

# Interaction Strengths Affect Whether Ecological Networks Promote the Initiation of Egalitarian Major Transitions

Sydney Leither<sup>1,2,3</sup>, Max Foreback<sup>1,2,3</sup>, David A. Baum<sup>4,5</sup> and Emily Dolson<sup>1,2,3</sup>

<sup>1</sup>Computer Science and Engineering, Michigan State University, East Lansing, MI

<sup>2</sup>Ecology, Evolutionary Biology, and Behavior, Michigan State University, East Lansing, MI

<sup>3</sup>BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI

<sup>4</sup>Botany, University of Wisconsin, Madison, WI

<sup>5</sup>Wisconsin Institute for Discovery, University of Wisconsin, Madison, WI

leithers@msu.edu

## Abstract

Identifying conditions that promote egalitarian major transitions, where unlike replicating units unite to form a higher-level unit, is an open problem with far-reaching implications. We propose that egalitarian major transitions can only begin in ecological communities that are conducive to them. To formalize this idea, we introduce the concept of “transitionability”, which describes the extent to which a community is poised to undergo an egalitarian major transition. We hypothesize that transitionability is a property of ecological interaction networks, which represent the set of pairwise interactions among members of a community. Using a digital artificial ecology that simulates interactions between species based on a static interaction network, we test the transitionability of interaction networks created by a range of graph-generation techniques, as well as some real-world ecological networks. To measure the extent to which a community is moving towards a major transition, we quantify the increase in community-level fitness relative to individual-level fitness across five different fitness proxies. We find that some network generation protocols produce more transitionable networks than others. In particular, interaction strengths (i.e. edge weights) have a substantial impact on transitionability, despite receiving low attention in the literature.

## Introduction

It has been theorized that ecological dynamics were an emergent property of pre-biotic chemistry. Auto-catalytic (i.e. effectively self-replicating) cycles of chemical reactions interacted with each other in ways analogous to predation, mutualism, and competition (Peng et al., 2020). However, these auto-catalytic cycles would not have been capable of mutation, and so would not have experienced evolution by natural selection. Indeed, one commonly-proposed dividing line for distinguishing between this soup of chemicals and the origin of life is the beginning of Darwinian evolution (Kauffman, 1993). From this perspective, the origin of life can be seen as an egalitarian major evolutionary transition in individuality in which a set of autocatalytic cycles came together to form a self-replicating individual (Queller and Strassmann, 2016). Populations of these self-replicating individuals would be

capable of evolution by natural selection because they would experience mutations (gain or loss of an autocatalytic cycle), differential reproductive success (via varying ability to spread), and heredity (due to their constituent auto-catalytic cycles self-replicating and co-dispersing).

Major evolutionary transitions in individuality happen when the nature of what it means to be an individual changes (Smith and Szathmary, 1997). Often, major transitions involve a higher-level evolutionary unit capable of inheritance emerging from the unification of independent lower-level units. Major transitions can be either egalitarian, where lower-level units are unrelated, or fraternal, where lower-level units are related (Queller, 1997). The origin of life via the emergence of adapting communities of autocatalytic cycles would be an egalitarian major transition (Queller and Strassmann, 2016). Understanding how these transitions occur is important for many open questions across evolutionary biology and artificial life, including open-ended evolution (Taylor et al., 2016; Banzhaf et al., 2016).

The idea that the emergence of collectives of self-replicating molecules in compartments can be thought of as major transitions dates back to Maynard Smith’s foundational book (Smith and Szathmary, 1997). Nevertheless, the origin of life differs from other egalitarian major transitions in an important way: co-evolution between constituent species, an important dynamic in other egalitarian major transitions, is impossible in the origin of life. We propose that this difference is actually useful for research purposes; the feedback loops inherent in egalitarian major transitions with co-evolution can be challenging to untangle. Thinking about the origin of life makes it clear that the factors affecting whether an egalitarian major transition occurs can be split into two categories: 1) those not involving co-evolution between lower-level species, and 2) those involving co-evolution between lower-level species. Factors in category one should act on all egalitarian major transitions. Here, we explore how they operate in the absence of co-evolution, an important first step towards eventually understanding how they operate in the presence of co-evolution.

What factors make a given chemistry more or less likely to lead to the origin of life? We suggest that these factors fall into two categories: 1) external (abiotic) factors such as diffusion, disturbance, and seeding of new chemicals, and 2) properties of the ecological interactions within the chemistry. For an exploration of the former, see Foreback et al. (2023a). Here, we focus on the latter. We hypothesize that answering this question in the context of chemical cycles is actually the same problem as identifying ecological communities that are poised to undergo egalitarian major transitions (i.e. those which could undergo a major transition purely on the basis of changes to their abiotic environment), a property which we term "transitionability." To facilitate the connection of this work to general ecological theory, we describe ecological communities using the abstraction of ecological networks, also called interaction matrices. These networks describe how units within the community interact with or affect each other. The structure of the network captures the extent to which the units' community possesses ecological features such as trophic levels, instances of facilitation, and instances of antagonism.

We use the framework of multi-level selection to conceptualize the extent to which selection is acting on the two different levels of organization (Bonner, 1998). In this context, identifying shifts in the level of organization on which selection is acting most strongly can be a proxy for identifying major transitions (Ratcliff et al., 2015). Egalitarian major transitions, then, can be qualitatively identified by observing fitness increasing primarily at the level of the community. Thus, interaction networks that are conducive to undergoing egalitarian major evolutionary transitions are those for which community-level fitness increases are possible.

Major transitions and ecological networks are both difficult to observe in nature, making it challenging to ascertain how properties of ecological networks affect major transitions using biological data. Real-world ecological networks are usually created through field observation, which can only capture a limited snapshot of the complete ecological reality. Consequently, this problem lends itself to an artificial life approach. We use an artificial ecology to explore what kinds of ecological networks are more likely to lead to a major transition. By studying both real-world interaction matrices and five schemes for stochastically generating matrices, we find evidence that both interaction network structure and interaction strengths affect transitionability. By identifying generation schemes that promote transitionability, we lay the groundwork for follow-up research to determine what exact properties of the networks are fostering the conditions for major transitions.

## Background

### Interaction Matrices

Interaction matrices, also known as ecological networks, are networks that define interactions between species in a com-

munity (Landi et al., 2018; Delmas et al., 2019). Interaction matrices are adjacency matrices representing graphs whose nodes are species and edges are interactions between species. Interactions can be unweighted (binary presence/absence of interaction) or weighted (strength of interaction). Ecologists define different types of interaction matrices by the set of interactions represented in them. Some common kinds of interaction matrices include food webs (D'Alelio et al., 2016), diet composition (Bulman et al., 2001), or plant-pollinator networks (Kaiser-Bunbury et al., 2009). Food webs typically have directed trophic interactions where the interaction between species  $i$  and species  $j$  is positive while the interaction between species  $j$  and species  $i$  is negative (Pomeranz et al., 2019). Diet composition networks may have weighted interactions representing the percent of a predator's diet a prey species makes up (Bulman et al., 2001), while plant-pollinator networks may have mutualistic interactions that track the number of visits by a pollinator to a plant (Kaiser-Bunbury et al., 2009). Interaction matrices can also include multiple kinds of interactions, although in practice these are less frequently measured. In this paper, interaction networks are directed and weighted. Interactions represent the non-symmetric effect of species  $i$  on species  $j$ , which can be beneficial or detrimental as indicated by the sign of the weight.

There are many open questions regarding ecological networks, such as how the number of interactions vary with the number of species (Landi et al., 2018). One of the most studied properties of ecological networks is stability. Prior work has suggested that network structure/topology (Lurgi et al., 2016) and interaction strengths (Tang et al., 2014) have an impact on the stability of a community. It is unclear whether we should expect stable communities to promote or inhibit egalitarian major transitions. On one hand, stability may be a prerequisite for fostering an environment capable of egalitarian major transitions, as an unstable community likely could not reliably self-replicate if functioning as an individual. On the other hand, too much stability may make the community resistant to seeding events, and thus unable to undergo the community-level equivalent of mutation. Such a community would have low community-level evolvability.

Real-world interaction matrices are measured through field observation and literature review. As such, they are expensive to create and represent an incomplete ecological reality based on limited observation. Models exist to generate networks that capture properties thought to be found in real-world networks, such as Barabási and Albert (1999), Watts and Strogatz (1998), and Klemm and Eguiluz (2002), which generate networks with scale-free behavior, small-world behavior, or both, respectively. However, the extent to which these behaviors are present in real-world ecological networks is controversial; while some studies have observed these properties in nature (Montoya and Solé, 2002), others have not (Dunne et al. (2002)). Furthermore, these net-

work generation models only generate matrix structure, so weights must be set manually or by using another method to generate weighted ecological networks. A study by Tang et al. (2014) finds that matching the distribution and correlation of interaction strengths to real-world ecological networks allows random network structures to preserve stability. This observation implies that interaction strengths may play an important role in producing specific ecological dynamics. However, relatively few studies have focused on interaction strengths, suggesting that they may be an understudied aspect of ecological networks. Network topology is much more commonly studied, likely due to the difficulty of measuring real-world interaction strengths.

### Role of Ecology in Major Transitions

It has been widely speculated that the presence of ecological dynamics is critical for open-ended evolution in general (Soros and Stanley, 2014; Taylor et al., 2016; Dolson et al., 2019). Moreover, hallmarks of open-ended evolution such as egalitarian major evolutionary transitions trivially require preexisting ecological dynamics, as such transitions require the presence of multiple distinct species (Dolson et al., 2019). However, while ecology in general is thought to be important in these contexts, less is known about the specific ecological dynamics necessary to bring these phenomena about. Here we seek to identify such dynamics through our examination of interaction networks.

The origin of life is a useful case study of a major egalitarian transition, because it allows us to make the helpful simplifying assumption that the interaction network itself does not evolve. Historically, efforts to study the impact of interaction network structure on adaptation have been stymied by the fact that these two factors are locked in a feedback loop; the interaction network influences adaptation and adaptation in turn changes the interaction network. If we conceive of the origin of life as being the origin of the first adapter, however, we can, for now, set aside the possibility of interaction networks changing in response to evolution.

## Methods

### Artificial Ecology

**Architecture** The code used in this paper is open source (Leither et al., 2023). We employ an artificial ecology to conduct *in silico* experiments on major evolutionary transitions. The artificial ecology uses a generalized Lotka-Volterra model to simulate growth of species within a meta-community (Leibold and Chase, 2017). The meta-community consists of a grid of local sites which contain a combination of species that interact with each other. Species are represented by continuous values indicating the sizes of their populations at each site. The growth of each species is directly proportional to the abundance of species beneficial to it and the scarcity of species detrimental to it. The meta-community goes through updates which modify the combi-

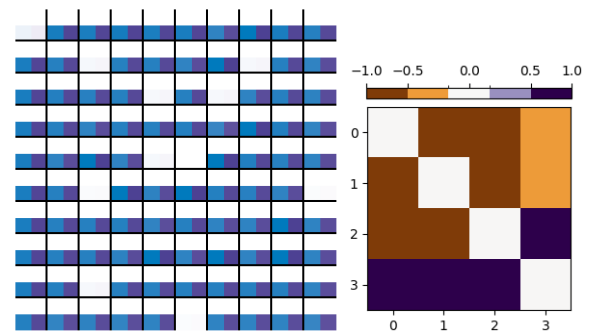


Figure 1: An example meta-community (left) and corresponding interaction matrix (right) with four species. The meta-community is split into local sites, with colored squares indicating the species at each site. The community composition 0011 is dominating, as most local sites have only species 2 and 3 present. This community composition is dominating because species 2 and 3 interact positively, allowing them to grow fast.

nation of species (i.e. the community composition) present in each local site. Community compositions are represented by binary strings indicating the presence or absence of each species. For an example of a meta-community, see Figure 1.

Species are differentiated by their unique interactions with each other as defined in a weighted and directed community interaction matrix. The matrix is non-symmetric, meaning that the effect of species  $i$  on species  $j$  can be different than the effect of species  $j$  on species  $i$ . Species can interact with themselves (diagonal matrix values) which defines their intrinsic growth rate. Interactions are continuous in the range  $[-1, 1]$ . As an interaction value  $a_{ij}$  approaches one, species  $i$  has an increasingly beneficial effect on species  $j$ , while a value approaching negative one has an increasingly detrimental effect. A value of zero indicates that species  $i$  does not interact with or affect species  $j$ .

Abiotic parameters, which define how species interact with the environment at each time step, also affect how the meta-community updates and have an impact on ecological dynamics. The artificial ecology has three continuous-valued abiotic parameters in the range  $[0, 1]$ : diffusion, seeding, and clear rate. Diffusion defines what proportion of the species present in the local site will spread into adjacent local sites on each time step, and is analogous to wind- or water-assisted dispersal in nature. Seeding controls the probability that new species are introduced into a local site; it acts separately on each species and if activated will increase the count of that species by one. An analogous process in nature is immigration from an outside population (often referred to in ecology as the “regional species pool”). From the perspective of community-level adaptation, seeding is analogous to mutation. Clear rate is the probability that a local site will have all its species counts reset to

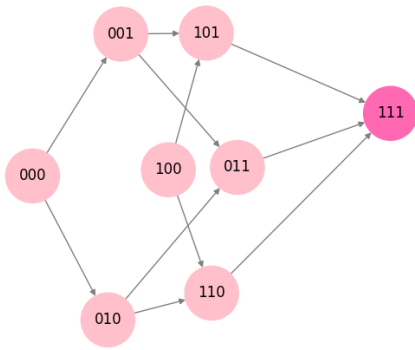


Figure 2: A simple example community assembly graph with three species. Community compositions are represented as bit strings where ones indicate species presence and zeros indicate species absence. For example, node 110 represents a community composition of species 0 and species 1. The edge from 110 to 111 indicates that species 2 can invade the community composition 110 and the new community composition 111 can exist stably.

zero on a given time step, which is analogous to disturbance from seasonality or weather events. Clear offers a way for a species to be removed from a local site without being driven to extinction by the other species in the local site. At each time step, the artificial ecology updates by growing the species via their interactions, clearing, diffusing, and then seeding. Other tunable parameters in the artificial ecology are meta-community size, number of time steps, and the maximum population size for each species in each local site, which we keep constant throughout the experiments in this paper at  $10 \times 10$ , 1000, and 10000 respectively.

**Measures of Adaptation** Earlier, we noted that major transitions can be identified by observing an increase in community-level fitness relative to the concurrent increase in individual-level fitness. Putting this definition into practice requires identifying a way to measure fitness. We identified five proxies for community-level fitness and analyzed each one. Since there are no individual members of any species in the artificial ecology, there cannot be mutation at the species level and any adaptation must occur at the community level. The five measures of adaptation are calculated from five fitness proxies: biomass, growth rate, heredity, invasion ability, and resilience. Biomass is the sum of each species' population size within each local site. It is a fitness proxy because, in the presence of diffusion, communities with high biomass will spread more of themselves to neighboring sites in the meta-community. Growth rate is how quickly a local site's biomass will grow given the current state of the community. It is perhaps the most direct fitness proxy, as it is essentially a measurement of how fast the community can copy itself. Heredity is calculated by placing the

species at each local site into an empty meta-community free of competition. Heredity is then measured as the similarity between the originally seeded community and its descendants that spread across the empty meta-community, where similarity is the Euclidean distance between the two communities. Heredity is not a direct fitness proxy, but some level of heritability is needed for true self-replication. Invasion ability is the speed at which a community can traverse empty local sites while maintaining its composition. Invasion ability is another fairly direct metric of the community's ability to replicate itself. Lastly, resilience is how resistant a community is to an invasion (addition of a new species) that would cause it to shift to a different stable community composition, where a stable community composition is one in which the present species could coexist permanently in the absence of perturbations. Like heredity, resilience is not a direct fitness measure, but is likely a necessary feature for continued existence in our system.

Since major transitions are generally understood to be rare, there is a somewhat high burden of proof associated with identifying them. Consequently, we want to err on the side of under-counting the number of major transitions rather than over-counting them. Thus, our null hypothesis is that no adaptive dynamics are present. If there are no community-level adaptive dynamics, then the only dynamics that should be present are purely ecological community assembly processes. To identify community-level adaptive dynamics (as we would expect to see in a major transition), we determine whether each of the fitness proxies are increasing beyond the extent expected under purely ecological dynamics. We create two models: one model predicting the expected dynamics if the changes in community compositions are governed by purely ecological forces and the other model predicting the expected dynamics if the changes are governed by purely community-level adaptive dynamics. Using a maximum likelihood approach, we compare the results of each run of our artificial ecology to these models and identify which model is more likely to have produced the observed results. The models are built as directed graphs with one (the community assembly graph) describing ecological dynamics and five (the fitness landscapes) describing adaptive dynamics under each of the fitness proxies. Ultimately, this process yields five measures of adaptation calculated based on the discrepancy between the community assembly graph prediction and each of the five fitness landscape predictions. A high probability of adaptation for any given score is indicative of adaptive dynamics working to maximize that fitness proxy, even if the other adaptation scores have low probabilities.

A community assembly graph contains nodes representing all stable communities that can theoretically exist and the possible transitions between them (Hang-Kwang and Pimm, 1993; Serván and Allesina, 2021). Nodes in our assembly graph represent a single community composition out of the

$2^n$  possible community compositions, where  $n$  is the number of species. Only stable community compositions are included in the graph. Edges between nodes represent possible invasions that could cause one community composition to transition to another. If the meta-community is being governed by purely ecological dynamics, it should behave like a random walk on the community assembly graph and eventually be populated only by community compositions that correspond to sink nodes (assuming the graph has sink nodes). Sink nodes are those which have an out degree of zero, meaning that they cannot be stably invaded. In some cases, the assembly graph has no sink nodes because there are no stable communities that are incapable of being invaded. In this case, graphs usually contain some basin of attraction that we would expect a random walk to ultimately conclude in. For example, a subset of nodes might contain only out-transitions amongst themselves. The five fitness landscape graphs are defined similarly to the community assembly graph, except they have an additional rule: a transition cannot occur from a node of higher community-level fitness to a node of lower community-level fitness. An example community assembly graph is shown in Figure 2.

To determine whether our results are best explained by ecological or adaptive dynamics, we determine whether stable community compositions present at the end of a run of the artificial ecology are more likely to be reached by a random walk on each of the fitness graphs or by a random walk on the assembly graph. We quantify the probability of a random walk ending up on a given node by calculating the PageRank (Page et al., 1999) of the nodes in each graph. The likelihood of each individual stable community composition being present under ecological or adaptive dynamics is given by the corresponding node's PageRank in the corresponding graph. The likelihood that observed results are primarily due to ecological dynamics,  $L(\text{ecological})$ , is the product of the community assembly graph's PageRanks for each stable community composition present at the end of a run. Similarly, the likelihood that the results are primarily due to adaptive dynamics according to each measure of adaptation,  $L(\text{adaptive})$ , is the product of the corresponding fitness graph's PageRanks for each stable community composition present. Formally, the adaptation score for a given fitness measure is the likelihood ratio:

$$\frac{\prod_{i=0}^{i=n} (\text{PageRank}_F(\text{community}_i))}{\prod_{i=0}^{i=n} (\text{PageRank}_A(\text{community}_i))}$$

where  $n$  is the number of stable community compositions,  $\text{PageRank}_A$  is the community composition's PageRank score on the assembly graph, and  $\text{PageRank}_F$  is the community composition's PageRank score on the given fitness landscape graph. Likelihood ratios greater than one suggest that our results are more likely under purely adaptive dynamics than under purely ecological dynamics. We describe the interaction matrices of these runs as "transitionable", as they

can produce communities that selection is capable of acting on under right combination of abiotic parameters, and so are likely able to foster an egalitarian major transition similar to the origin of life. We hypothesize that they also foster other egalitarian major transitions, but further research is necessary to test this hypothesis.

## Matrix Generation Schemes

In order to identify interaction matrices that lead to greater rates of transitionability, we employ three matrix generation schemes and also test real-world networks. Testing different matrix generation schemes allows us to see how different matrix properties, both in terms of topology and in terms of interaction strength, affect adaptive dynamics in our artificial ecology. It also allows us to explore different matrices systematically, which would not be tractable through enumeration due to the large parameter space of the matrices. Interaction matrices contain  $N$  species consist of  $N^2$  elements, each with a continuous value in the range  $[-1, 1]$ . We used  $N = 9$  for our generated matrices, although some of the real-world matrices were larger. Furthermore, the three abiotic parameters diffusion, seeding, and clear rate also affect the presence of adaptive dynamics, so each matrix generation scheme must be tested with differing abiotic parameters. To do so we rigorously sample the parameter space (i.e. the matrix generation parameters in combination with the abiotic parameters) of each matrix generation scheme using latin hypercube sampling (McKay et al., 2000). This approach tests a representative sample of the parameter space without conducting an expensive exhaustive search. Previously, we utilized evolutionary computation to explore the parameter space of one matrix generation scheme (plus abiotic parameters) to find local regions of adaptive dynamics (Foreback et al., 2023b). We run the artificial ecology on 50,000 samples from the parameter space over ten replicates, each with a different random seed, ensuring there are ten unique matrices from each sampled part of the landscape.

**Real World** We investigate whether real-world interaction matrices exhibit transitionability within the artificial ecology when combined with the right abiotic parameters by utilizing the Mangal database (Poisot et al., 2016). The Mangal database serializes ecological networks from across the world into a consistent format for accessible and efficient analysis. We chose to use all dietary composition matrices with less than twenty species<sup>1</sup>, a total of four unique networks. Mangal's dietary matrices by default have interactions directed from predator to prey species. The weight of the interaction indicates the percent of the predator's diet that the prey comprises. As this direction is the opposite of what our system expects, we switch the direction of the

<sup>1</sup>This limitation was necessary due to the exponential scaling of the size of the community assembly graph.

Matrix Scheme	Biomass	Growth Rate	Heredity	Invasion Ability	Resilience
Mangal	0.22	0.06	0.22	0.05	0.08
Klemm-Eguíluz + LFR	0.27	0.25	0.02	0.03	0.20
Klemm-Eguíluz + Random	0.22	0.20	0.05	0.02	0.22
Motifs	0.19	0.17	0.03	0.02	0.22
Random	0.20	0.20	0.03	0.03	0.18

Table 1: Proportion of transitionable runs for each generation scheme.

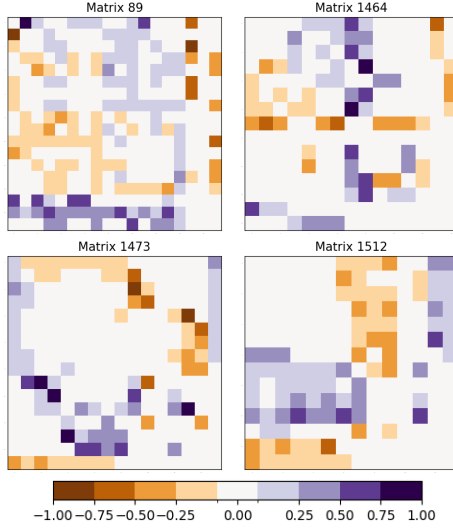


Figure 3: The dietary matrices analyzed. Negative interactions are randomized. Labels show Mangal network IDs.

interaction so it goes from prey to predator, indicating how much the prey species positively affects the predator species. We add negative interactions from the predator species to the prey species for predation and herbivory interaction types. Negative interactions are not added for the other two interaction types present in the dietary matrices: unspecified and detritivore. Because we do not know the strength of the predator’s effect on the prey, each dietary matrix is replicated five times with different random weights for these interactions drawn from a normal distribution centered on the original prey-predator interaction strength with a standard deviation of 0.1. One of the five replicated matrices for each of the four dietary composition networks are shown in Figure 3, labeled with their Mangal network ID. By replicating matrices with different negative interaction strengths we can study the effect of the strength of the predator’s interaction on the prey. Each of the replicated matrices are run through the artificial ecology with combinations of the three abiotic parameters sampled from a representative set of values [0.001, 0.333, 0.666, 0.999]. As we have four interaction matrices each replicated five times on  $3^4$  combinations of abiotic parameters, this amounts to a total of

$4 \times 5 \times 64 = 1,280$  runs of the artificial ecology.

**Klemm-Eguíluz** We generate interaction matrix topologies using the Klemm-Eguíluz network generation scheme (Klemm and Eguíluz, 2002). Weights are set using two different methods: Lancichinetti–Fortunato–Radicchi (LFR) (Lancichinetti and Fortunato, 2009) and random. Klemm-Eguíluz matrices are scale-free networks with small world behavior as they hold the following three properties commonly found in networks representing real-world systems: a small shortest average path length, a large clustering coefficient, and a scale-free degree distribution (Klemm and Eguíluz, 2002). Real-world ecological networks may or may not hold these properties, as discussed in the background section, but we leave it to future work to explore how deviating from these properties affects the artificial ecology. Klemm-Eguíluz matrices are generated by initially creating a clique of a given size  $m$ , called the active nodes, and sequentially connecting the remaining nodes in a preferential attachment process. Each remaining node is connected to either one of the active nodes or a non-active node with probability  $\mu$ . Once connected, the node becomes active and another active node is deactivated, where active nodes with lower degrees are more likely to be deactivated.

Weights are assigned to Klemm-Eguíluz matrices using LFR’s directed graph weight-setting algorithm, originally designed to generate benchmark graphs for testing community detection algorithms (Lancichinetti and Fortunato, 2009). The weight-setting algorithm takes in an existing matrix structure, community memberships of each node  $i$ , a parameter  $\beta$  controlling the sum of the edge weights of each node, and a parameter  $\mu_w$  controlling the sum of edge weights between nodes in the same communities. Each node has a power-law relation  $s_i = k_i^\beta$  between its sum of edge weights  $s_i$  and its sum of in and out degrees  $k_i$ . This property is thought to be desirable, as a power-law relation between  $s_i$  and  $k_i$  has been observed in networks representing real-world systems (Barrat et al., 2004). The sum of edge weights between nodes within the same communities is set to  $s_i^{(in)} = (1 - \mu_w)s_i$  and the sum of edge weights between nodes of differing communities is set to  $s_i^{(out)} = \mu_w s_i$ . To apply LFR weight-setting to Klemm-Eguíluz matrices, we consider the original active nodes a community and the remaining nodes another community. Min-max normalization

is applied to the final edge weights to keep them in the range  $[0, 1]$ . We set a proportion  $p_{in}$  of the weights between nodes in the same community as negative and a different proportion  $p_{out}$  of the weights between nodes in differing communities as negative. There are six matrix parameters for this generation scheme:  $m$ ,  $\mu$ ,  $\beta$ ,  $\mu_w$ ,  $p_{in}$ , and  $p_{out}$ .

As an additional matrix generation scheme, Klemm-Eguíluz matrix weights are set randomly in order to explore how beneficial LFR weights are. Random weights are drawn from a uniform distribution in range  $[-1, 1]$ . This scheme has two matrix parameters:  $m$  and  $\mu$ .

**Motifs** We generate motif-based interaction matrices based on the motifs outlined in Losapio et al. (2021). Certain three-species motifs have been found to appear in ecological networks with differing frequency than expected by chance (Delmas et al., 2019), indicating that motifs may be an ecologically significant aspect of real-world interaction matrices. Losapio et al. (2021) explored 13 motifs found in alpine plant community ecological networks and found that three motifs representing competition between facilitated species or facilitation between inferior competitors were over-represented in the ecological networks. We create interaction matrices from the defined motifs by adding a motif to each three-species block in the interaction matrix, for a total of nine motifs in each nine species interaction matrix. Motifs are drawn with replacement from a pool including a motif consisting of all zeros. Motifs can be set with low, medium, or high interaction strengths; the strength of the interactions in the motifs is a random value chosen from the range  $[0, 0.33]$ ,  $[0.33, 0.66]$  or  $[0.66, 1]$  respectively, depending on the selected strength category. There are eighteen matrix parameters for the motif scheme: nine controlling which motif goes in each of the three species blocks and nine controlling the strength of each motif.

**Random** As a control, we also test random interaction matrices. Similar to the above schemes, we create 100,000 matrices with nine species, each with random abiotic parameters. The random matrices have a parameter that controls the probability that each possible edge has a nonzero weight drawn from a uniform distribution in range  $[-1, 1]$ .

## Results and Discussion

The proportion of runs that had transitionable matrices for each matrix generation scheme is shown in Figure 1. The Klemm-Eguíluz + LFR generation scheme had the highest proportion of transitionable matrices in four out of the five fitness measures, while the other generation schemes all performed similarly across fitness measures. The Klemm-Eguíluz + LFR matrices have the greatest control of interaction strengths over all other generation schemes, as they had power-law relation between the sum of edge weights and node degree while all other schemes had strengths set semi-

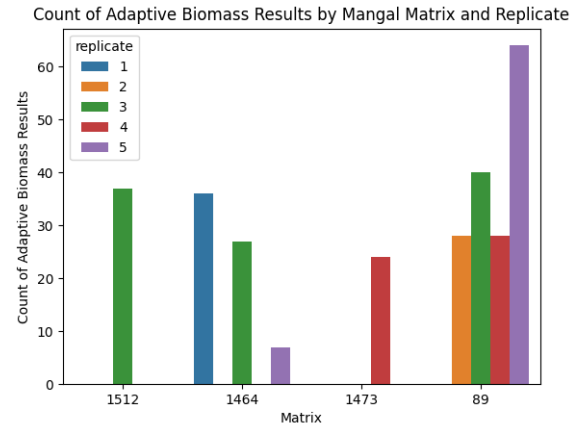


Figure 4: A barplot showing the count of likely biomass transitionable matrices for each Mangal matrix, colored by replicate. Each replicate had unique negative interactions, so the variation in results showing likely transitionable biomass dynamics within the same matrix structure implies that interaction strengths have a substantial effect.

randomly. Existing literature on interaction matrices focuses more on the structure of networks rather than weights, but Klemm-Eguíluz + LFR being able to find more transitionable matrices in most adaptation measurements than Klemm-Eguíluz + random implies that non-random interaction strengths may be important for finding transitionable matrices. Further evidence for the importance of interaction strengths can be seen by examining the Mangal results.

Each of the four Mangal matrices were replicated five times with random negative weights and each of those five were run through the artificial ecology with 64 different combinations of abiotic parameters. We find that the negative interaction strengths are an important aspect of whether a given matrix is transitionable or not. Even though each of the five replicates had the same underlying structure, the proportion of their runs that were transitionable varied based on the strengths of the negative weights (see Figure 4). We observed this dynamic for the other four adaptation measures as well. Whether or not a given replicate Mangal matrix is transitionable does not correlate with how often the negative weight  $a_{ij}$  is larger than the respective positive weight  $a_{ji}$ . Instead it appears to depend on the specific placement of the stronger negative weight; the matrix may be less likely to be transitionable if the negative interaction is preventing key species from coexisting. Furthermore, the fact that Mangal matrix 89 had a majority of its replicates transitionable while the other Mangal matrices did not suggests that there are properties of network structure important for transitionability that go beyond just exhibiting real-world structure. This idea is further supported by Mangal matrices not having the

highest proportion of transitionable matrices. Another interesting aspect of the Mangal results is that they had by far the highest proportion of transitionable heredity results and the lowest proportion of transitionable growth rate results compared to the other matrix generation schemes. Furthermore, all of the transitionable biomass results were also transitionable in heredity which did not happen in the other schemes.

The distribution of adaptation scores for biomass, heredity, and invasion ability are shown in Figure 5. While Klemm-Eguíluz + LFR matrices had the largest proportion of results that are more likely than not to be transitionable, the likelihood of those results actually being transitionable was lower than for the other generation schemes. On the other hand, while Mangal did not have the highest proportion of transitionable matrices, among those that were more likely than not to be transitionable it consistently had the highest likelihood of transitionability. According to a set of Kruskal-Wallis Tests (Kruskal and Wallis, 1952) ( $p < .05$  for all fitness metrics) with followup pairwise Wilcoxon rank-sum tests (Mann and Whitney, 1947) with a Bonferroni correction for multiple comparisons (Dunn, 1961), the medians of each score of each matrix scheme were significantly different from each other ( $p < 0.005$ , the Bonferroni-corrected significance level).

## Conclusion

We find that instances of transitionability are likely to occur in all adaptation measurements across all matrix schemes. The Klemm-Eguíluz + LFR matrices, which has the greatest control over the interaction strengths, had the highest proportion of transitionable results. Interestingly, the real-world matrices, although not substantially more likely to be transitionable than randomly-generated matrices, had the highest likelihoods of transitionability among those matrices that were transitionable. Note that these findings do not imply the communities these networks were measured in are undergoing major transitions, only that they might under the right conditions. Our results suggest that systematic weights and non-random structure in interaction matrices increase the probability that they foster community-level adaptation. For future work we plan to identify specific properties of networks that promote transitionability across generation schemes. We will analyze more generation schemes and real-world networks, including bipartite ones.

## Acknowledgements

We thank David E. Smith and the Michigan State University ECODE lab for feedback on the ideas in this paper. This material is based upon work supported by the National Science Foundation under Grant No. 2218818. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

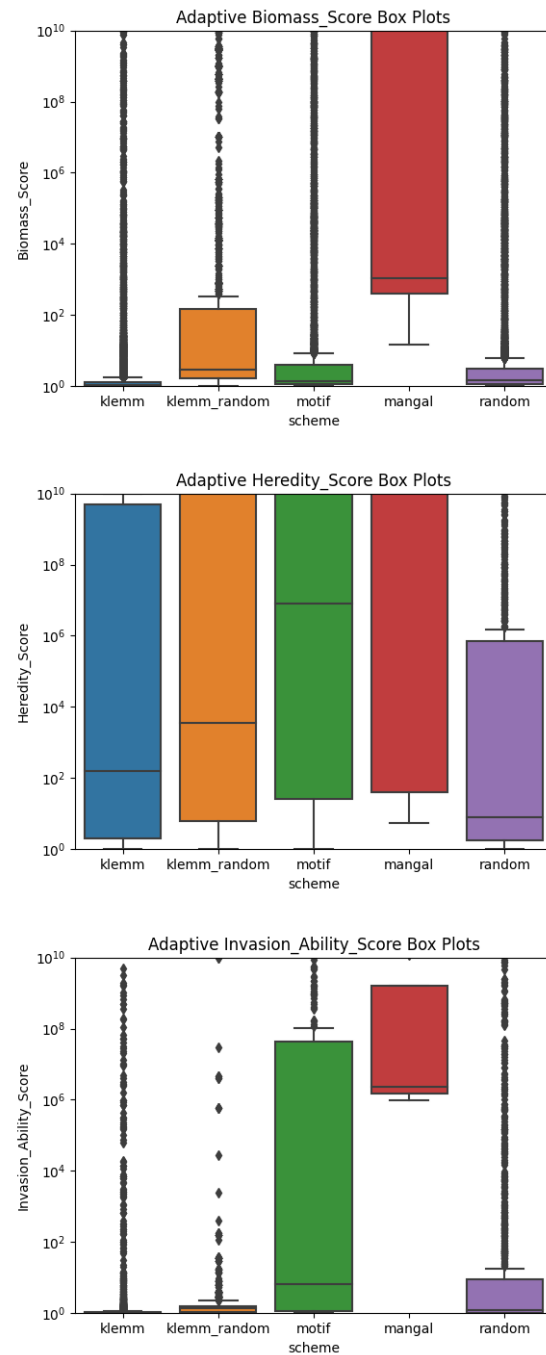


Figure 5: Boxplots of transitionable adaptation scores for biomass, heredity, and invasion ability (cut off at  $10^{10}$ ). Growth rate and resilience (not shown) are similar to biomass.



## References

- Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R., Foster, J. A., McMullin, B., De Melo, V. V., Miconi, T., Specator, L., Stepney, S., et al. (2016). Defining and simulating open-ended novelty: requirements, guidelines, and challenges. *Theory in Biosciences*, 135:131–161.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439):509–512.
- Barrat, A., Barthelemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the national academy of sciences*, 101(11):3747–3752.
- Bonner, J. T. (1998). The origins of multicellularity. *Integrative Biology: Issues, News, and Reviews: Published in Association with The Society for Integrative and Comparative Biology*, 1(1):27–36.
- Bulman, C., Althaus, F., He, X., Bax, N., and Williams, A. (2001). Diets and trophic guilds of demersal fishes of the south-eastern australian shelf. *Marine and Freshwater Research*, 52(4):537–548.
- Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Dalla Riva, G. V., Fortin, M.-J., Gravel, D., Guimarães Jr, P. R., Hembry, D. H., Newman, E. A., et al. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, 94(1):16–36.
- Dolson, E. L., Vostinar, A. E., Wiser, M. J., and Ofria, C. (2019). The modes toolbox: Measurements of open-ended dynamics in evolving systems. *Artificial life*, 25(1):50–73.
- Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American statistical association*, 56(293):52–64.
- Dunne, J. A., Williams, R. J., and Martinez, N. D. (2002). Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences*, 99(20):12917–12922.
- D'Alelio, D., Libralato, S., Wyatt, T., and Ribera d'Alcalà, M. (2016). Ecological-network models link diversity, structure and function in the plankton food-web. *Scientific Reports*, 6(1):21806.
- Foreback, M., Leither, S., Baum, D. A., and Dolson, E. (2023a). The role of abiotic parameters in the promotion of egalitarian major evolutionary transitions. In *ALIFE 2023: The 2023 Conference on Artificial Life*. MIT Press.
- Foreback, M., Leither, S., and Dolson, E. (2023b). Using evolutionary computation to find parameters that promote egalitarian major evolutionary transitions. In *Genetic and Evolutionary Computation Conference Companion (GECCO '23 Companion), July 15–19, 2023, Lisbon, Portugal*, New York, NY, USA. ACM.
- Hang-Kwang, L. and Pimm, S. L. (1993). The assembly of ecological communities: a minimalist approach. *Journal of Animal Ecology*, pages 749–765.
- Kaiser-Bunbury, C. N., Memmott, J., and Müller, C. B. (2009). Community structure of pollination webs of mauritian heathland habitats. *Perspectives in Plant Ecology, Evolution and Systematics*, 11(4):241–254.
- Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford University Press, USA.
- Klemm, K. and Eguiluz, V. M. (2002). Growing scale-free networks with small-world behavior. *Physical Review E*, 65(5):057102.
- Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):583–621.
- Lancichinetti, A. and Fortunato, S. (2009). Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E*, 80(1):016118.
- Landi, P., Minoarivelo, H. O., Brännström, Å., Hui, C., and Dieckmann, U. (2018). Complexity and stability of ecological networks: a review of the theory. *Population Ecology*, 60:319–345.
- Leibold, M. A. and Chase, J. M. (2017). Metacommunity ecology, volume 59. In *Metacommunity Ecology, Volume 59*. Princeton University Press.
- Leither, S., Foreback, M., and Dolson, E. (2023). Chemeco\_alife\_2023 github repository. <https://doi.org/10.5281/zenodo.7915917>.
- Losapio, G., Schöb, C., Staniczenko, P. P., Carrara, F., Palamara, G. M., De Moraes, C. M., Mescher, M. C., Brooker, R. W., Butterfield, B. J., Callaway, R. M., et al. (2021). Network motifs involving both competition and facilitation predict biodiversity in alpine plant communities. *Proceedings of the National Academy of Sciences*, 118(6):e2005759118.
- Lurgi, M., Montoya, D., and Montoya, J. M. (2016). The effects of space and diversity of interaction types on the stability of complex ecological networks. *Theoretical ecology*, 9:3–13.
- Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60.
- McKay, M. D., Beckman, R. J., and Conover, W. J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1):55–61.
- Montoya, J. M. and Solé, R. V. (2002). Small world patterns in food webs. *Journal of theoretical biology*, 214(3):405–412.
- Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. In *The Web Conference*.
- Peng, Z., Plum, A. M., Gagrani, P., and Baum, D. A. (2020). An ecological framework for the analysis of prebiotic chemical reaction networks. *Journal of theoretical biology*, 507:110451.
- Poisot, T., Baiser, B., Dunne, J. A., Kéfi, S., Massol, F., Mouquet, N., Romanuk, T. N., Stouffer, D. B., Wood, S. A., and Gravel, D. (2016). mangal-making ecological network analysis simple. *Ecography*, 39(4):384–390.
- Pomeranz, J. P., Thompson, R. M., Poisot, T., and Harding, J. S. (2019). Inferring predator-prey interactions in food webs. *Methods in Ecology and Evolution*, 10(3):356–367.

- Queller, D. C. (1997). Cooperators since life began.
- Queller, D. C. and Strassmann, J. E. (2016). Problems of multi-species organisms: endosymbionts to holobionts. *Biology & Philosophy*, 31:855–873.
- Ratcliff, W. C., Fankhauser, J. D., Rogers, D. W., Greig, D., and Travisano, M. (2015). Origins of multicellular evolvability in snowflake yeast. *Nature communications*, 6(1):6102.
- Serván, C. A. and Allesina, S. (2021). Tractable models of ecological assembly. *Ecology Letters*, 24(5):1029–1037.
- Smith, J. M. and Szathmáry, E. (1997). *The major transitions in evolution*. OUP Oxford.
- Soros, L. and Stanley, K. (2014). Identifying necessary conditions for open-ended evolution through the artificial life world of chromaria. In *ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 793–800. MIT Press.
- Tang, S., Pawar, S., and Allesina, S. (2014). Correlation between interaction strengths drives stability in large ecological networks. *Ecology letters*, 17(9):1094–1100.
- Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., et al. (2016). Open-ended evolution: Perspectives from the oee workshop in york. *Artificial life*, 22(3):408–423.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *nature*, 393(6684):440–442.