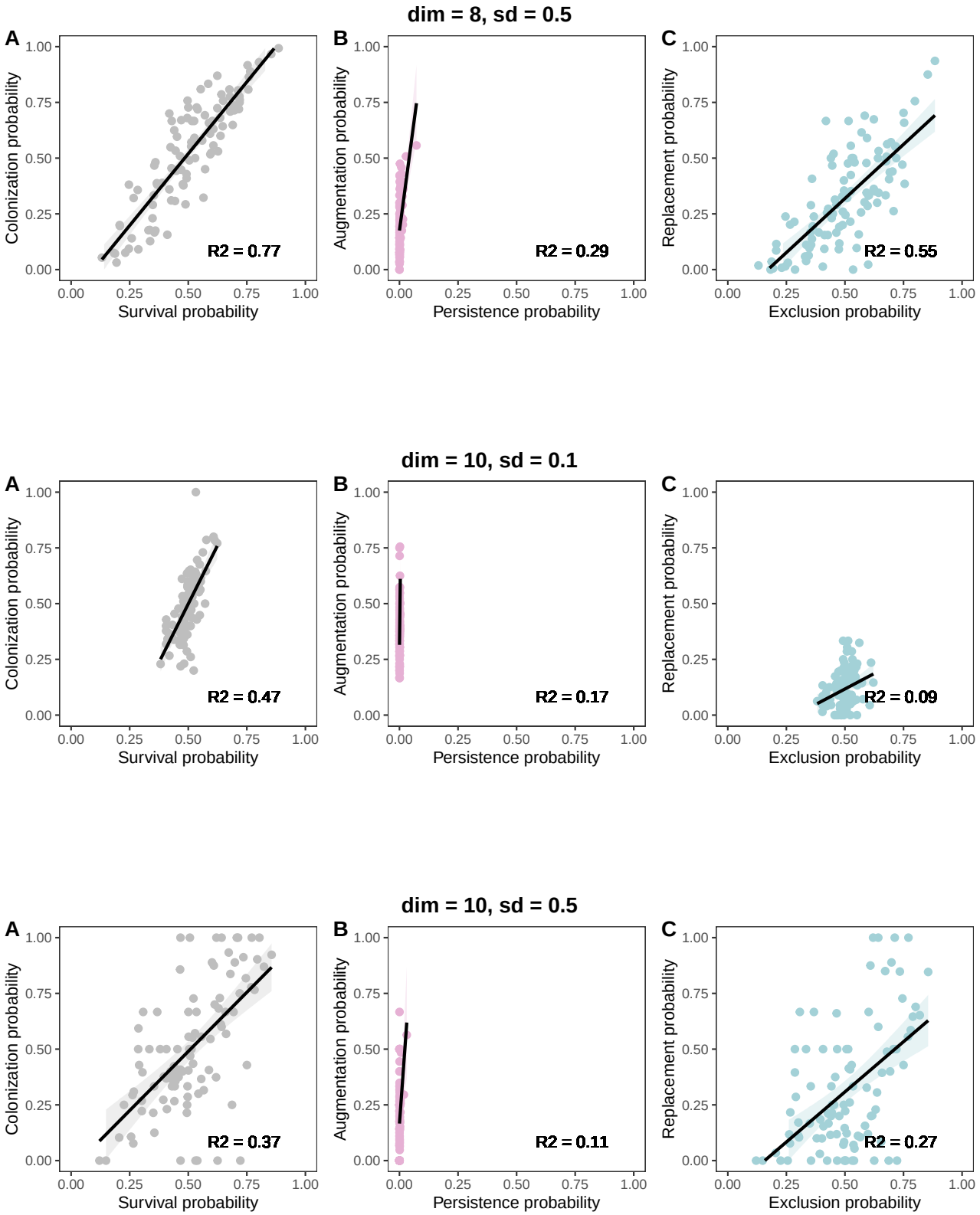
S2 Robustness of results

In this section, we reproduce Fig. 3 with different numbers of species (dim) and different standard deviations of the interactions (sd). The figures illustrate the numerical results of three pairs of analogous probabilistic outcomes using 100 randomly generated communities. Specifically, each point corresponds to a community with the number of species dim $\in \{3, 4, 5, 8, 10\}$. The corresponding interaction matrices $\mathbf{A}$ are randomly generated using a normal distribution with mean $\mu = 0$ and standard deviations sd $\in \{0.1, 0.5, 1, 1.5\}$ (diagonal entries $a_{ii} = -1$). Specific parameter combinations are shown in the caption of each figure. Probabilities (y-axes) are calculated numerically using an ensemble of 10^4 effective growth rates $\boldsymbol{\theta}$ sampled uniformly inside the feasibility domain of residents in isolation $\boldsymbol{\theta}^*(\bar{i})$ (Eqs. (13)-(15)) and then simulating the gLV dynamics (see main text for details). Then, the probabilities can be approximated by the frequencies that certain species compositions persist together over a finite period of time and threshold for species extinction (Deng et al., 2022). Consistent with the main text, we use the R^2 to measure the translatability of analogous outcomes between simultaneous and sequential assembly outcomes. With different combinations of parameters, the numerical results are qualitatively the same as Fig. 3.




dim = 3, sd = 1.5
A
Colonization probability
Survival probability
R2 = 0.91
B
Augmentation probability
Persistence probability
R2 = 0.61
C
Replacement probability
Exclusion probability
R2 = 0.79
dim = 4, sd = 0.5
A
Colonization probability
Survival probability
R2 = 0.94
B
Augmentation probability
Persistence probability
R2 = 0.58
C
Replacement probability
Exclusion probability
R2 = 0.67
dim = 4, sd = 1
A
Colonization probability
Survival probability
R2 = 0.87
B
Augmentation probability
Persistence probability
R2 = 0.45
C
Replacement probability
Exclusion probability
R2 = 0.7
dim = 4, sd = 1.5
A
Colonization probability
Survival probability
R2 = 0.76
B
Augmentation probability
Persistence probability
R2 = 0.47
C
Replacement probability
Exclusion probability
R2 = 0.58
0.00
0.25
0.50
0.75
1.00

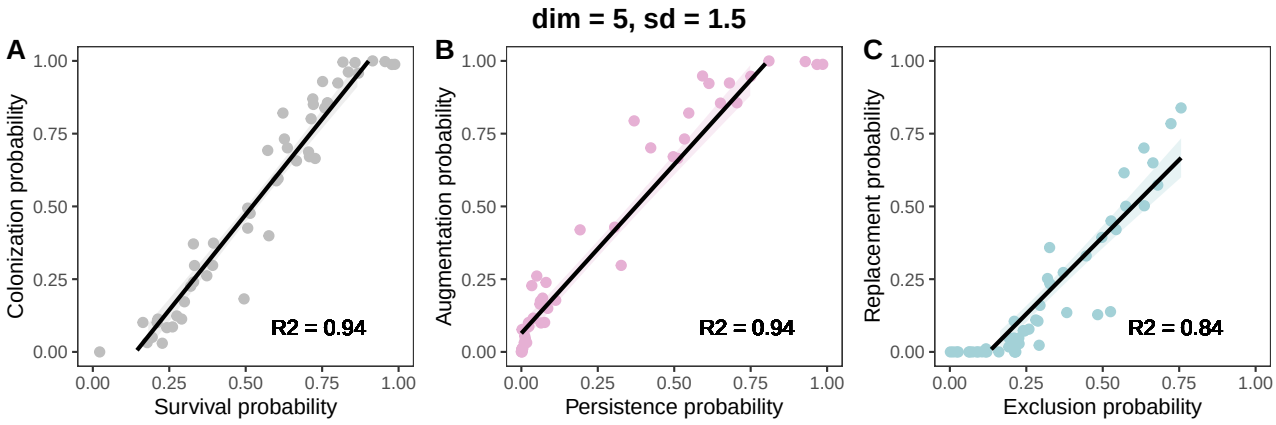

dim = 5, sd = 0.5
A
Colonization probability
Survival probability
R2 = 0.88
B
Augmentation probability
Persistence probability
R2 = 0.42
C
Replacement probability
Exclusion probability
R2 = 0.45
dim = 5, sd = 1
A
Colonization probability
Survival probability
R2 = 0.8
B
Augmentation probability
Persistence probability
R2 = 0.55
C
Replacement probability
Exclusion probability
R2 = 0.56
dim = 5, sd = 1.5
A
Colonization probability
Survival probability
R2 = 0.83
B
Augmentation probability
Persistence probability
R2 = 0.42
C
Replacement probability
Exclusion probability
R2 = 0.68
dim = 8, sd = 0.1
A
Colonization probability
Survival probability
R2 = 0.84
B
Augmentation probability
Persistence probability
R2 = 0.26
C
Replacement probability
Exclusion probability
R2 = 0.1
0.00
0.25
0.50
0.75
1.00



Supplementary Figure S1: **Reproducing Fig. 3 with different numbers of species (dim) and different standard deviations of the interactions (sd).** The figures illustrate the numerical results using gLV simulations of three pairs of analogous probabilistic outcomes using 100 randomly generated communities. The x-axes and y-axes correspond to the outcomes under simultaneous and sequential assembly processes, respectively. Each point corresponds to a community with the number of species dim $\in \{3, 4, 5, 8, 10\}$. The interaction matrices $\mathbf{A}$ are randomly generated using a normal distribution with mean $\mu = 0$ and standard deviations sd $\in \{0.1, 0.5, 1, 1.5\}$ (diagonal entries $a_{ii} = -1$). Specific parameter combinations are shown in the caption of each figure. In summary, the numerical results with different combinations of parameters are qualitatively the same as Fig. 3.

S3 Locally stable communities

Considering three-species locally stable communities, we sampled 50 interaction matrices $\mathbf{A}$ from a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 2$ (diagonal entries $a_i i = -1$). Then, for each locally stable community, we sampled 5×10^3 vectors of effective growth rates $\boldsymbol{\theta}$ uniformly on a 3-dimensional closed unit sphere corresponding to the unknown and changing environments. Also, we set the initial abundance of the invader (focal) species i as 10^{-3}, and draw that of the other two species from a uniform distribution with limits min $= 0$ and max $= 1$. Next, to check the surviving species in each community, we simulate the Lotka-Volterra model (Eq. (1)) with total time 200 and step size 10^{-2}. The extinction threshold is 10^{-6}. In particular, for sequential assembly, we perform separate simulations for the resident communities with interaction sub-matrices $\mathbf{A}(\bar{i}, \bar{i})$, truncated effective growth rates $\boldsymbol{\theta}(\bar{i})$, and corresponding initial conditions to guarantee that they can coexist in isolation (see the definition of the feasibility domain of residents in isolation $\boldsymbol{\theta}^*(\bar{i})$ in Eq. (9) and the geometric illustration in Fig. 1A). The simulations are conducted by the Runge-Kutta method. Hence, the probability $F(\mathcal{C}, \mathcal{S})$ in simultaneous assembly (Eq. (8)) and the probability $I_i(\mathcal{C}, \mathcal{S})$ in sequential assembly (Eq. (9)) of a composition $\mathcal{C}$ within a multispecies community $\mathcal{S}$ can be approximated by the frequency with which all species in the composition $\mathcal{C}$ coexist together. Consistent with the main text, we use the R^2 to measure the translatability of analogous outcomes between simultaneous and sequential assembly processes. Fig. S2 illustrates the numerical results of three pairs of analogous probabilistic outcomes. Specifically, Panel A shows the pair of survival and colonization probabilities ($R^2 = 0.94$). Panel B shows the pair of persistence and augmentation probabilities ($R^2 = 0.94$). Panel C shows the pair of exclusion and replacement probabilities ($R^2 = 0.84$).



Supplementary Figure S2: **Relationship of analogous outcomes between simultaneous and sequential assembly processes in locally stable communities.** The figure illustrates the numerical results of three pairs of analogous probabilistic outcomes using 50 locally stable three-species communities that are randomly generated. The x-axes and y-axes correspond to the outcomes under simultaneous and sequential assembly processes, respectively. Each point corresponds to a three-species community, whose interaction matrix $\mathbf{A}$ is randomly generated using a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 2$ (diagonal entries $a_{ii} = -1$). All probabilities are calculated numerically using an ensemble of 5×10^3 effective growth rates $\boldsymbol{\theta}$ sampled uniformly inside the parameter space (i.e., a 3-dimensional closed unit sphere for simultaneous assembly and the feasibility domain of resident in isolation for sequential assembly). The results are qualitatively similar to the globally stable cases in Fig. 3 and Supplementary Section S2.