Physica A 651 (2024) 130000

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Minimal specialization: Coevolution of network structure and dynamics

Annika King [a], Dallas Smith [b], Benjamin Webb [a,*]

[a] Department of Mathematics, Brigham Young University, Provo, UT 84602, USA
[b] Department of Mathematics, Utah Valley University, Orem, UT 84058, USA

ARTICLE INFO

Keywords:
Complex networks
Network growth models
Specialization
Equitable partitions
Bottlenecks

ABSTRACT

The changing topology of a network is driven by the need to maintain or optimize network function. As this function is often related to moving quantities such as traffic, information, etc., efficiently through the network, the structure of the network and the dynamics on the network directly depend on the other. To model this interplay of network structure and dynamics we use the *dynamics on* the network, or the dynamical processes the network models, to influence the *dynamics of* the network structure, i.e., to determine where and when to modify the network structure. We model the dynamics on the network using Jackson network dynamics and the dynamics of the network structure using *minimal specialization*, a variant of the more general network growth model known as *specialization*. The resulting model, which we refer to as the *integrated specialization model*, coevolves both the structure and the dynamics of the network. We show that by reducing dynamic bottlenecks, i.e. reducing the network's maximal asymptotic load, this model simultaneously produces networks with real-world like properties. This includes right-skewed degree distributions, sparsity, the small-world property, and non-trivial equitable partitions. Additionally, when compared to other growth models, the integrated specialization model creates networks with small diameter, minimizing distances across the network. The result are networks that have both dynamic and structural features that allow quantities to more efficiently move through the network.

1. Introduction

Networks studied in the biological, social, and technological sciences are inherently dynamic in that the state of the network's components evolve overtime. Technological and traffic networks show phase-transition type dynamics [1,2], gene regulatory networks experience boolean dynamics [1,3], metabolic networks exhibit flux-balance dynamics [1,4], and the human brain has been shown to have synchronous and other dynamic behavior [1,5].

The first type of dynamics is most often referred to as the *dynamics on the network*, referring to the changing states of the network's components. The second type of dynamics, which is the evolving *topology* of the network, is referred to as the *dynamics of the network*. To a large extent the study of network dynamics focuses on one of these two types of dynamics, meaning either the network's structure is fixed and the dynamics on the network are studied, or the dynamics on the network are ignored and the evolving structure of the network is studied. However, in real-world networks, these two types of dynamics typically influence one another [6–8]. For instance, as traffic increases in a traffic network, the Internet, a supply chain, etc., new routes are added, which

* Corresponding author.
E-mail addresses: aking@mathematics.byu.edu (A. King), dallas.smith@uvu.edu (D. Smith), bwebb@mathematics.byu.edu (B. Webb).

https://doi.org/10.1016/j.physa.2024.130000
Received 27 November 2023; Received in revised form 22 July 2024
Available online 29 July 2024

in turn creates new traffic patterns. In other areas the same principle holds. For instance, as money, etc. moves through a banking network new paths emerge depending on the need to efficiently route these assets [9].

Here, we consider the interplay between these two types of dynamics. Specifically, we introduce a model that uses the dynamics on the network to determine where to evolve the network's structure. This change in structure in turn alters the dynamics on the network, and this coevolving back and forth of structure and dynamics is analyzed. For the dynamics on the network, we use a Jackson network model, which describes queuing systems, e.g. transportation networks, website traffic, etc. [10]. Under certain conditions, Jackson networks have a *stationary distribution*, which can be thought of as a *globally attracting fixed point* that the dynamics on the network tend to. This allows us to determine asymptotically high-stress areas, or areas of *maximal asymptotic load*, within the network [11]. These areas with maximal asymptotic load are where we perform the structural evolution of the network via *minimal specialization*, which maintains the function of the network while separating the number of tasks, i.e. load, of these high-stress areas.

Specialization, as a mechanism of growth, is a phenomenon observed in many real-world networks including biological [12–14], social [15], and airline hubs in transportation networks [16]. Specialization allows a network to perform increasingly complex tasks by copying parts of the network and dividing the original network connections between these copies. The resulting specialized network maintains the functionality of the network by preserving all the network paths so that the ability to route information, etc., is maintained. Recently, network growth models for specialization have been studied in [17–19], where the authors describe how specialization creates real-world properties in a network over time [17], maintains intrinsic stability [18], and creates synchronous dynamics [19].

The specific model we propose uses what we refer to as *minimal specialization*. Minimal specialization is effectively the specialization defined in [17], the difference is that it modifies the least number of nodes and edges possible when specializing a network. After defining minimal specialization in Section 3 we prove some useful facts regarding its spectral and structural behavior. Specifically, we prove that if we start with a strongly connected or primitive network, minimal specialization of the network will result in a strongly connected or primitive network (see Lemma 3.1.1 and Theorem 3.1.1). Additionally, we show how eigenvalues and eigenvectors evolve with minimal specialization (see Lemma 3.1.2). We also prove that minimal specialization creates and preserves non-trivial equitable partitions (see Theorem 7.0.1). We will use minimal specialization to evolve the structure of the Jackson network in the area of highest maximal asymptotic load (see Sections 3 and 4). As this method incorporates both structure and dynamics we refer to it as the *integrated specialization model*. We show that the integrated specialization model creates networks that have structures observed in many real-world networks, including right-skewed degree distributions, sparsity, and the small-world property. It is worth emphasizing that these features emerge from a topological mechanism, i.e. specialization, designed to reduce the network's maximal asymptotic load. Moreover, when compared to other growth models, the integrated specialization model creates networks with small diameter, i.e., networks that have relatively small distances across the network. Aside from these structural features, the integrated specialization model sequentially removes the network's largest bottlenecks as the network evolves. This results in networks whose structure and dynamics are both increasingly well-adapted to allow quantities to move efficiently through the network.

The paper is structured as follows. Section 2 introduces basic notation, Jackson networks, and the equations used to model the dynamics on Jackson networks. In Section 3, we motivate and introduce the concept of *minimal specialization*, showing that certain topological and dynamical properties of the original network are maintained under these operations. In Section 4, we use the dynamics, i.e., the stationary distribution, on a Jackson network to determine where to perform minimal specialization of the network. We then extend this method, which we refer to as *minimal dynamic specialization*, to more general networks using the network's *eigenvector centrality* to specialize the network structure. In Section 5, we show that growing a Jackson network repeatedly using minimal dynamic specialization, referred to as the *integrated specialization model*, creates structural properties observed in many real-world networks. We also show that the maximum eigenvector centrality of the network decreases, meaning the integrated model sequentially reduces areas of high stress in the network, on average. These experiments are carried out for Erdös-Rényi (ER) graphs in this section and later in the Appendix for random geometric (RG), Barabśi-Albert (BA), and Watts–Strogatz (WS) graphs. Section 6 compares the integrated specialization model to other growth models, where we show that the integrated specialization model is comparable in creating common real-world properties while being more efficient in creating small diameter networks and networks with equidistributed traffic loads. Section 7 introduces equitable partitions and shows that minimal specialization creates and preserves non-trivial equitable partition elements, which are related to symmetries in the network and are common structures observed in many real-world networks. A natural question is whether for other initial graphs types we see the same qualitative behavior as seen here. This question is further explored in the Appendix

2. Background

Real-world networks perform specific functions. The underlying structure of a network, represented by a graph, is key to its performance [20–22]. Formally, this is given by the graph $G = (V, E, W)$, where V is a *vertex set* (or node set) and E is an *edge set*. The vertices in V represent the network components, or objects, while the edges in E represent the connections or interactions between these objects. We let $V = \{1, 2, \dots, n\}$ with i representing the ith component of the network. An edge from i to j, denoted e_{ij} is used to represent the ith node affecting the jth node. Here, the graphs we consider are directed graphs, meaning edges of the graph are directed, noting that undirected graphs can be viewed as a special case of directed graphs. The function $W : E \to \mathbb{R}$ gives the edge weight of each $e_{ij} \in E$, where $W(e_{ij})$ can represent the strength of the interaction, transitions probabilities, etc. If G is an unweighted graph, then $W(e_{ij}) = \delta_{ij}$ and we write $G = (V, E)$.

The underlying graph structure $G = (V, E, W)$ of a network can alternatively be represented by its *adjacency matrix* $A = A(G) \in \mathbb{R}^{n\times n}$, with entries

$$A_{ij} = \begin{cases} W(e_{ji}) & \text{if } e_{ji} \in E \\ 0 & \text{else.} \end{cases}$$

We note that this orientation corresponds to right multiplication by a column vector. We let $\sigma(A)$ denote the *eigenvalues* of the matrix A, and $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$ the *spectral radius* of A. Since there is a one-to-one relationship between a graph and its adjacency matrix, we will use the two interchangeably.

To study the interplay between the dynamics on and of a network, we consider Jackson networks, which are used to model queuing systems [10]. The flow of a given quantity, e.g. traffic, information, etc., through a Jackson network is modeled as the discrete-time affine dynamical system

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} + \boldsymbol{\gamma}$$

where $\mathbf{x}^{(k)} \in \mathbb{R}^n$ is the *state* of the network giving the state $x_i^{(k)}$ of each component at time $k \geq 0$. The graph $G = (V, E, W)$ associated with the Jackson network is the graph with the adjacency matrix $A \in [0,1]^{n\times n}$, called the system's *transition matrix* where $A_{ij} \in [0,1]$ is the probability of transitioning from vertex j to vertex i. The vector $\boldsymbol{\gamma} \in \mathbb{R}^n$ gives the *external inputs* to the system. A Jackson network can also have *internal loss*, meaning there is a non-zero probability of information, traffic, etc., leaving the system. This occurs when the probability of transitioning from node i to any other node is less than 1, i.e., the ith column of the transition matrix sums to less than 1. For simplicity, we assume there is no external input or internal loss, meaning $\boldsymbol{\gamma} \equiv \mathbf{0}$ and A is column stochastic. Thus,

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} = A^k\mathbf{x}^{(0)} \tag{1}$$

for the *initial vector* of quantities $\mathbf{x}^{(0)} \in \mathbb{R}^n$. This configuration induces the discrete-time linear dynamical system $(A, \mathbb{R}^n)$ where the matrix $A = A(G)$ is the adjacency matrix of the graph $G = (V, E, W)$.

The asymptotic behavior of linear systems, such as the Jackson networks $(A, \mathbb{R}^n)$, are fairly well understood. If A is primitive, then the *spectral radius* is an algebraically simple eigenvalue of A and $\lim_{k\to\infty}(\rho^{-1}A)^k = \mathbf{x}\mathbf{y}^\intercal$ where $\mathbf{x}$ and $\mathbf{y}$ are the right and left leading eigenvectors associated with ρ, respectively, with $\|\mathbf{x}\|_1 = 1$ and $\mathbf{x}^\intercal\mathbf{y} = 1$. Moreover, since A is stochastic, we have $\rho = 1$ and $\mathbf{y} = \mathbb{1}$. That is, for the initial condition $\mathbf{x}^{(0)} \succeq 0$, we have

$$\lim_{k\to\infty} A^k\mathbf{x}^{(0)} = \mathbf{x}\mathbb{1}^\intercal\mathbf{x}^{(0)} = \|\mathbf{x}^{(0)}\|_1\mathbf{x}. \tag{2}$$

Thus the sum of the quantities moving through the network is constant in time [23]. Additionally the *asymptotic state* of the system is a scaled version of the leading eigenvector $\mathbf{x}$, which when normalized with the 1-norm, is referred to as the systems *stationary distribution* [11]. For a Jackson network $(A, \mathbb{R}^n)$ with primitive matrix A, the stationary distribution is a *globally attracting fixed point* of the system (see Eq. (2)). This stationary distribution indicates the long-term dynamics of the Jackson network $(A, \mathbb{R}^n)$. Specifically, it tells us which nodes, on average, carry the highest amount of information, traffic, stress, etc. The main idea behind the model we propose is to use this globally attracting fixed point to determine where the long-term, high-stress areas of the network are (see Section 4). To alleviate this stress, we modify the structure of the network accordingly, using the notion of *minimal specialization*.

3. Topological network dynamics: *Minimal specialization*

In order to model the interplay between the dynamics on and the dynamics of the network, we need to determine *how* to evolve the structure of the network. As described in the introduction, network specialization has been observed in a number of real-world networks [12–16], and specialization models have been the focus of a number of recent papers [17–19]. In this work, we explore two main deviations from these specialization models. The first is related to the observation that most real-world specializations occur at small scales. For example, if a transportation route experiences high use, usually the route is only modified at its point of highest traffic. To reflect this, our model of minimal specialization adds the fewest number of nodes and edges required to modify the flow through the network while maintaining functionality.

Definition 3.0.1 (*Minimal Specialization*)**.** For the graph $G = (V, E, W)$ with $|V| > 1$, let $i \in V$ such that i has at least two out-going edges. Let $e_{ij} \in E$ with $i \neq j$ and let $w = W(e_{ij})$. Let $\overline{G} = \left(\overline{V}, \overline{E}, \overline{W}\right)$ be the graph where

(i) $\overline{V} = V \cup \{\bar{i}\}$;

(ii) $\overline{E} = \left(E \setminus \{e_{ij}\}\right) \cup \{e_{\bar{i}j}\} \cup \{e_{k\bar{i}} \mid \exists\, e_{ki} \in E\}$; and

(iii) $$\overline{W}(e_{\alpha\beta}) = \begin{cases} 1 & \text{if} \quad \alpha = \bar{i},\ \beta = j \\ (1-w)^{-1}W(e_{i\beta}) & \text{if} \quad \alpha = i,\ \beta \neq j \\ (1-w)W(e_{\alpha i}) & \text{if} \quad \alpha \neq i,\ \beta = i \\ wW(e_{\alpha i}) & \text{if} \quad \alpha \neq i,\ \beta = \bar{i}. \\ W(e_{ii}) & \text{if} \quad \alpha = i, \beta = i \\ w(1-w)^{-1}W(e_{ii}) & \text{if} \quad \alpha = i, \beta = \bar{i} \\ W(e_{\alpha\beta}) & \text{else.} \end{cases}$$

We refer to the graph $\overline{G} = \left(\overline{V}, \overline{E}, \overline{W}\right)$ associated with the system $\left(\overline{A}, \mathbb{R}^{(n+1)}\right)$ as the *minimal specialization* of the graph G over $i \in V$ with edge $e_{ij} \in E$. We refer to node $\bar{i}$ as the specialized copy of node i.

In Definition 3.0.1, $\bar{i} \in \overline{V}$ can be thought of as a copy of $i \in V$, whose specialized function is to maintain the weighted connections to $j \in V$ that were previously maintained by i through e_{ij}. This results in vertices specialized into two parts, where i performs its previous task with the exception of routing traffic to j, which is now executed by its specialized copy $\bar{i}$ (see Example 3.1.1).

It is important that $i \in V$ has at least two out-going edges in order to be specialized. If it has only one out-going edge, then after specialization, it would have no out-going edges, becoming a *dangling node*, essentially meaning that information, traffic, etc., gets trapped at this node.

We note that the edge weight update in Definition 3.0.1 maintains the stochastic nature of the Jackson network (see Section 3.1). This manner of updating could be done for any $0 < w < 1$ and still produce a stochastic system, but choosing $w = W(e_{ij})$ maintains the proportion of network quantities passing through i and its copy $\bar{i}$ after specialization that previously passed through i before specialization.

3.1. Topological and dynamical properties of minimal specialization

Minimal specialization induces a new system $(\overline{A}, \mathbb{R}^{n+1})$ with new dynamics. Here we show this new system is also a Jackson network that inherits the structural and stability properties of the original Jackson network $(A, \mathbb{R}^n)$ under mild conditions. In particular, we will show that if A is primitive then $\overline{A}$ is primitive, so that the specialized Jackson network $(\overline{A}, \mathbb{R}^{n+1})$ has a stationary distribution $\overline{\mathbf{x}} \in \mathbb{R}^{n+1}$. To prove this requires the following two lemmata.

Lemma 3.1.1 (*Preservation of a Strongly Connected Graph*). *If $\overline{G} = \left(\overline{V}, \overline{E}, \overline{W}\right)$ is a minimal specialization of the graph $G = (V, E, W)$ and G is strongly connected, then $\overline{G}$ is strongly connected.*

Proof. We note that under minimal specialization, the only deleted edge in G is e_{ij}, so any path in G that does not contain e_{ij} will also be in $\overline{G}$. Any path that contained e_{ij} can now use the edge $e_{\bar{i}j}$ since all the nodes having edges to i now have edges to $\bar{i}$. Thus we only need to verify three things: (1) there is a path in $\overline{G}$ from i to j, (2) there are paths in $\overline{G}$ from any node to $\bar{i}$, and (3) from $\bar{i}$ to any node.

First, we note that by constraints in Definition 3.0.1, $i \in V$ has an edge $e_{ik} \in E$ with $k \neq j$. Moreover, since G is strongly connected, there is at least one edge $e_{hi} \in E$ leading to i. Thus, there is a path from k to h and we have $e_{h\bar{i}} \in \overline{E}$. Therefore, we have the path $i \to k \to \cdots \to h \to \bar{i} \to j$.

Second, there is a path in G from any node to i and this path does not contain e_{ij} since it would create a cycle. Thus, we can take the penultimate node in that path and traverse the edge from that node to $\bar{i}$, and thus we have a path in $\overline{G}$ ending at $\bar{i}$.

Finally, it is easy to see there is a path from $\bar{i}$ to any node, since we have the edge $e_{\bar{i}j}$, and there is a path from j to any node. □

Lemma 3.1.2 (*Spectral Evolution of Jackson Networks*). *Suppose $\overline{G} = \left(\overline{V}, \overline{E}, \overline{W}\right)$ is the minimal specialization of $G = (V, E, W)$ over $i \in V$ with edge $e_{ij} \in E$. If $w = W(e_{ij})$ and $(\lambda, \mathbf{x})$ is an eigenpair of $A = A(G)$ then $(\lambda, \overline{\mathbf{x}})$ is an eigenpair of $\overline{A} = A(\overline{G})$ where,*

$$\overline{x}_\ell = \begin{cases} (1-w)x_\ell & \text{if } \ell = i \\ w x_\ell & \text{if } \ell = \bar{i} \\ x_\ell & \text{else.} \end{cases}$$

In particular, $x_i, \overline{x}_i$, and $\overline{x}_{\bar{i}}$ have the same sign with $x_i = \overline{x}_i + \overline{x}_{\bar{i}}$ so $\|x\|_1 = \|\overline{x}\|_1$. The additional eigenvalue of $\overline{A}$ is 0, meaning $\sigma(\overline{A}) = \sigma(A) \cup \{0\}$.

Proof. We will first consider how minimal specialization transforms the adjacency matrix. Without loss of generality, we can order the rows and columns of our matrix so that we have

$$A = \left[\begin{array}{cccc|c} & & A[\{ji\}^c, \{i\}^c] & & \begin{matrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{(j-1)i} \end{matrix} \\ \hline A_{j1} & A_{j2} & \cdots & A_{jj} & w \\ A_{i1} & A_{i2} & \cdots & A_{ij} & A_{ii} \end{array}\right] \in \mathbb{R}^{n \times n} \tag{3}$$

where $A[\{j,i\}^c, \{i\}^c]$ denotes the matrix A with the jth and ith rows excluded and the ith column excluded.

When we perform minimal specialization, the structural changes and weight updates found in Definition 3.0.1 are reflected in the new adjacency matrix as follows:

$$\overline{A} = \left[\begin{array}{ccc|c|c} & A[\{j,i,\bar{i}\}^c,\{i,\bar{i}\}^c] & & \begin{matrix}\frac{A_{1i}}{(1-w)}\\ \frac{A_{2i}}{(1-w)}\\ \vdots \\ \frac{A_{(j-1)i}}{(1-w)}\end{matrix} & \mathbf{0} \\ \hline A_{j1} & \dots & A_{jj} & 0 & 1 \\ (1-w)A_{i1} & \dots & (1-w)A_{ij} & A_{ii} & 0 \\ wA_{i1} & \dots & wA_{ij} & \frac{wA_{ii}}{1-w} & 0 \end{array}\right] \in \mathbb{R}^{(n+1)\times(n+1)}. \tag{4}$$

Let $\overline{\mathbf{x}}$ be the vector whose entries are given by

$$\overline{x}_\ell = \begin{cases} (1-w)x_\ell & \text{if } \ell = i \\ wx_\ell & \text{if } \ell = \bar{i} \\ x_\ell & \text{else} \end{cases}.$$

Then for $k \in \{1, \dots j-1\}$, the kth entry of the product $\overline{A}\overline{\mathbf{x}}$ is

$$\left(\overline{A}\overline{\mathbf{x}}\right)_k = \sum_{s=1}^{j} A_{ks}x_s + \frac{A_{ki}}{1-w}(1-w)x_i + 0 \cdot wx_i = \sum_{s=1}^{j} A_{ks}x_s + A_{ki}x_i = (A\mathbf{x})_k = \lambda x_k.$$

For the jth entry of the product $\overline{A}\overline{\mathbf{x}}$ we have,

$$\left(\overline{A}\overline{\mathbf{x}}\right)_j = \sum_{s=1}^{j} A_{js}x_s + 0 + wx_i = \sum_{s=1}^{j} A_{js}x_s + A_{ji}x_i = (A\mathbf{x})_j = \lambda x_j.$$

For the ith entry of the product $\overline{A}\overline{\mathbf{x}}$ we have,

$$\left(\overline{A}\overline{\mathbf{x}}\right)_i = \sum_{s=1}^{j}(1-w)A_{is}x_s + A_{ii}(1-w)x_i + 0 = (1-w)\sum_{s=1}^{i} A_{is}x_s = (1-w)\,(A\mathbf{x})_i = \lambda(1-w)x_i.$$

Finally, for the $\bar{i}th$ entry of the product $\overline{A}\overline{\mathbf{x}}$ we have,

$$\left(\overline{A}\overline{\mathbf{x}}\right)_{\bar{i}} = \sum_{s=1}^{j} wA_{is}x_s + \frac{wA_{ii}}{1-w}(1-w)x_i + 0 = w\sum_{s=1}^{i} A_{is}x_s = w\,(A\mathbf{x})_i = \lambda w x_i.$$

Thus we have the product $\overline{A}\overline{\mathbf{x}} = \lambda\overline{\mathbf{x}}$ and therefore $\overline{\mathbf{x}}$ is an eigenvector with eigenvalue λ.

Finally, the last two rows of $\overline{A}$ are scalar multiples of each other. Since the last row is a new row, a new linear dependence is created in the matrix, and thus the additional eigenvalue is 0. □

Theorem 3.1.1 (*Preservation of Asymptotic Dynamics*)**.** *Let $G = (V, E, W)$ and assume $A = A(G) \in [0,1]^{n\times n}$ is primitive and stochastic. Then $\overline{A} = A(\overline{G}) \in [0,1]^{(n+1)\times(n+1)}$, associated with any minimal specialization $\overline{G} = (\overline{V}, \overline{E}, \overline{W})$, is primitive and stochastic. Therefore,*

$$\lim_{n\to\infty} \overline{A}^k \overline{\mathbf{x}}^{(0)} = \|\overline{\mathbf{x}}^{(0)}\|_1 \overline{\mathbf{x}} \quad \textit{for any} \quad \overline{\mathbf{x}}^{(0)} \in \mathbb{R}^{n+1}_{\geq 0}$$

where $\overline{\mathbf{x}}$ is the stationary distribution of $\overline{A}$.

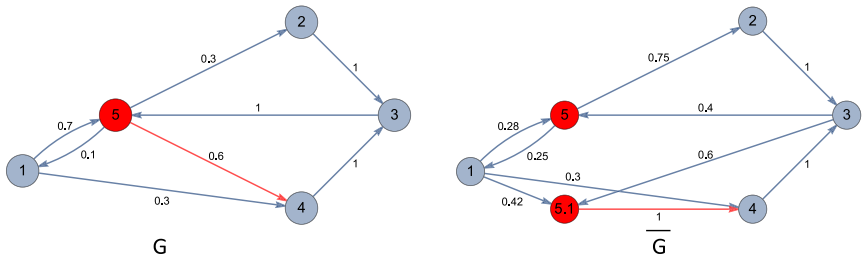
Proof. Given that A is stochastic and using Equation 3.1.2 in Lemma 3.1.2, we see that for $k \in \{1, 2, \dots, j\} \cup \{\bar{i}\}$, the kth column of $\overline{A}$ sums to 1. For column i, it follows from combining A_{ii} and $w(1-w)^{-1}A_{ii}$ and recognizing that the sum of the ith column of A excluding A_{ji} is $1-w$. Thus, $\overline{A}$ is stochastic.

From Lemma 3.1.1, we have that $\overline{G}$ is strongly connected. From Lemma 3.1.2, we have $\sigma(\overline{A}) = \sigma(A) \cup \{0\}$. Since A is primitive, $\rho(A) > 0$ is the only eigenvalue of maximum modulus. Thus $\rho(\overline{A}) = \rho(A)$ is the only eigenvalue of $\overline{A}$ of maximal modulus. Thus $\overline{A}$ is primitive (see [23] for the definition of primitive and equivalent characterizations). The dynamic consequences now follow from arguments in Section 2. □

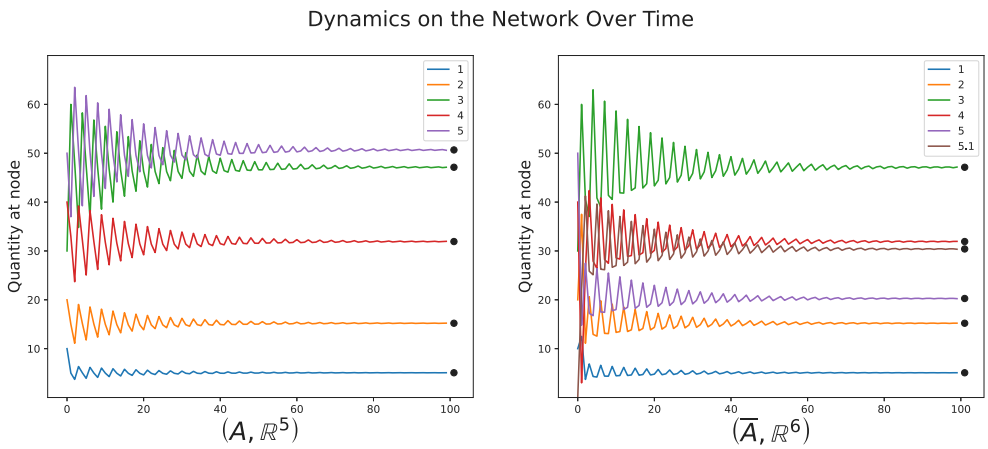
Example 3.1.1 (*Minimal Specialization*)**.** Consider the graph $G = (V, E, W)$ in Fig. 1(a) (left) with vertex set $V = \{1, 2, 3, 4, 5\}$. After performing minimal specialization on node 5 over the edge $e_{54} \in E$, the result is the graph $\overline{G} = (\overline{V}, \overline{E}, \overline{W})$ shown in Fig. 1(a) (right) with vertex set $\overline{V} = \{1, 2, 3, 4, 5, 5.1\}$, where $\bar{i} = 5.1$ is the copy of vertex $i = 5$. The dynamics of the original and specialized Jackson networks $(A, \mathbb{R}^5)$ and $(\overline{A}, \mathbb{R}^6)$ are shown in Fig. 1(b), left and right, respectively.

As $A \in [0,1]^{5\times 5}$ is primitive, the Jackson network $(A, \mathbb{R}^5)$ has the stationary distribution

$$\mathbf{x} = [.0337838, .101351, .314189, .212838, .337838]^{\mathsf{T}}.$$



(a) An example of *minimal specialization* where the graph $\overline{G}$ on the right is a specialized version of the graph G on the left. Here, node 5 with edge e_{54} is specialized, via minimal specialization, resulting in the node 5.1 (red on left). An edge from node 5.1 to node 4 is added (red on right). Nodes 1 and 3, which have an edge into node 5 now also have edges into node 5.1. Additionally, the edge from node 5 to node 4 is deleted but all other outgoing edges of 5 remain. The weights are updated according to the given scheme (see Definition 3.0.1).



(b) The dynamics on the network over time, for the networks in Figure 1a. For the initial network, the system's dynamics are shown for the initial condition $\mathbf{x}^{(0)} = [10, 20, 30, 40, 50]$ (left). Here, the long-term behavior at each node tends to a scaled version of $\mathbf{x}$, the stationary distribution the Jackson network $(A, \mathbb{R}^5)$, which is represented by the black dots. After specialization (right), we see that the network has similar behavior to the original network. Specifically, using the initial condition $\overline{\mathbf{x}}^{(0)} = [10, 20, 30, 40, 50, 0]$, the dynamics on the network converge to a scaled version of $\overline{\mathbf{x}}$, the stationary distribution of $\left(\overline{A}, \mathbb{R}^6\right)$, and the load at node 5 decreases.

Fig. 1. Example of minimal specialization.

Similarly, $\overline{A} \in \mathbb{R}^6$ is primitive and the associated Jackson network has the stationary distribution

$$\overline{\mathbf{x}} = [.0337838, .101351, .314189, .212838, .135135, .202703]^\intercal.$$

In Fig. 1(b) the initial conditions $\mathbf{x}^{(0)} = [10, 20, 30, 40, 50]^\intercal$ and $\overline{\mathbf{x}}^{(0)} = [10, 20, 30, 40, 50, 0]^\intercal$ of the systems $\left(A, \mathbb{R}^5\right)$ and $\left(\overline{A}, \mathbb{R}^6\right)$, respectively, lead to similar dynamics. On the left, $A^k\mathbf{x}^{(0)}$ is calculated for $k = 0, 1, \ldots, 100$. We see that the long-term dynamics approach $\|\mathbf{x}^{(0)}\|_1\mathbf{x}$. After minimal specialization, we calculate $\overline{A}^k\overline{\mathbf{x}}^{(0)}$ for $k = 0, 1, \ldots, 100$. Again, the asymptotic state of the system is $\|\overline{\mathbf{x}}^{(0)}\|_1\overline{\mathbf{x}}$. Notice that $\|\mathbf{x}\| = \|\overline{\mathbf{x}}\|$, which is not a coincidence. This follows from Lemma 3.1.2, and is explored further in Section 4.

The second deviation from the original specialization models we propose is how we choose where to evolve the graph structure of the Jackson network. Previously, the structure was evolved stochastically by choosing vertices randomly from the graph to specialize [17–19]. Here our goal is to use the dynamics on the network to determine where to specialize the structure.

4. Coevolution of structure and dynamics

In real-world systems, the changing topology of the network is driven by the need to optimize the network's function, which is often related to moving quantities efficiently through the network. Dynamical processes such as traffic flow, information transfer, etc., put pressure on the network's topology to evolve in specific ways. In order to model this behavior, we use the dynamics on the network, or the dynamical processes the network models, to determine where to modify the network topology.

The node that experiences the highest traffic volume in a network is a natural candidate for the part of the network under the most stress. The idea is that once specialized, this node and its specialized copy now share the load, lessening the stress on the original node. A complicating factor in most real-world networks is that it may be difficult to determine which node experiences the most traffic. This may be the case, for instance, if the network dynamics are irregular, e.g., aperiodic, chaotic, etc. The reason we choose Jackson networks for our initial coevolution model is that, under mild assumptions, there is a clear hierarchy of which nodes experience more traffic. This is given by the network's stationary distribution.

4.1. Integrated specialization model

Suppose $(A, \mathbb{R}^n)$ is a Jackson network where $A \in [0,1]^{n\times n}$ is primitive. Then A has a unique stationary distribution $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$, where x_i is the *asymptotic use* of $i \in V$ in the associated network or graph $G = (V, E, W)$. The node that experiences the *maximal asymptotic load* is the node i such that $x_i = \max_\ell \{x_\ell\}_{\ell=1}^n$. To specialize the network relative to its dynamics, we choose node i with maximal asymptotic load from among the nodes with more than 2 out-going edges. Given $i \in V$ we choose a node $j \in V, j \neq i$, such that $W(e_{ij}) = \max\{W(e_{i\ell}) | e_{i\ell} \in E, \ell \neq i\}$, i.e., we choose the edge that transitions the most traffic away from node i. Specializing $(A, \mathbb{R}^n)$ over node i with edge e_{ij} results in the specialized Jackson network $(\overline{A}, \mathbb{R}^{n+1})$, which we refer to as the *minimal dynamic specialization* of $(A, \mathbb{R}^n)$.

We note that if the node with maximal asymptotic load is not unique, say nodes $i, k \in V$, specializing i with edge $e_{ij} \in E$ and then specializing k with edge $e_{kh} \in \overline{E}$ will result in a different (non-isomorphic) graph structure than if the network is specialized in the opposite order. However, the asymptotic dynamics for all nodes after these two minimal dynamic specializations, in either order, will be the same. Similarly, if the largest edge weight is not unique, the graph structure will be different depending on which edge is chosen, but the resulting asymptotic dynamics for all nodes will be the same. In practice, non-uniqueness rarely occurs. If it does, we randomly choose one of the nodes (or edges) to be used in the minimal dynamic specialization process.

If $A \in [0,1]^{n\times n}$ is primitive, by Theorem 3.1.1 it is possible to sequentially specialize the primitive Jackson network $(A, \mathbb{R}^n)$ via minimal dynamic specialization. For such networks we can define the following coevolving Jackson model which integrates both the dynamics on and the dynamics of the network.

Definition 4.1.1 (*Integrated Specialization Model*)**.** Let $(A_0, \mathbb{R}^n)$ be a Jackson network where $A_0 \in [0,1]^{n\times n}$ is primitive. We define the sequence of Jackson networks $\left\{\left(A_m, \mathbb{R}^{n+m}\right)\right\}_{m=0}^{\infty}$ to be the *integrated specialization model* with initial Jackson network $(A_0, \mathbb{R}^n)$ and $A_{m+1} = \overline{A}_m$ for $m \geq 0$ is the minimal dynamic specialization of A_m.

4.2. Irreducibility

Having a Jackson network $(A, \mathbb{R}^n)$ where $A \in [0,1]^{n\times n}$ is primitive is a strong condition. Primitivity is a property that is difficult to establish since it typically involves calculating eigenvalues, taking large powers of matrices, calculating path lengths, etc., all of which are computationally intensive for large networks. However, a condition that is more reasonable is for A to be irreducible, which is equivalent to having a strongly-connected graph (see, for instance, [23]).

Most real-world networks have a largest strongly-connected component that comprises the majority of the network [24]. Thus, the underlying graph structure of the network is strongly connected if we restrict our attention to the graph's largest strongly-connected component. This guarantees that the network's adjacency matrix A is irreducible. With irreducibly, we still have $\rho(A)$ as an algebraically simple eigenvalue with positive leading eigenvector, but we lose having a unique stationary distribution (i.e. stable asymptotic dynamics). This positive leading eigenvector is the network's *eigenvector centrality*, which gives a ranking of the nodes that takes into account the importance of a node relative to the importance of its neighbors [24]. A necessary condition for primitivity is irreducibly, so in the setting of primitivity, we still have a notion of *eigenvector centrality*. In fact, if we have a primitive, stochastic matrix A, the eigenvector centrality is equivalent (up to scaling) to the associated Jackson network's stationary distribution.

Similar to our minimal dynamic specialization, we can use this ranking to determine which node to use in minimal specialization, i.e., we choose a node with maximal eigenvector centrality, eligible for minimal specialization, and choose the edge in the same manner as in minimal dynamic specialization. Moreover, Lemma 3.1.1 tells us that if A is strongly connected, then so is $\overline{A}$. Thus, like with the integrated specialization model, we can use minimal dynamic specialization to create a sequence of irreducible Jackson networks.

Example 4.2.1 (*Minimal Dynamic Specialization*)**.** Consider the Jackson network $(A, \mathbb{R}^5)$ shown in Fig. 1(a) (left). The Jackson network's stationary distribution is

$$\mathbf{v} = \left[.0337838, .101351, .314189, .212838, .337838\right]^{\intercal},$$

so $\max\{x_i\} = x_5$ is the maximum load and $\max\{W(e_{5\ell})\} = W(e_{54})$. Thus the minimal dynamic specialization of $(A, \mathbb{R}^5)$ is the Jackson network $(\overline{A}, \mathbb{R}^6)$ shown in Fig. 1(a) (right). Note that the specialized network has the stationary distribution

$$\overline{\mathbf{v}} = \left[.0337838, .101351, .314189, .212838, .135135, .202703\right]^{\intercal}.$$

Fig. 1(b) shows the dynamics on the original network from Fig. 1(a) (left) and the specialized network (right). Over time, the quantities at each node tend to converge to the (scaled) quantities of $\mathbf{v}$, represented as black dots. After specialization, the dynamics on the network behave in the same way, converging to (a scaled) $\overline{\mathbf{v}}$. The total traffic load in these systems is conserved, but the maximal asymptotic load from $\left(A, \mathbb{R}^5\right)$ to $\left(\overline{A}, \mathbb{R}^6\right)$ is reduced.

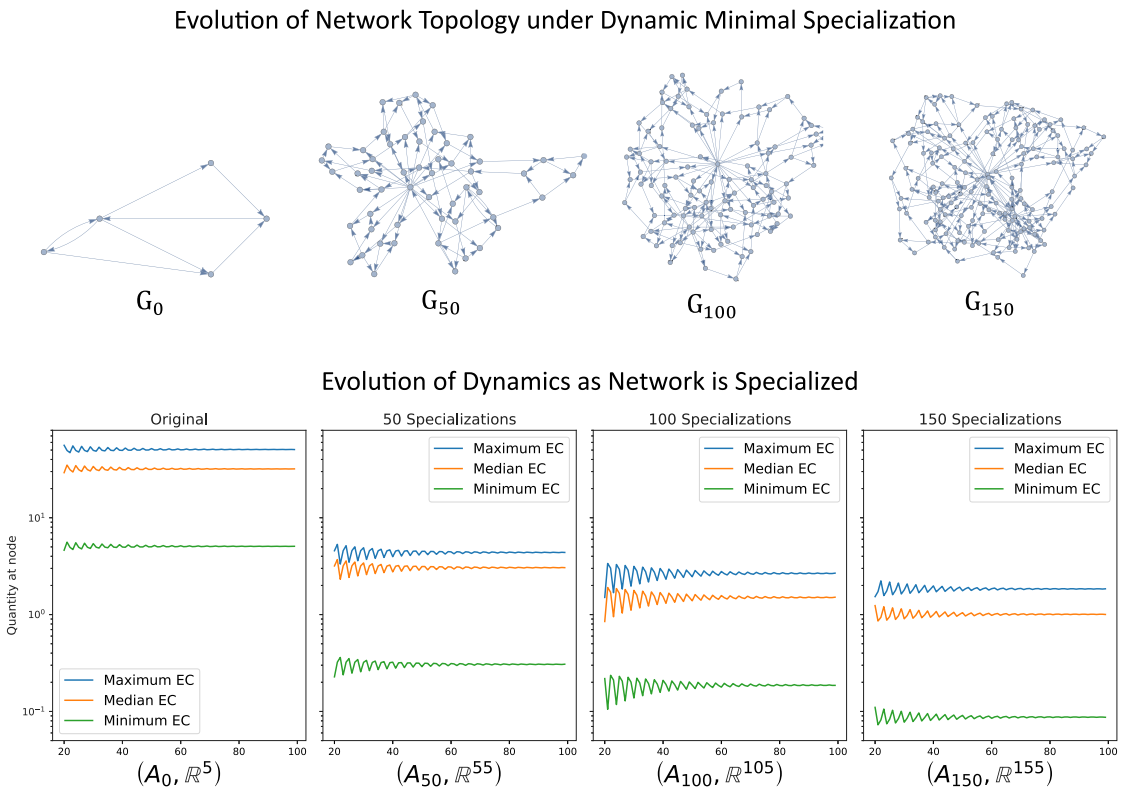


Fig. 2. The long-term dynamics associated with the Jackson network in Fig. 1(a) is shown in part, where the dynamics of the network are shown above and the dynamics on the network are shown below. The dynamics are generated using the integrated specialization model. Each plot on the bottom has the dynamics for the node with the highest eigenvector centrality, the node with the median eigenvector centrality, and the node with minimum eigenvector centrality, plotted on a log scale. As the network is repeatedly specialized, these quantities decrease by an order of magnitude.

Lemma 3.1.2 gives us a way of tracking how the maximum eigenvector centrality changes with the integrated specialization model. In particular, the sequence $\left\{\left(A_m, \mathbb{R}^{m+n}\right)\right\}_{m=0}^{\infty}$ has the sequence of leading eigenvectors $\mathbf{v}^{(0)}, \mathbf{v}^{(1)}, \mathbf{v}^{(2)} \cdots$, where

$$v_\ell^{(k+1)} = \begin{cases} (1-w)v_i^{(k)} & \text{if } \ell = i \\ wv_i^{(k)} & \text{if } \ell = \bar{i} \\ v_\ell^{(k)} & \text{else} \end{cases}$$

with $\|\mathbf{v}^{(0)}\|_1 = \|\mathbf{v}^{(1)}\| = \cdots \|\mathbf{v}^{(k)}\| = \cdots$. This allows us to recursively compare the maximum eigenvector centrality as we specialize. Consequently, Lemma 3.1.2 describes the evolution of the eigenvector centrality vector as a graph G with a column stochastic adjacency matrix is specialized under minimal dynamic specialization. Specifically, we have

$$\max_{1\le k\le n} v_k^{(m)} \ge \max_{1\le k\le n+1} v_k^{(m+1)} \quad \text{for } m \ge 0$$

in our sequence of leading eigenvectors. Thus, we are targeting and decreasing the areas of high stress or high importance on the network.

Example 4.2.2. Sequentially specializing the Jackson network $\left(A, \mathbb{R}^5\right)$ in Example 1(a) using the integrated specialization model results in the sequence $\left\{\left(A_m, \mathbb{R}^{5+m}\right)\right\}_{m=0}^{150}$. Fig. 2 (bottom) shows the long-term dynamics for $m = 0, 50, 100, 150$. Each plot has the dynamics for the node with the highest eigenvector centrality, the node with the median eigenvector centrality, and the node with minimum eigenvector centrality, plotted on a log scale. As the network is specialized, these quantities are getting increasingly closer together and decrease by at least an order of magnitude. That is, as we specialize high-stress areas of the network, we are decreasing network's maximal asymptotic load, and repeatedly doing so creates more equidistributed asymptotic loads.

5. Network model properties of the integrated specialization model

There are numerous models that generate networks which achieve specific properties observed in many real-world networks. These properties include (i) right-skewed degree distributions, (ii) sparsity, (iii) the small-world property etc. An Erdös-Rényi

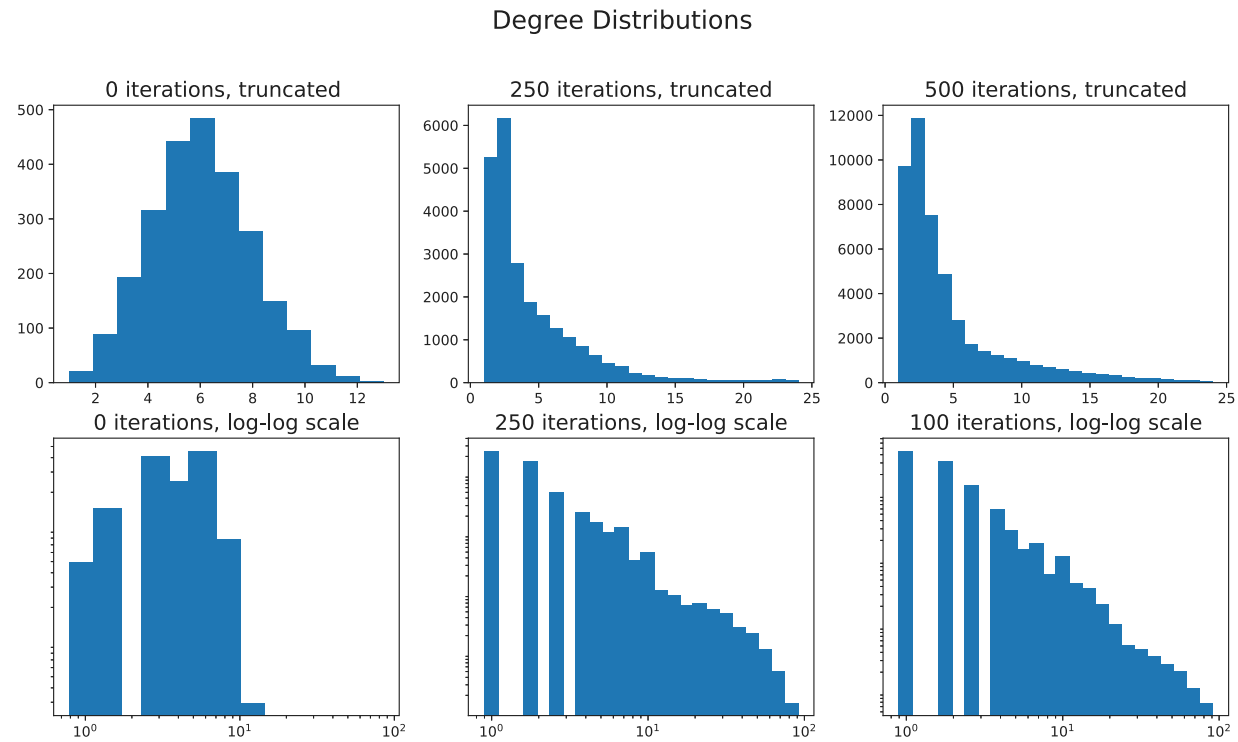


Fig. 3. Averaged results describing how the out-degree distribution of the network evolves using the integrated specialization model. The top row shows the degree distribution (here the tails are truncated). The bottom row shows the degree distribution on a log–log scale, where the binning is done logarithmically (where the tails are not truncated). As the network evolves, the out-degree distribution becomes increasingly right-skewed, a hallmark of many real-world networks. On a log–log scale, the degree distribution appear increasingly linear as the system evolves.

network has a giant component, a Barabási-Albert network has a scale-free right-skewed degree distribution, and a Watts–Strogatz network exhibits the small-world property [24,25]. In recent research, models have been proposed that create scale-free networks, the small-world property, and high clustering coefficients. See for instance, [26–28].

Since the sequence of vertices that are specialized for a given initial network $(A_0, \mathbb{R}^n)$ is difficult to determine, it is then difficult to determine the asymptotic structure of a network as it evolves under the integrated model. In this section we discuss some numerical results. We give statistical evidence that the integrated specialization model exhibits many observed network model properties including right-skewed degree distributions with some scale-free like properties, the small-world property, and sparsity. We also present, what is likely the most novel property of this model, which is the decrease in maximal asymptotic load of the network as a network is repeatedly specialized.

In our experiments, we begin with a Jackson network $(A_0, \mathbb{R}^{25})$ given by a directed Erdös-Rényi graph $G = (V, E, W)$, with $|V| = 25$ nodes and a density of $G_{den} = .25$, where for each ordered pair of nodes $\{i, j\}$ the edge e_{ij} belongs to the graph with probability G_{den}. The edge weights are uniformly assigned a value from 1 to 100 and normalized so out-going edges sum to 1. We use the integrated specialization model to produce the sequence $\left\{\left(A_m, \mathbb{R}^{25+m}\right)\right\}_{m=0}^{500}$, growing the network to 525 nodes. For this sequence of graphs, we collect statistics (i)-(iii) and repeat this experiment for 100 such initial Jackson networks. The averaged statistics with standard-deviations are described in the following subsections.

5.1. Degree distribution

Many real-world networks exhibit a right-skewed degree distribution, meaning many nodes have low degree and few nodes have high degree [24]. Since we are analyzing directed networks in our numerical experiment, we consider both the in-degree and out-degree distributions.

Fig. 3 shows our results regarding how the out-degree evolves, on average, over our 100 trials. The top left histogram, which has a distinct binomial shape, is the average histogram for 100 directed Erdös-Rényi graphs. This binomial shape is very unlike the right-skewed distribution found in real-world networks [24]. The top middle panel shows the averaged histogram for the networks $(A_{100}, \mathbb{R}^{125})$ over 100 simulates, and the top right panel is the averaged histogram of $(A_{200}, \mathbb{R}^{225})$ over 100 simulations. As these graphs are specialized, we see, on average, an increasingly right-skewed degree distribution, i.e., an increasingly more real-world like degree distribution.

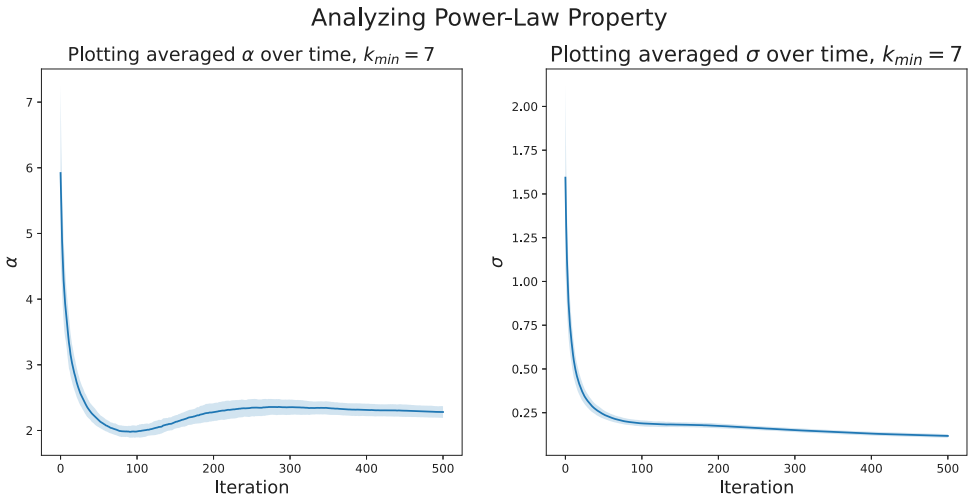


Fig. 4. Left: A plot of the evolution of average $\alpha(m)$-value and over one-hundred sequences $\left\{\left(A_m, \mathbb{R}^{25+m}\right)\right\}_{m=0}^{500}$ of the integrated model is shown. Right: A plot of the associated statistical error $\sigma(m)$-value average over the same sequences is shown. In each one standard-deviation is indicated.

There is a simple heuristic that helps explain why we see a right-skewed out-degree distribution. When we use the integrated specialization model, a node is created with only one out-going edge. As we sequentially specialize, it is possible for nodes with a small number of out-going edges to gain more out-going edges, but we are still adding a node with an out-degree of 1 at each iteration. Thus, although the degree-distribution does not reach a steady state as we are always increasing the out-degree of some nodes after minimal specialization, the distribution of the network's out-degree becomes increasingly right skewed.

Additionally, along with a right-skewed degree distribution, many real-world networks exhibit a power-law distribution, meaning this distribution is $d_k \sim Ck^{-\alpha}$ for some $C > 0$ where $\alpha \in [2,3]$ for many real-world networks [24]. Equivalently, the log–log distribution is linear with a slope of $-\alpha$ where $\alpha \in [2,3]$. The histograms shown in the bottom row of Fig. 3 are those in the top row but shown on a log–log scale, from left to right, respectively where the binning is done logarithmically. As the network evolves these degree distributions appear increasingly linear, specifically in the tail of the distribution (cf. Fig. 4).

For the integrated model we determine $\alpha = \alpha(m)$ for the out-degree distribution of each $\left\{\left(A_m, \mathbb{R}^{25+m}\right)\right\}_{m=0}^{500}$ using

$$\alpha = 1 + n\left[\sum_{i=1}^{n} \ln\left(\frac{k_i}{k_{min} - \frac{1}{2}}\right)\right]^{-1}$$

where we let $k_{min} = 7$ (see [24], p. 325). We also use the associated *statistical error* $\sigma = \sigma(m)$ for this sequence given by $\sigma = (\alpha - 1)/\sqrt{n}$.

Fig. 4 shows the exponent $\alpha = \alpha(m)$ (left) and statistical error $\sigma = \sigma(m)$ (right) in a power-law analysis of the out-degree distribution of the integrated model $\left\{\left(A_m, \mathbb{R}^{25+m}\right)\right\}_{m=0}^{500}$, averaged over 100 simulations. In Fig. 4 left we approach the value $\alpha \approx 2.3$ with little variance across each simulation, and decreasing statistical error σ as we specialize. Additionally, the bottom row of Fig. 3 shows the log-histogram of the degree distribution as we specialize. We can see that as we specialize, the log-histograms become increasingly linear.

We note that one drawback of the integrated specialization model is that the in-degree distribution does not evolve into a right-skewed distribution. To solve this, we could alternate between specializing G and G^T, where G^T is the graph with adjacency matrix A^T. However, we would not have the same theoretical properties as the integrated specialization model defined here.

5.2. Small-world property

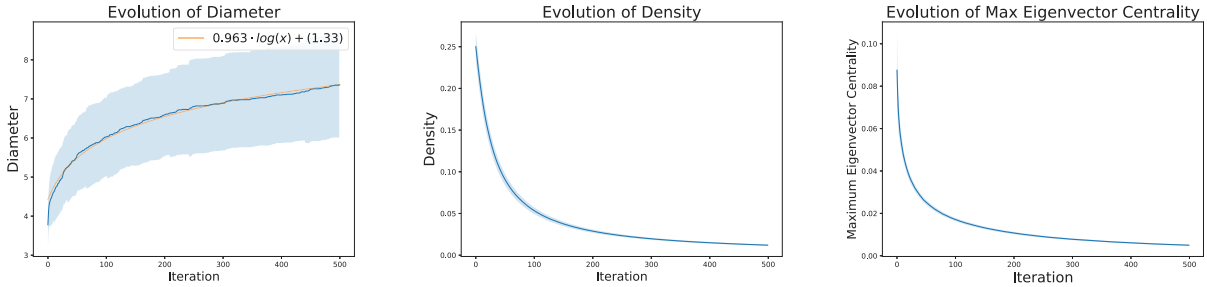
The diameter of a network is the network's largest geodesic, i.e., its longest shortest path. Intuitively, the length of the diameter is the farthest distance traffic, information, assets, etc., will travel in a network. It has been observed that as a real-world network evolves over time, its diameter grows logarithmically. This phenomenon is known as the *small-world Property* [24].

Fig. 5(a) shows how the diameter of the sequence $\left\{\left(A_m, \mathbb{R}^{25+m}\right)\right\}_{m=0}^{500}$ grows, averaged over 100 such sequences. The black curve is the evolution of the average diameter, with the shaded region representing one standard deviation. The orange curve is fitted to the black curve. The growth appears to be logarithmic, suggesting that the integrated specialization model has the small-world Property, at least for initial networks with an Erdös-Rényi topology.

5.3. Density

The density of a directed network $G = (V, E, W)$ is defined to be

$$G_{den} = \frac{m}{n(n-1)},$$



(a) Averaged results describing how the diameter of the network evolves using the integrated specialization model. One standard deviation is shown with the shaded region. The orange curve is a logarithmic equation fitted to the data. The trend of the diameter is logarithmic, i.e., the model exhibits the *small-world property*, a phenomenon that is typical of real-world networks.

(b) Averaged results describing how the density of the network evolves using the integrated specialization model. One standard deviation is shown, with the maximum standard deviation as 0.017 and the minimum as 0.0008. As the network evolves, the density decreases overall, a hallmark of many real-world networks.

(c) Averaged results describing how the maximum eigenvector centrality of the network evolves using the integrated specialization model. One standard deviation is shown, with the maximum standard deviation as 0.0178 and the minimum as 0.0003. As the network evolves, the maximum eigenvector centrality decreases, indicating that we are reducing areas of high stress in the network.

Fig. 5. Analyzing real-world properties.

where $m = |E|$ is the number of edges and $n = |V|$ is the number of nodes. The density can be thought of as the ratio of edges to possible edges in a graph. If the ratio tends to 0 as the network grows, the network is said to be *sparse*; otherwise, it is said to be *dense*. A hallmark of real-world networks is that they appear to be sparse when compared to random networks [29].

Fig. 5(b) shows how the density changes as a network is grown using the integrated specialization model. For the graphs we consider, the average density begins at $G_{den} = .25$ and the sparsity rapidly drops towards zero, at a nearly exponential rate. One standard deviation is shaded, but is too small to observe, indicating a very constrained evolution towards sparsity.

5.4. Maximum eigenvector centrality

In Section 4.1, we describe how eigenvector centrality is used in our dynamical model. To reiterate, real-world networks tend to specialize in areas of the network with high traffic or stress. The integrated specialization model is designed to specialize nodes with high eigenvector centrality, maximizing dynamic functionality by reducing high-stress areas. A consequence of Lemma 3.1.2 is that the maximal eigenvector centrality, i.e., the maximal asymptotic load, of a network cannot increase as the network is specialized.

Fig. 5(c) shows the averaged results for how the maximal asymptotic load evolves with the integrated specialization model. One standard deviation is shown, but is very small. On average, we see a rapid decrease in the maximum eigenvector centrality; thus, statistically, we do much better than the theory informs, efficiently targeting areas of high traffic in the network and successfully reducing the network's maximal asymptotic load, i.e. areas of stress.

We note that in this section we consider the integrated model $\{(A_m, \mathbb{R}^{25+m})\}_{m=0}^{500}$ for the initial condition $\{A_0, \mathbb{R}^{25}\}$ where A_0 corresponds to a directed Erdös-Rényi graph as described in Section 5.1. A natural question is whether for other initial graphs types we see the same qualitative behavior as seen here. This question is further explored in the Appendix where the same experiments are run for random geometric (RG), Barabśi-Albert (BA), and Watts–Strogatz (WS) graphs.

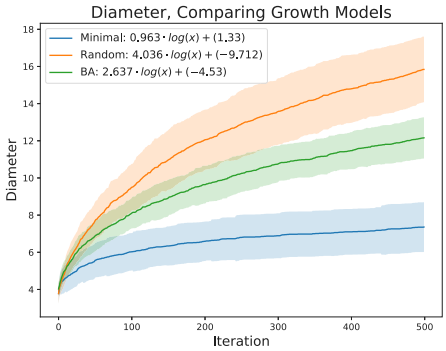
As the outcomes to these experiments are similar in each case, this suggest that, irrespective of the initial condition, the integrated model produces Jackson models with similar structure and dynamics.

6. Comparison to other models

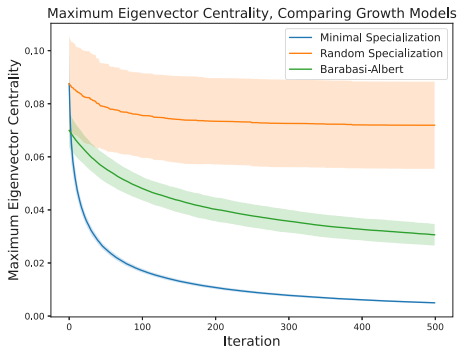
The most novel feature of the integrated specialization model is that it uses the dynamics on the network to determine where to evolve the structure of the network. In the previous section, we presented statistical evidence that the integrated specialization model creates real-world structural properties. To understand how this coevolution leads to structural differences, we compare the integrated specialization model to two other models: (i) *random minimal specialization* and (ii) the *Barabśi-Albert* (BA) model.

In random minimal specialization, we remove the integrated specialization model's dependence on the network's dynamics by performing minimal specialization over a random node and its out-going edge with the highest weight. This creates a sequence of Jackson networks $\left\{\left(R_m, \mathbb{R}^{m+n}\right)\right\}_{m\geq 0}$. In the Barabśi-Albert model we consider a single node preferentially added via a single edge at each step, resulting in a sequence of simple graphs $\{B_m\}_{m\geq 0}$, where $|B_m| = m + n$.

As in the previous section, we begin with an Erdös-Rényi graph $G_0 = (V, E, W)$ with $|V| = 25$ and $G_{den} = .25$ for each of the integrated specialization, random minimal specialization, and Barabśi-Albert Models. For the models involving minimal specialization, the initial Erdös-Rényi graph is directed, with edge weights uniformly assigned values in $[1, 100]$ and then normalized so out-going edge weights sum to 1. For the Barabśi-Albert Model, the initial Erdös-Rényi graph is undirected, and no edge-weights



(a) Figure comparing the evolution of the diameter for the integrated specialization, random minimal specialization, and Barabśi-Albert model (adding one edge at a time). 100 simulations of networks grown to 525 nodes were averaged. A line of best fit is shown for each model. In each case the models appear to have logarithmic growth, i.e., the *small-world* property.



(b) Figure comparing the evolution of the maximum eigenvector centrality for the integrated specialization, random minimal specialization, and Barabśi-Albert model (adding one edge at a time). 100 simulations of networks grown to 525 nodes were averaged. For the maximum eigenvector centrality over time, the integrated specialization model does a better job of decreasing the maximum eigenvector centrality and does so with little variance between simulations. Thus, the integrated specialization model is more efficient at targeting high-stress areas of the network and relieving that stress.

Fig. 6. Comparing models.

are assigned. For each model, we evolve the networks 500 times to create the sequences $\{(A_m, \mathbb{R}^{n+m})\}_{m=0}^{500}$, $\{(R_m, \mathbb{R}^{n+m})\}_{m=0}^{500}$, and $\{B_m\}_{m=0}^{500}$, respectively. This experiment is performed 100 times for each model and the data is averaged. We note that for each of these models, the evolution of the density is essentially the same, rapidly decreasing to zero at nearly the same rate. Similarly, the degree distributions all evolve to a right-skewed degree distribution. The three models also exhibit the small-world property, but show different growth rates (see Fig. 6(a)).

In the networks we consider, the maximum eigenvector centrality corresponds to the node that, on average, has the most information or traffic, i.e., the maximal asymptotic load. Thus, a decrease in maximum eigenvector centrality corresponds to a decrease in areas of high stress on the network. Fig. 6(b) shows the evolution of the maximum eigenvector centrality, or maximal asymptotic load, in each of the three growth models, where each eigenvector is normalized by the 1-norm. The integrated specialization model rapidly decreases the maximal asymptotic load, which is expected. Random minimal specialization and BA do not decrease as rapidly, or reach as low of a value as the integrated specialization model. Moreover, the variance across simulations for the integrated specialization model is negligible, while there is higher variance in the other two models.

The differences in the evolution of the diameter and maximal asymptotic load give evidence that the integrated specialization model creates efficient networks, both in terms structure and dynamics. Specifically, the integrated specialization model creates networks with small diameter, so that distances across the network are minimized, and equidistributed traffic, so that traffic bottlenecks are reduced throughout the network. Moreover, the negligible variance between trials for the integrated specialization model suggests that we achieve these results in a near-optimal way.

7. Equitable partitions

A hallmark of real-world networks is the high occurrence of symmetric structures [30]. Each such symmetry is given by an equitable partition, which is a generalization of the notion of a graph symmetry. Historically, equitable partitions are defined for *simple graphs*: unweighted, undirected graphs without loops. However, equitable partitions can be defined for unweighted directed graphs $G = (V, E)$ as follows [19]:

Definition 7.0.1 (*Equitable Partition*)**.** Let $G = (V, E)$ be a graph with adjacency matrix $A = A(G)$. Let $\pi = \{V_1, V_2, \dots, V_k\}$ be a partition of the vertices V. Then π is an equitable partition if the sum

$$\sum_{j \in V_b} A_{ij} = D_{ab} \tag{5}$$

is constant for any $i \in V_a$. If $|V_i| = 1$, we call V_i trivial, else we call it non-trivial. We call π trivial if each element in π is trivial and non-trivial if there exists a non-trivial element in π. We call the matrix $D \in \mathbb{N}^{k \times k}$ the divisor matrix of A associated with π, and G_π the divisor graph associated with D.

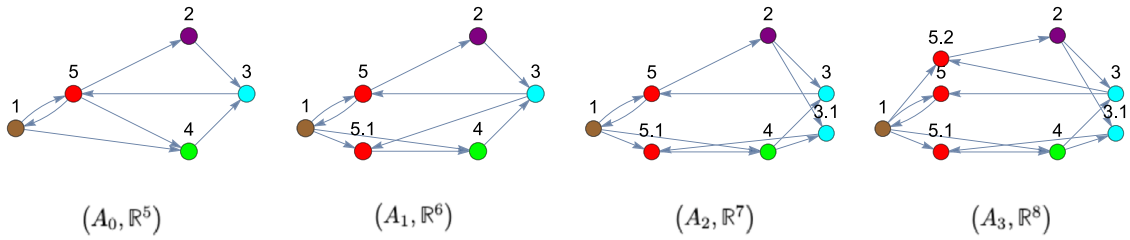


Fig. 7. An example illustrating the formation of non-trivial equitable partitions. On the left is the original network. From left to right, node 5 is specialized, with node 5.1 as its copy. These are in the same equitable partition element (red). Next, node 3 is specialized with node 3.1 as its copy, which are in the same partition element (cyan). Finally, node 5 is specialized again, with node 5.2 as its copy, increasing the partition element size from two to three (red).

We emphasize that our definition uses the unweighted adjacency matrix, meaning we are focusing on the topology of the network and not the edge weights. A Jackson network $(A, \mathbb{R}^n)$ has an equitable partition π if its associated unweighted graph $G = G(A)$ where $G = (V, E)$ has the equitable partition π.

The main result of this section is that minimal specialization creates and preserves non-trivial equitable partition elements. Later, we study the consequences of repeated minimal specialization on the size and type of equitable partitions a specialized network has (see Corollary 7.0.1 and Fig. 7).

Theorem 7.0.1 (*Preservation of Equitable Partitions*). *Let $\left(\overline{A}, \mathbb{R}^{n+1}\right)$ be the minimal specialization of the Jackson network $(A, \mathbb{R}^n)$ over any eligible vertex i with edge $e_{ij}, i \neq j$. If $(A, \mathbb{R}^n)$ has an equitable partition $\pi = \{V_1, V_2, \dots, V_k\}$, then $\left(\overline{A}, \mathbb{R}^{n+1}\right)$ has an equitable partition $\overline{\pi} = \{\overline{V}_1, \overline{V}_2, \dots \overline{V}_k\}$ where*

$$\overline{V}_a = \begin{cases} V_a \cup \{\bar{i}\} & \text{if } i \in V_a \\ V_a & \text{otherwise} \end{cases} \qquad \text{for } a = 1, 2, \dots, k$$

where $\bar{i}$ is the node created during minimal specialization.

Proof. We first note that Definition 7.0.1 is equivalent to saying we can partition our adjacency matrix so that the row sum in each partition is constant. Thus, we will consider the partitioned adjacency matrices corresponding to π and $\overline{\pi}$. Without loss of generality, let $j \in V_{k-1}$ and $i \in V_k$. The partitioned adjacency matrix is

$$A = \left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1k} \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline A_{(k-1)1} & \cdots & A_{(k-1)(k-1)} & A_{(k-1)k} \\ \hline A_{k1} & A_{k2} & A_{k(k-1)} & A_{kk} \end{array}\right] \in \mathbb{R}^{n \times n} \tag{6}$$

where each $A_{\ell\ell}$ is a block matrix that represents the connections within a part and each $A_{\ell h}, \ell \neq h$ a block matrix that represents the connections from part h to part ℓ.

Let $\tilde{A}_{(k-1)k}$ denote the matrix that has the same entries as $A_{(k-1)k}$ except the entry corresponding to e_{ij} in the adjacency matrix is changed from a 1 to a 0. This represents deleting the edge from i to j, which is done during minimal specialization.

Let $\tilde{A}_{k\ell}$ for $1 \leq \ell \leq k$ denote the matrix that has the same entries as $A_{k\ell}$ but now we add an extra row that is a copy of the row corresponding to node i. This represents $\bar{i}$ having the same in-edges as i.

Since $\bar{i}$ has only one out-going edge, which is to node j, the column in $\overline{A}$ corresponding to $\bar{i}$ is all zeros except a 1 in the jth entry.

Thus, the partitioned adjacency matrix of $\overline{A}$ is

$$\overline{A} = \left[\begin{array}{c|c|c|cc} A_{11} & A_{12} & \cdots & A_{1k} & 0 \\ \hline \vdots & \vdots & \vdots & & \vdots \\ \hline A_{(k-1)1} & \cdots & A_{(k-1)(k-1)} & \tilde{A}_{(k-1)k} & e_j \\ \hline \tilde{A}_{k1} & \tilde{A}_{k2} & \tilde{A}_{k(k-1)} & \tilde{A}_{kk} & 0 \end{array}\right] \in \mathbb{R}^{(n+1)\times(n+1)} \tag{7}$$

where e_j is the vector that is all zeros except for a 1 that corresponds to the row associated with j. In this form, it is clear that the upper left $(k-1) \times (k-1)$ block matrix is the as the upper left $(k-1) \times (k-1)$ block matrix of A, so the row sums are constant on each part. Moreover, in the last partition column, it is clear that partition rows 1 to $k-2$ have constant row sum since we are only adding a column of zeros. For the $(k-1), k$ partition of $\overline{A}$, if we are not examining the row corresponding to node j, the row sum is the same as the row sum in $A_{(k-1)k}$. If we are looking at the row corresponding to j, then the row in $\tilde{A}_{(k-1)k}$ has the same entries as $A_{(k-1)k}$ except for the entry corresponding to e_{ij}, which is 0 for $\tilde{A}_{(k-1)k}$ and 1 for $A_{(k-1)k}$. But, we have an additional 1 in the row corresponding to j in $\overline{A}$. Thus the row sums for $\left[\tilde{A}_{(k-1)k} \quad e_j\right]$ are the same as $A_{(k-1)k}$, and thus the row sums are constant. Finally, for the last partition row, since the rows of $\tilde{A}_{k\ell}$ are copies of the rows of $A_{k\ell}$ (some repeated), and we are only adding a column of zeros in the case of $\tilde{A}_{kk}$, the row sums are constant.

Thus, the partitioned $\overline{A}$ has constant row sums on each part, so equivalently, $\overline{\pi}$ is an equitable partition.

We note that if i and j are in the same partition element, then the argument is similar, but instead of having an $\tilde{A}_{(k-1)k}$ and $\tilde{A}_{kk}$, there is just $\tilde{A}_{kk}$ with an extra row that is a copy of the row corresponding to i in A_{kk} and changing the appropriate entry from a 1 to a 0. □

Using Theorem 7.0.1, it follows that the size of a network's equitable partition remains constant under minimal specialization, i.e., the following holds.

Corollary 7.0.1. *Let $\left\{\left(A_m, \mathbb{R}^{n+m}\right)\right\}_{m=0}^{\infty}$ be a sequence of Jackson networks created via repeated minimal specialization. Denote the trivial partition of $(A_0, \mathbb{R}^n)$ by π_0. Then (A_m, R^{n+m}) has an equitable partition π_m where $|\pi_m| = |\pi_0|$ for all $m \geq 0$.*

Proof. This follows from an inductive argument. Our base case is π_0 being the trivial equitable partition of $(A_0, \mathbb{R}^n)$. Thus $|\pi| = |V|$. Now assume inductively that for $(A_{m-1}, \mathbb{R}^{m-1+n})$, we have that $|\pi_{m-1}| = |V|$. Let $(A_m, \mathbb{R}^{m+n})$ be the minimal specialization of $(A_{m-1}, \mathbb{R}^{m-1+n})$. From Theorem 7.0.1, we have that $|\pi_n| = |\pi_{n-1}| = |V|$. □

The number of partition elements remains fixed as a network is specialized, meaning that as the network grows, the partition elements are what grow in size. A related question, not answered here, is as the network grows, how do the sizes of the partition elements grow? For the sake of illustration, we consider this in the following example.

Example 7.0.1. Fig. 7 shows an example of an evolving Jackson network given by the sequence $\left\{\left(A_m, \mathbb{R}^{5+m}\right)\right\}_{m=0}^{3}$. The Jackson networks are sequentially specialized over the vertices $i = 5, 3,$ and 5 again, respectively. Here the original network has the trivial equitable partition $\pi_0 = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$. The result of specializing in this manner results in the partitions

$$\pi_1 = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 5.1\}\}$$
$$\pi_2 = \{\{1\}, \{2\}, \{3, 3.1\}, \{4\}, \{5, 5.1\}\}$$
$$\pi_3 = \{\{1\}, \{2\}, \{3, 3.1\}, \{4\}, \{5, 5.1, 5.2\}\},$$

where each partition element is colored brown, purple, black, green, and red, respectively. Note that $|\pi_0| = |\pi_1| = |\pi_2| = |\pi_3|$.

7.1. Evolution of equitable partitions and comparisons

In Corollary 7.0.1, we proved that minimal specialization either creates a new non-trivial partition element or grows the size of a non-trivial element. Thus, the percentage of non-trivial equitable partition elements never decreases. Here, we explore experimentally the rate at which non-trivial elements are created. The experiments here are the same as those discussed earlier in Section 6, beginning with a directed Erdös-Rényi graph of $|V| = 25$ nodes, $G_{den} = .25$ density, and normalized edge weights for dynamic and random minimal specialization models. We also examined two Barabśi-Albert models, one where we preferentially attach one node with one edge, and the other with one node and two edges. For these models, we begin with an undirected, unweighted Erdös-Rényi graph with $|V| = 25$ nodes, $G_{den} = .25$ density. We iterate each model 500 times, calculating the percentage of non-trivial equitable partition elements at each time step. We repeat this for 100 simulations and average the results.

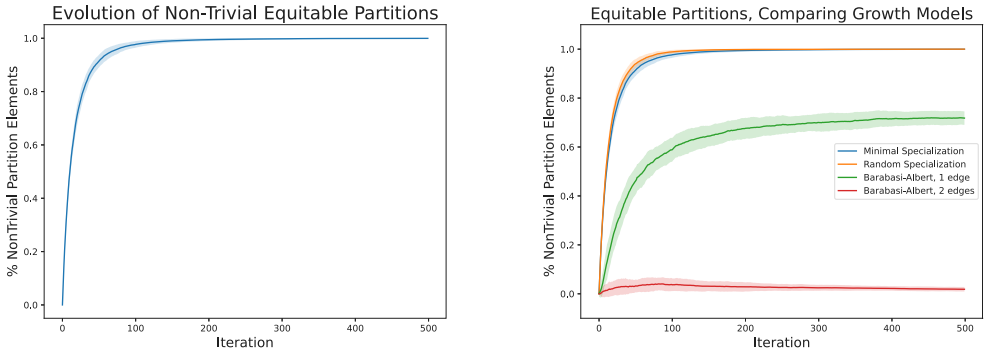
For the integrated specialization model, we achieve 100% non-triviality quite rapidly with little standard deviation, as seen in Fig. 8(a). We see similar results for repeated random minimal specialization (Fig. 8(b)). We can see that, on average, the BA network adding one node and edge evolves to have roughly 70% non-triviality, with higher variance. For BA networks grown by adding one node and two edges, there is no consistent creation of equitable partitions (Fig. 8(b)). (As far as the author's know, there is no theory describing the occurrence of equitable partitions in BA models.) Our model is potentially useful in the sense that it can be used to evolve a network to any percentage of non-trivial partition elements, given that the percentage is in the of the form $\frac{k}{n}$ for $0 \leq k \leq n$ where n is the size of the starting network.

8. Conclusion

In this work, we develop a model that considers the interplay between the dynamics on and the dynamics of the network. We use a modified Jackson network, i.e. stochastic linear dynamics, for the dynamics on the network, and we develop a mechanism called minimal specialization to model the dynamics of the network. We then repeatedly use the dynamics on the network to determine where to minimally specialize the network, which we refer to as the integrated specialization model. We proved that minimal specialization preserves structural and dynamic properties of the network. Specifically, we proved that if we start with a strongly connected or primitive network, minimal specialization will result in a strongly connected or primitive network. We also show how eigenvalues and eigenvectors evolve with minimal specialization.

Numerically, we show that the integrated specialization model creates networks with properties observed in many real-world networks. This includes a right-skewed degree distribution, sparsity, the small-world property, and nontrivial equitable partitions. While other network growth model can produce similar structural features, what is novel about the integrated model is it produces these features through a topological mechanism, i.e. specialization, designed to reduce the maximal asymptotic load of a network. This suggests that locally reducing a networks load may be an important feature in the formation of real-world networks.

Last, we also compare the integrated specialization model to other growth models, showing evidence that the integrated specialization model creates more efficient networks in terms of both dynamics and structure. Specifically, we see that the integrated specialization model creates networks with small diameter, meaning distances across the network are minimized, which is done while simultaneously minimizing maximal load, meaning bottlenecks are reduced.



(a) An example illustrating the evolution of the percentage of non-trivial equitable partition elements using the integrated specialization model. As the network evolves, the percentage of non-trivial equitable partition elements increases. Little standard deviation is observed, with the maximum being .0378 and the minimum being 0 (in every simulation, the final network had 100% non-triviality).

(b) The evolution of the maximum percentage of non-trivial equitable partition elements for the growth models, the integrated specialization model, random minimal specialization, and Barabśi-Albert model (adding one and two edges at a time). 100 simulations of networks grown to 525 nodes are averaged. The integrated specialization model and random specialization perform virtually the same, with little variance in both models. The BA model adding one edge doesn't achieve as high of a percentage, and exhibits more variance between simulations. For the BA model adding two edges, there is no consistent creation of non-trivial equitable partitions.

Fig. 8. Non-trivial equitable partitions.

CRediT authorship contribution statement

Annika King: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation. **Dallas Smith:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Benjamin Webb:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

B. W. was partially supported by the NSF grant #2205837 in Applied Mathematics. A. K. was partially supported by the NSF Graduate Research Fellowship Program under Grant #2238458. 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.

Appendix. Other initial networks

The simulations in Section 5 used directed random (directed Erdös-Rényi) graphs as the initial graphs for the integrated model. Here, we briefly present analogous results using three different initial graphs. The first type of graph is a random geometric graph (RG) with $n = 25$ nodes and a radius of $r = .25$. The second graph type is a Barabśi-Albert graph (BA) with $n = 25$ nodes and $m = 5$ connections. The third graph is a connected Watts–Strogatz graph (WS) with $n = 25$ nodes, $k = 5$ neighbors, and $p = .25$ probability of re-wiring connections. Each of these graphs is initially undirected and unweighted. We then *modify* each graph type by directing then deleting half the edges (rounding down) uniformly assign edge weights $W(e_{ij}) \in [1, 100]$ then normalize out-going edges to sum to 1. We check that the resulting graph is strongly connected.

Analogous to the previous simulations on Erdös-Rényi graphs we run one-hundred simulations for each type of graph, growing the networks from 25 to 525 nodes. The same statistics as before are averaged and presented here (see Figs. 3 and 5). For these experiments, Fig. 9 shows how the out-degree distribution evolves using the integrated model with the three modified initial graphs. In each case the out-degree distribution becomes increasingly right-skewed under the integrated model, on average.

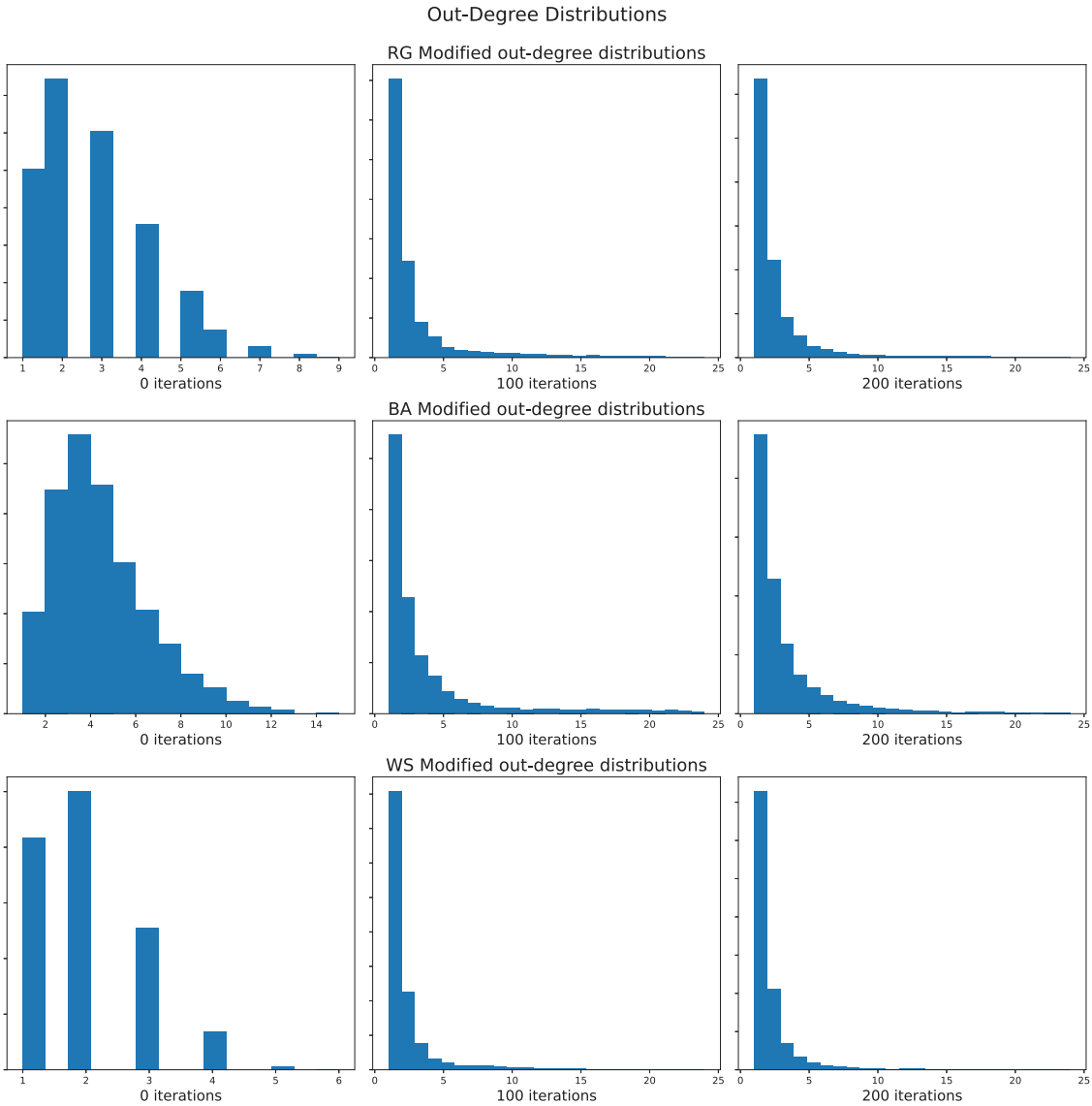


Fig. 9. The evolution of the out-degree distribution using the integrated model is shown for the modified random geometric (RG), Barabśi-Albert (BA), and Watts–Strogatz (WS) graphs. These are shown in the top, middle and bottom rows, respectively, for 0, 100, and 200 iterations, with the distribution tails truncated.

Fig. 10 shows how the diameter evolves using the integrated model starting with the modified RG, BA, and WS graphs. Each experiment exhibits a different type of logarithmic growth, suggesting that for each has a distinct small-world property.

Fig. 11 shows how the density evolves using the integrated model starting with the modified RG, BA, and WS graphs. Regardless of the initial graph, the graph becomes increasingly sparse as the network is repeatedly specialized where the density appears to be approaching zero.

Fig. 12 shows how the maximum eigenvector centrality of a network evolves using the integrated model starting with the modified RG, BA, and WS graphs. As with network density, regardless of the initial graph the maximal centrality decreases as the network is repeatedly specialized where the maximum in each case appears to be approaching zero.

Each of the figures in this and in Section 5 provide evidence that the integrated model creates networks with some of most studied properties observed in many real-world networks including (i) right-skewed degree distributions, (ii) sparsity, and (iii) the small-world property. While the specifics of (i)-(iii) depend on the initial graph used in the integrated model, the overall pattern appears to be independent of the initial condition. Additionally, for each type of graph, the integrated model deceases the maximal load on the network as the network is specialized.

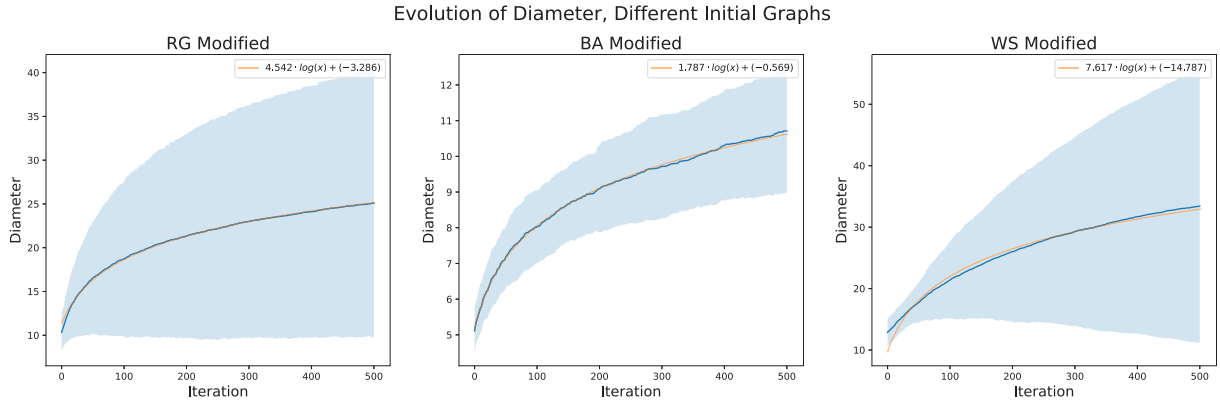


Fig. 10. The evolution of network diameter using the integrated model is shown for the modified random geometric (RG), Barabśi-Albert (BA), and Watts–Strogatz (WS) graphs, from left to right respectively. A line of best fit is plotted in orange in each plot, and one standard deviation is shown.

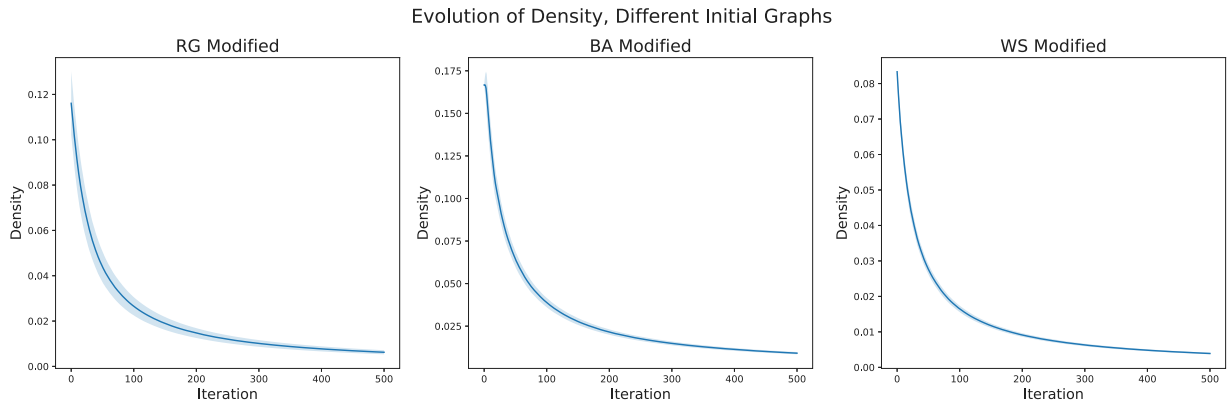


Fig. 11. The evolution of network density using the integrated model is shown for the modified random geometric (RG), Barabśi-Albert (BA), and Watts–Strogatz (WS) graphs, from left to right respectively. The average and one standard deviation is shown.

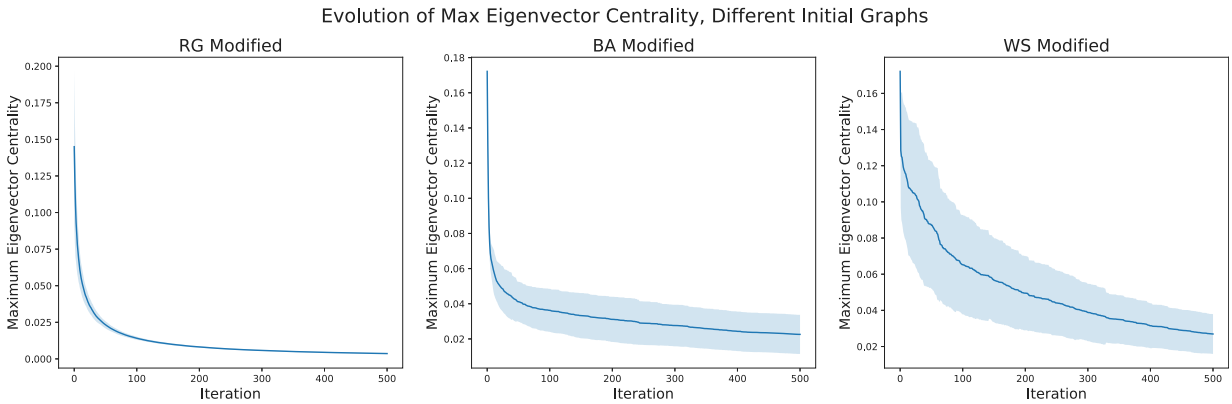


Fig. 12. The evolution of maximum eigenvector centrality using the integrated model is shown for the modified random geometric (RG), Barabśi-Albert (BA), and Watts–Strogatz (WS) graphs, from left to right respectively. The average and one standard deviation is shown.

References

[1] Alain Barrat, Marc Barthelemy, Alessandro Vespignani, Dynamical Processes on Complex Networks, Cambridge University Press, 2008.

[2] Toru Ohira, Ryusuke Sawatari, Phase transition in a computer network traffic model, Phys. Rev. E 58 (1998) 193–195.

[3] Shawn Martin, Zhaoduo Zhang, Anthony Martino, Jean-Loup Faulon, Boolean dynamics of genetic regulatory networks inferred from microarray time series data, Bioinformatics 23 (7) (2007) 866–874.

[4] Karthik Raman, Nagasuma Chandra, Flux balance analysis of biological systems: applications and challenges, Brief. Bioinform. 10 (4) (2009) 435–449.
[5] Leon Glass, Synchronization and rhythmic processes in physiology, Nature 410 (6825) (2001) 277–284.
[6] Ed Bullmore, Olaf Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems, Nat. Rev. Neurosci. 10 (3) (2009) 186–198.
[7] Albert-Laszlo Barabasi, Zoltan N. Oltvai, Network biology: understanding the cell's functional organization, Nature Rev. Genet. 5 (2) (2004) 101–113.
[8] Gerald R. Ash, Dynamic network evolution, with examples from AT&T's evolving dynamic network, IEEE Commun. Mag. 33 (7) (1995) 26–39.
[9] Walter E Beyeler, Robert J Glass, Morten L Bech, Kimmo Soramäki, Congestion and cascades in payment systems, Phys. A 384 (2) (2007) 693–718.
[10] James R. Jackson, Networks of waiting lines, Oper. Res. 5 (4) (1957) 518–521.
[11] Peter Müller, Handbook of Dynamics and Probability, Springer, 2022.
[12] Olaf Sporns, Network attributes for segregation and integration in the human brain, Curr. Opin. Neurobiol. 23 (2) (2013) 162–171.
[13] Ricard V. Solé, Sergi Valverde, Spontaneous emergence of modularity in cellular networks, J. R. Soc. Interface 5 (18) (2008) 129–133.
[14] Carlos Espinosa-Soto, Andreas Wagner, Specialization can drive the evolution of modularity, PLoS Comput. Biol. 6 (3) (2010) e1000719.
[15] Ronald J. Salz, David K. Loomis, Kelly L. Finn, Development and validation of a specialization index and testing of specialization theory, Hum. Dimens. Wildl. 6 (4) (2001) 239–258.
[16] Guillaume Burghouwt, Long-haul specialization patterns in European multihub airline networks–An exploratory analysis, J. Air Transp. Manage. 34 (2014) 30–41.
[17] L. Bunimovich, D. Smith, B. Webb, Specialization models of network growth, J. Complex Netw. 7 (2017).
[18] Leonid Bunimovich, DJ Passey, Dallas Smith, Benjamin Webb, Spectral and dynamic consequences of network specialization, Int. J. Bifurcation Chaos 30 (06) (2020) 2050091.
[19] Erik Hannesson, Jordan Sellers, Ethan Walker, Benjamin Webb, Network specialization: A topological mechanism for the emergence of cluster synchronization, Phys. A 600 (2022) 127496.
[20] Jörg Stelling, Steffen Klamt, Katja Bettenbrock, Stefan Schuster, Ernst Dieter Gilles, Metabolic network structure determines key aspects of functionality and regulation, Nature 420 (6912) (2002) 190–193.
[21] Annica Sandström, Lars Carlsson, The performance of policy networks: the relation between network structure and network performance, Policy Stud. J. 36 (4) (2008) 497–524.
[22] Ilse Lindenlaub, Anja Prummer, Network structure and performance, Econ. J. 131 (634) (2020) 851–898.
[23] Roger A. Horn, Charles R. Johnson, Matrix Analysis, Cambridge University Press, 2012.
[24] Mark Newman, Networks, Oxford University Press, 2018.
[25] Richard Durrett, Random Graph Dynamics, Citeseer, 2007.
[26] Qinghua Chen, Shenghui Chen, A highly clustered scale-free network evolved by random walking, Phys. A 383 (2) (2007) 773–781.
[27] Chuankui Yan, Network model with scale-free, high clustering coefficients, and small-world properties, J. Appl. Math. 2023 (2023) 1–11.
[28] Zhe-Ming Lu, Yu-Xin Su, Shi-Ze Guo, Deterministic scale-free small-world networks of arbitrary order, Phys. A 392 (17) (2013) 3555–3562.
[29] Thilo Gross, Hiroki Sayama, Adaptive Networks Theory, Models and Applications, Springer Berlin, 2009.
[30] Ben D. MacArthur, Rubén J. Sánchez-García, James W. Anderson, Symmetry in complex networks, Discrete Appl. Math. 156 (18) (2008) 3525–3531.