Supernetworks and Systemic Risk

Frank Page\textsuperscript{1}
Systemic Risk Centre
London School of Economics
London WC2A 2AE
UK
fpage.supernetworks@gmail.com

Jing Fu
Department of System Management
Fukuoka Institute of Technology
3-30-1 Wajiro-higashi, Higashi-ku, Fukuoka, 811-0295
JAPAN
Joanna.Fuu@gmail.com

March 20, 2019\textsuperscript{2}

\textsuperscript{1}Visiting Professor and Co-Investigator, Systemic Risk Centre, London School of Economics, London WC2A 2AE, UK. Also, Centre d’Economie de la Sorbonne, Université Paris 1 (Panthéon-Sorbonne). Permanent address: Indiana University, Bloomington, IN 47405, USA

\textsuperscript{2}The authors are grateful to Ann Law, J. P. Zigrand and Jon Danielsson of LSE for all their support and hospitality during several visits to the SRC. The first author acknowledges financial support from the Systemic Risk Centre at LSE under ESRC grant number ES/K002309/1.
Abstract

We give a formal definition of systemic risk that is firmly grounded in network dynamics and using our definition we show that the dynamics driving network formation generate a finite set of basins of attraction. The presence of basins of attraction has profound implications for our understanding of systemic risk. The main implications for systemic risk are the following: First, departing from any initial network, the stochastic process of network formation reaches a network contained in one of the basins in finite time with probability 1 - and once there, the process remains in the basin, visiting each node in the basin infinitely often. Second, we show that each basin is homogeneous with respect to its network failure characteristics (i.e., if a basin contains a network with a particular set of failed nodes, then all networks in the basin have the same set of failed nodes). Third, using our definition of systemic risk, we formally define and characterize the notion of a tipping point network. We then show that the profile of basins comes equipped with a unique set of tipping point networks. Each tipping point network is the gateway to some sequence of future networks leading inexorably to some basin, and depending on the failure characteristics of this basin, this sequence might best be described as a failure cascade. Thus, tipping point networks are the supernetwork’s early warning system for network failure. Finally, we define and characterize the notions of systemic and killer nodes. The big picture take away from our approach to systemic risk is that what is critical in assessing systemic risk is the allocation of failed nodes across the basins of attraction.

Key Words: endogenous systemic risk, equilibrium supernetworks, basins of attraction, tipping points, spheres of influence, systemic nodes, killer nodes, first passage probabilities, hitting probabilities, stationary Markov processes, discounted stochastic games

JEL classifications: C73, C78, D85, G21
1 Introduction

We give a formal definition of systemic risk that is firmly grounded in network dynamics. Underlying our definition is the unique Markov supernetwork representation of the stochastic dynamics driving network formation. In a Markov supernetwork the nodes are networks and the probability-weighted directed arcs pointing from one network (or node) to another represent the Markov transition probabilities governing the process. Viewing the supernetwork as a map of the transportation system over which the process of network formation will travel, we are naturally lead to define systemic risk as the probability that the stochastic process of network formation, starting at a given network (i.e., at a given node in the supernetwork), will arrive at a failed network (i.e., another node in the supernetwork), at or before a given time. In classical terminology, we define systemic risk as the first passage probability to a failed network from a given network. Thus, our notion of systemic risk is one inextricably linked to the underlying process of network formation as represented by the supernetwork. Following our approach, rather than there being a single measure of systemic risk, there is instead a schedule of systemic risk measures which lists the probabilities of various arrival times at various failed networks in the supernetwork, departing from any given network in the supernetwork. It is the structure and stochastic properties of this transportation system which determine systemic risk.

Our approach to systemic risk allows us to bring into the analysis - in a very direct and simple manner - the strategic underpinnings of systemic risk (e.g., see Gong, Page, and Zigrand, 2015). For example, if we model network formation as a discounted stochastic game, then the players’ Nash equilibrium in stationary Markov strategies of network formation give rise to an equilibrium stationary Markov process of network formation - and therefore, to an equilibrium Markov supernetwork. Using the structure and properties of the equilibrium supernetwork, we are able to compute the schedule of endogenous systemic risks - that is, the schedule of systemic risks that arises from the dynamic, strategic behavior of the players in forming networks. Moreover, because such a game-theoretic approach lays out and makes visible the direct feedback connections that exist between strategic behavior in forming networks and the dynamic risks of network failure, our approach opens up the possibility of designing decentralized incentive mechanisms that induce players in the game of network formation to minimize systemic risks endogenously. If the networks we are analyzing are banking networks, for example, then our approach to systemic risk - by allowing strategic behavior to enter the analysis - is useful in formulating policies which incentivize risk-minimizing behavior in banking interactions.

Our approach to systemic risk, via Markov supernetworks and first passage probabilities, has several useful properties which allow us to greatly refine our understanding of systemic risk. First, even in the case where the feasible set of networks is uncountable, under mild regularity conditions on the supernetwork’s transition probabilities, the Markov supernetwork will have a unique, finite profile of basins of attractions. More importantly, with probability 1, the network formation process moving through the supernetwork will reach, in finite time, one of these finitely many basins of attraction and once there, will stay there (e.g., see Tweedie, 2001, and
The presence of basins of attraction has profound implications for our understanding of the dynamic behavior of systemic risk, and therefore, for our ability to measure and control systemic risk. This is true for two reasons. First, as we will show here, the basins of attraction generated by the network formation process are homogenous with respect to their network failure characteristics. In particular, if a basin contains a network having a particular nonempty set of failed nodes (failed banks in the case of banking networks), then all networks in the basin will have precisely the same set of failed nodes - no more, no less - and these nodes can be identified and computed. The stratification of basins by network failure characteristics is very useful because it allows us to describe in a very precise way the severity of the failure level exhibited by a particular basin. Thus, a supernetwork's unique finite collection of basins together with its unique finite profile of sets of failed nodes across basins is the systemic risk signature of the underlying stochastic process of network formation. Second, given the structure of a particular supernetwork, the corresponding profile of basins comes equipped with a unique set of tipping points. Here, based on our definition of systemic risk, we will formally define the notion of a tipping point network. Each tipping point network is the gateway to some sequence of future networks leading inexorably to some basin, and depending on the failure characteristics of this basin, this sequence might best be described as a failure cascade. Thus, tipping points are the supernetwork's early warning system for network failure. A tipping point network in a supernetwork that is a point of departure for the process's journey to a severely failed basin is truly a systemic network - and such systemic networks can easily be identified.

The big picture take away from our approach to systemic risk is that what is critical in assessing systemic risk is the allocation of failure levels across the basins of attraction. For example, if each basin of attraction contains failed networks (i.e., a network with a nonempty subset of failed nodes), then the long run fate of all networks is some level of failure. If some basin contains networks with all failed nodes, then the tipping point for this basin is the gateway to total network failure. If there is only one basin of attraction and if this basin contains totally failed networks, then essentially, we are looking at the gambler's ruin problem for networks (i.e., the network ruin problem). Alternatively, if there is only one basin and if this basin contains only partially failed networks (i.e., networks with a fixed set of failed nodes smaller than the full set of nodes), then we are looking at a network formation process which endogenously limits the severity of network failure. Can the formation of such a self-limiting network be incentivized by smartly designed policies?

Under our approach, the Markov supernetwork representation of network dynam-
ics allows us to see the current position of the network in the supernetwork relative to the “bad” tipping points, “bad” failure cascades, and “bad” basins of attraction. Thus, our approach provides a network visualization device for viewing the qualitative properties of network dynamics and therefore for visualizing and computing systemic risk schedules. Moreover, using the supernetwork representation of network dynamics as the connection point between strategic behavior and systemic risk, our approach provides a paradigm for studying the feedback effects between network structure, strategic behavior and risk - providing therefore a method to carry out network formation policy studies (see Gong and Page, 2016). Here we will focus on providing formal definitions - as well as characterizations - of systemic risk and tipping points. For a game theoretic analysis of the strategic underpinnings of systemic risk see Gong, Page, and Zigrand (2016).

2 Networks

2.1 Primitives

Connections are the basic building blocks of networks. Here we will focus on networks made up of directed connections. A directed connection consists of two nodes connected by a labeled arc (i.e., a labeled arrow pointing from one node to the other) indicating the connection type. If we were to draw a picture of a directed connection, it would look like the following:

![Figure 1: A directed connection](image)

Written out long hand, the directed connection depicted in Figure 1 is given by \((a, (i, j))\), indicating that nodes \(i\) and \(j\) are connected by an arc, \(a\), running from node \(i\) to node \(j\). Formally, the basic ingredients making up a directed connection are,

\[
\begin{align*}
\text{(Nodes and Arcs)} \\
N &= \text{a finite set of nodes containing } n \text{ nodes, with typical elements } i \text{ and } j, \\
A &= \text{a finite set of arcs containing } m \text{ arcs, with typical element } a.
\end{align*}
\]

Let \(N \times N := N^2\) denote the collection of all node pairs, \((i, j)\), or pre-connections, and let \(P(N^2)\) denote the collection of all nonempty subsets of \(N \times N\), with typical element \(g\) called a pre-network. In order for a pre-connection to become a connection, it must be assigned a connection type (i.e., it must be assigned an arc). The set of arcs, \(A\), provides us with the set of connection types. Given pre-network, \(g \in P(N^2)\), let \(g(i)\) be \(j\)-section of \(g\) at \(i\) given by

\[g(i) := \{ j \in N : ij \in g \}.\]
For pre-network \( g \in P(N^2) \), we define the domain of \( g \), denoted by \( D(g) \), to be the nodes \( i \) in \( N \) having nonempty \( j \)-sections in \( N \). Thus,

\[
D(g) := \{ i \in N : g(i) \neq \emptyset \}.
\]

Finally, let \( P^d(N^2) \subset P(N^2) \) be the subcollection of all pre-networks consisting only of diagonal pre-connections. Thus,

\[
P^d(N^2) := \{ g \in P(N^2) : \text{for all } i \in D(g), g(i) = i \}.
\]

**Definition 1 (Connections)**

Given node set \( N \) arc set \( A \) a connection is an ordered pair \( (a, (i, j)) \in A \times (N \times N) \) consisting of an arc type \( a \in A \) and a pre-connection, \( (i, j) \in N \times N \).

The set of all directed connections is given by the Cartesian product

\[
K := A \times (N \times N).
\]

A directed network is defined as follows:

**Definition 2 (Networks)**

Given node set, \( N \), arc set, \( A \), a directed network, \( G \), is a nonempty subset of the set directed connections, \( K \). The set of all directed networks is given by \( P(K) \), the collection of all nonempty subsets of \( K \).

Under our definition of a directed network, loops are allowed - a loop being a connection where an arc goes from a given node back to that given node.\(^3\) Also, under our definition an arc can be used multiple times in a given network and multiple arcs (even uncountably many) can go from one node to another. However, under our definition no arc \( a \) is allowed to go from a node \( i \) to a node \( j \) multiple times.

The following notation is useful in describing networks. Given directed network \( G \in P(K) \), let

\[
\begin{align*}
G(a) & := \{(i, j) \in N \times N : (a, (i, j)) \in G\}, \\
G(ij) & := \{a \in A : (a, (i, j)) \in G\}.
\end{align*}
\]

In network \( G \),

\[
G(a) \text{ is the set of node pairs connected by arc } a, \quad \text{and} \quad G(ij) \text{ is the set of arcs from node } i \text{ to node } j.
\]

\(^3\)By allowing loops we are able to represent a network having no connections between distinct nodes as a network consisting entirely of loops at each node.
The set \( \mathcal{G}(a) \) is called the section of network \( \mathcal{G} \) at arc \( a \), while \( \mathcal{G}(ij) \) is called the section of network \( \mathcal{G} \) at pre-connection \( (i, j) \). If for some arc \( a \in A \), \( \mathcal{G}(a) \) is empty, then arc \( a \) is not used in network \( \mathcal{G} \). Also, if for some node \( i \in N \), \( \mathcal{G}(ij) \) and \( \mathcal{G}(ji) \) are empty for all \( j \neq i \), then node \( i \) is isolated.

Corresponding to a network \( \mathcal{G} \in \mathcal{P}(K) \), there is an arc mapping, \( ij \rightarrow \mathcal{G}(ij) \), from pre-connections into subsets of arcs. Letting \( 2^A \) denote the collection of all \( \text{nonempty} \) subsets of \( A \), we have \( \mathcal{G}(\cdot) : N^2 \rightarrow 2^A \).

The domain of network \( \mathcal{G} \) is given by

\[
\mathcal{D}(\mathcal{G}) := \{ (i, j) \in N^2 : \mathcal{G}(ij) \neq \emptyset \}.
\]  

(3)

Note that the arc mapping, \( \mathcal{G}(\cdot) \), induced by network \( \mathcal{G} \in \mathcal{P}(K) \), can also be viewed as a mapping from \( \mathcal{G} \) domain, \( \mathcal{D}(\mathcal{G}) \), into \( \mathcal{P}(A) \). Also note that given any pre-network \( g \in \mathcal{P}(N^2) \), the set-valued mapping,

\[
\mathcal{G}(\cdot) : g \rightarrow \mathcal{P}(A),
\]

has a graph,

\[
\text{Gr} \mathcal{G}(\cdot) := \{ (a, (i, j)) \in K : a \in \mathcal{G}(ij) \},
\]

contained in \( \mathcal{P}(K) \). Thus a network induces an arc mapping and an arc mapping induces a network.

Given network \( \mathcal{G} \in \mathcal{P}(K) \), with arc mapping, \( ij \rightarrow \mathcal{G}(ij) \), there is also an induced cardinality mapping, \( ij \rightarrow |\mathcal{G}(ij)| \), where \( |\mathcal{G}(ij)| \) denotes the cardinality of the set of arcs, \( \mathcal{G}(ij) \) (we adopt the convention that \( |\mathcal{G}(ij)| = 0 \) if and only if \( \mathcal{G}(ij) = \emptyset \)).

Finally, to complete our description of the primitives, we will assume that associated with the set of nodes, \( N \), and the set of arcs, \( A \), there is a feasible arc correspondence, \( ij \rightarrow A(ij) \), defined on the set of pre-connections, \( N^2 \) taking nonempty values in \( A \). Thus, \( A(\cdot) : N^2 \rightarrow \mathcal{P}(A) \).

### 2.2 Failed Networks

In order to define and analyze systemic risk, we must define what we mean by a failing and a failed network.

\[ \text{A-2} \] (Failed Networks)

We will assume that the set of arcs is given by \( A := A_c \cup A_f \), where \( A_c \) and \( A_f \) are disjoint and where the arcs in \( A_f \) are called failure arcs, while the arcs in \( A_c \) are called continuation arcs. Also, we will assume that the feasible set of networks, \( \mathcal{G} \subset \mathcal{P}(K) \), is such that for each network \( \mathcal{G} \in \mathcal{G} \), with induced arc mapping, \( ij \rightarrow \mathcal{G}(ij) \), the following conditions are satisfied: \( \mathcal{G}(ij) \subset A_c \) for all pre-connections \( ij \in \mathcal{D}(\mathcal{G}) \) with \( i \neq j \), and for all diagonal pre-connections, \( ii \in \mathcal{D}(\mathcal{G}) \), either \( \mathcal{G}(ii) \subset A_c \) or \( \mathcal{G}(ii) \subset A_f \). Finally, we will assume that the feasible set, \( \mathcal{G} \), has the following properties:
(1) If $G \in \mathbb{G}$ is such that for some diagonal pre-connection, $ii \in D(G)$, $G(ii) \subset A_f$, then $G$ is said to be a failed network.

(2) If $G \in \mathbb{G}$ is a failed network with $G(ii) \subset A_f$ for some $ii \in D(G)$, then for all pre-connections, $ij$ and $ji$ with $j \neq i$, $G(ij) = \emptyset = G(ji)$.

Let

$$D_f(G) := \cup_{a \in A_f} G(a) \subset P^i(N^2),$$

and

$$D_c(G) := \cup_{a \in A_c} G(a) \subset P(N^2).$$

where recall $P^i(N^2)$ is the collection of nonempty subsets of the set of diagonal pairs, $ii \in N^2$ given by

$$P^i(N^2) := \{g \in P(N^2) : \text{for all } i \in D(g), \ g(i) = i\}.$$

We will call the set of diagonal pre-connections, $D_f(G)$, the failure domain, and we will call the set of pre-connections, $D_c(G)$, the continuation domain. Note that under $[A-2](2)$, $D_c(G) = D(G) \setminus D_f(G)$. Finally, define

$$N_f(G) := \{i \in N : ii \in D_f(G)\},$$

and

$$N_c(G) := \{i \in N : ii \notin D_f(G)\}.$$  

Under $[A-2]$, a network $G \in \mathbb{G}$ is a failed network if and only if $N_f(G) \neq \emptyset$. We will denote by $\mathbb{F}$ the subset of all failed networks in $\mathbb{G}$.

**Definition 3 (k-Level Networks)**

We say that network $G \in \mathbb{G}$ is a $k$-level network if $|N_f(G)| = k$, for $k = 0, 1, 2, \ldots, n$.

Thus, an 0-level network $G$ has no failed nodes. For $1 \leq k \leq n$, a $k$-level network, $G$, is a network with $k$ failed nodes (i.e., $|N_f(G)| = k$). Initially we will measure the severity of the failure by simply counting the number of failed nodes - the higher the $k$ to more severe is the failure. Obviously, not all $k$-level failures are equal - the severity of the failure in the long run depends on which nodes fail now. We will refine our failure measure once we have in place the notions of tipping points and basins of attraction.

By definition, the set of failed networks $\mathbb{F}$ is given by,

$$\mathbb{F} := \{G \in \mathbb{G} : N_f(G) \neq \emptyset\} := \cup_{i=1}^{n} \{G \in \mathbb{G} : |N_f(G)| = k\}. $$

We will denote by $\mathbb{F}^k$, $k = 1, 2, \ldots, n$, the set of failed networks consisting of $k$-level networks. While the set of safe networks, $\mathbb{S}$, is given by,

$$\mathbb{S} := \mathbb{G} \setminus \mathbb{F} := \{G \in \mathbb{G} : N_f(G) = \emptyset\}. $$
The state space of networks, \( \mathcal{G} \), can be partitioned into safe and failed networks - and the space of failed networks can be further partitioned into failure classes according to the level of failure. Thus, we have

\[
\mathcal{G} := \mathcal{S} \cup \mathcal{F} := \mathcal{S} \cup \mathcal{F}^1 \cup \mathcal{F}^2 \cup \ldots \cup \mathcal{F}^m. \tag{9}
\]

In order to complete our formal description of failure, we must state our assumptions concerning the behavior of the law of motion with respect to failed networks. This we will do below when we introduce the law of motion.

### 3 Stationary Markov Transition Probabilities

Assume that the stochastic process of network formation, \( \{G^n\}_n \), is governed by a Markov transition probability function given by

\[
G \rightarrow p(\cdot \mid G) \in \Delta(\mathcal{G}), \tag{10}
\]

defined on the feasible set of networks, \( \mathcal{G} \), taking values is the set of probability measures, \( \Delta(\mathcal{G}) \), defined on \( \mathcal{G} \). Thus, for each subset of feasible networks, \( \mathcal{E} \subset \mathcal{G} \), if at time \( t \) the network formation process is visiting network \( G \in \mathcal{G} \), then the process will visit some network in \( \mathcal{E} \) at time \( t + 1 \) with probability

\[
p(\mathcal{E} \mid G) := \sum_{G' \in \mathcal{E}} p(G' \mid G) \in [0, 1], \tag{11}
\]

and because the process is stationary Markov, expression (11) holds for all \( t \).

In order to further simplify the notation, let

\[
\mathcal{G} := \{G_1, G_2, \ldots, G_H\}, \tag{12}
\]

and denote states by \( l \) and \( h \), where \( l \) and \( h \) are elements of

\[
\mathcal{H} := \{1, 2, \ldots, H\}. \tag{13}
\]

We will let \( \mathcal{F} \subset \mathcal{H} \) be the subset of indices corresponding to failed networks and \( \mathcal{S} \subset \mathcal{H} \) the subset of indices corresponding to safe networks. Thus, \( l \in \mathcal{H} \) if and only if \( G_l \in \mathcal{G} \), and for any nonempty subset \( \mathcal{E} \) of \( \mathcal{H} \), \( l \in \mathcal{E} \) if and only if \( G_l \in \mathcal{G}_\mathcal{E} \subset \mathcal{G} \). Thus, \( G_l \in \mathcal{G}_\mathcal{F} := \mathcal{F} \) is a failed network if and only if \( l \in \mathcal{F} \), \( G_l \in \mathcal{G}_\mathcal{F^k} := \mathcal{F}^k \) is a \( k \)-level network if and only if \( l \in \mathcal{F}^k \), and \( G_l \in \mathcal{G}_\mathcal{S} := \mathcal{S} \) is a safe network if and only if \( l \in \mathcal{S} \).

Given current state \( G_l \in \mathcal{G} \), the conditional probability with which nature chooses network \( G_h \in \mathcal{G} \) is given by \( p(G_h \mid G_l) \). To save writing, we will denote transition probabilities, \( p(G_h \mid G_l) \), by

\[
p(G_h \mid G_l) := p_{lh}. \tag{14}
\]
We can then write down all the transition probabilities in matrix form as an \( H \times H \) Markov transition matrix given by

\[
P := \begin{pmatrix}
p_{11} & \cdots & p_{1h} & \cdots & p_{1H} \\
\vdots & & \vdots & & \vdots \\
p_{h1} & \cdots & p_{hh} & \cdots & p_{hH} \\
\vdots & & \vdots & & \vdots \\
p_{H1} & \cdots & p_{HH} & \cdots & p_{HH}
\end{pmatrix}.
\]

We will assume that once a node fails, it stays failed.

**[A-3]** (The persistence of failed networks)

We will assume that the Markov transition probabilities are such that if the current network, \( G_\ell \in \mathcal{G}_\mathcal{F} \), is failed with failed node set, \( N_f(G_\ell) \), then with probability 1, the coming network, \( G_{\ell+1} \), as well as all future networks, \( G_{\ell+r} \), beyond \( G_{\ell+1} \) will be such that

\[
N_f(G_\ell) \subset N_f(G_{\ell+1}) \subset N_f(G_{\ell+r}).
\]  

(15)

Thus, once a failed network appears (i.e., once a network \( G_\ell \) with \( N_f(G_\ell) \neq \emptyset \) is reached), even one with only one failed node, with probability one, all future networks, \( G_{\ell+r} \), generated by the Markov transition, \( p(\cdot|\cdot) \), will be failed networks with each future network, \( G_{\ell+r} \), in succession having a nondecreasing number of failed nodes - the result of a sequence of nondecreasing nested sets of failed nodes. In fact, it is possible that in finite time with positive probability a network will be reached having a failed network with all failed nodes.

Note that the \( l^{th} \) row of the Markov transition matrix, \( P \), is given by

\[
p^l = (p_{l1}, p_{l2}, \ldots, p_{lH}) := (P)^l,
\]  

while the \( h^{th} \) column of \( P \) is given by \( p_h = \begin{pmatrix} p_{1h} \\
p_{2h} \\
\vdots \\
p_{Hh} \end{pmatrix} := (P)_h. \)

Note that \( p^l \) is a conditional probability measure on the state space \( \mathcal{G} \), that is, \( p^l \in \Delta(\mathcal{G}) \), where in this case the set of probability measures \( \Delta(\mathcal{G}) \) is given by

\[
\Delta(\mathcal{G}) = \left\{ p = (p_1, \ldots, p_H) \in \mathbb{R}^H : p_l \geq 0 \text{ and } \sum_{l=1}^H p_l = 1 \right\}.
\]  

(17)

Given initial probability measure \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_H) \in \Delta(\mathcal{G}) \) prescribing the probability with which the initial network \( G_{l_0} \) is chosen, the probability that the process reaches network \( G_{l_n} \) after \( \tau \) moves is given by

\[
(\gamma P^\tau)_{l_n} = \sum_{l_0=1}^H \gamma_{l_0} (P^\tau)_{l_0 l_n},
\]
where

\[ P^\tau \] is the matrix obtained by multiplying \( P \) by itself \( \tau \) times,

\[(\gamma P^\tau)_{t,\tau} := \langle \gamma, (P^\tau)_{t,\tau} \rangle \] is the \( t^{th} \) component of the row vector \( \gamma P^\tau \),

and

\[(P^\tau)_{l_0,\tau} \] is the \((l_0, l_\tau) \in \mathcal{H} \times \mathcal{H}\) entry of the matrix \( P^\tau \).

4 Systemic Risk

Corresponding to the Markov transition matrix \( P \) there is a unique directed network \( \mathbf{M} \) - a supernetwork (see Page, Wooders, and Kamat, 2005, and Page and Wooders, 2009a and b) - where

\[ \mathbf{M} \subset [0, 1] \times (\mathbb{G} \times \mathbb{G}), \]

with typical connection \((p_{lh}, (G_l, G_h))\) where \( p_{lh} \in [0, 1] \) is the probability that the process of network formation departing network \( G_l \) at time point \( t \) arrives at network \( G_h \) at time point \( t + 1 \) (i.e., \( p_{lh} \) is the \( lh^{th} \) entry in the Markov transition matrix \( P \)), where networks \( G_l \) and \( G_h \) are contained in the state space \( \mathbb{G} \). The connection

\[(p_{lh}, (G_l, G_h)) \in \mathbf{M} \]

is active if and only if the process of network governed by Markov transition \( P \) is such that for all time points \( t = 1, 2, \ldots, \)

\[ \Pi\{ \tilde{G}^t = G_h | \tilde{G}^{t-1} = G_l \} = p_{lh} > 0.5 \]

\[ ^4\text{We will adopt the following notation: for the } \mathcal{H} \times \mathcal{H}, \text{ Markov transition matrix, } P, \text{ let} \]

\[ p^k := (P)^k = k^{th} \text{ row} \]

\[ p_h := (P)_h = h^{th} \text{ column} \]

\[ p_{kh} := (P)_{kh} = kh^{th} \text{ entry.} \]

For the row vector \( \gamma P \) gotten by pre-multiplying the matrix \( P \) by the (row) probability vector \( \gamma \),

\[(\gamma P)_k = k^{th} \text{ component of the vector} \]

\[ (\langle \gamma, (P)_1 \rangle, \ldots, \langle \gamma, (P)_H \rangle). \]

\[ ^5\text{Recall, } \Pi \text{ is the probability measure defined on the measurable space } (\Theta, \mathcal{F}) \text{ underlying the Markov process of network formation. Thus, we are assuming that the Markov process } \{\tilde{G}^n\} \text{ is a sequence of network-valued, } \mathcal{F}-\text{measurable functions} \]

\[ \tilde{G}^n : \Theta \rightarrow \mathbb{G}, \]

such that,

\[ \Pi\{ \tilde{G}^t \in \mathbb{E} | \tilde{G}^0 = G^0, \tilde{G}^1 = G^1, \ldots, \tilde{G}^{t-1} = G^{t-1} \} \]

\[ = \Pi\{ \tilde{G}^t \in \mathbb{E} | \tilde{G}^{t-1} = G^{t-1} \} \]

\[ = \Pi G^{t-1} (\tilde{G}^t \in \mathbb{E}). \]
4.1 A Definition of Systemic Risk

We will denote by \( f_{th}^\tau \) the conditional probability that the first passage from network \( G_t \) to network \( G_h \) occurs exactly at time point \( \tau \) in Markov supernetwork \( M \). Thus, for any two network indices \( l \) and \( h \) in \( \mathcal{H} \) and any (time point) integer \( \tau = 1, 2, \ldots \),

\[
 f_{th}^\tau := \Pi \{ V_h(\tau) | \tilde{G}^0 = G_l \},
\]

where

\[
 V_h(\tau) := \{ \tilde{G}^\tau = G_h, \tilde{G}^t \neq G_h \text{ for } t = 1, \ldots, \tau - 1 \}.
\]

The underlying set of states, \( V_h(\tau) \), is the event that the first time among the times \( t = 1, 2, \ldots \), at which the process of network formation, \( \{ \tilde{G}^t \}_t \), visits network \( G_h \) is at time \( \tau \). This event is then conditioned by the underlying set of states \( \{ \tilde{G}^0 = G_l \} \) to obtain the first passage probability.\(^6\)

4.1.1 Systemic Risk: Networks and Sets of Networks

We begin with our definition.

**Definition 4 (Systemic Risk, \( SR_\tau(G_t, G_h) \))**

In the Markov supernetwork, \( M \), the systemic risk of the network formation process, \( \{ \tilde{G}^t \}_t \), departing network \( G_t \in \mathcal{G} \) and arriving at failsd network \( G_h \in \mathcal{F} \) at exactly time \( t = \tau \) (i.e., in exactly in \( \tau \) steps) denoted by \( SR_\tau(G_t, G_h) \), is given by

\[
 SR_\tau(G_t, G_h) := f_{th}^\tau.
\]

Given any subset of failed networks, \( \mathcal{F} \subset \mathcal{G} \), the systemic risk of leaving network \( G_t \) and arriving at network \( G_h \) contained in \( \mathcal{F} \) at exactly time \( t = \tau \) (i.e., in exactly in \( \tau \) steps) is given by

\[
 SR_\tau(G_t, \mathcal{F}) := \sum_{h \in \mathcal{F}} f_{th}^\tau.
\]

Let \( P := [p_{th}] \), where \( p_{th} \) is the probability that the network formation process, leaving network \( G_t \) at time \( t = \tau \), will arrive at network \( G_h \) at exactly time \( t = \tau + 1 \) (i.e., in exactly in 1 step). We will denote by

\[
 (P^\tau)_l \text{ the } l^{th} \text{ row of } P^\tau,
\]

\[
 (P^\tau)_h \text{ the } h^{th} \text{ column of } P^\tau.
\]

\(^6\)Note that, given the probability space \((\Theta, \mathcal{F}, \Pi)\) underlying the Markov process, \( \{ \tilde{G}^t \}_t \), where each network-valued random variable, \( \theta \rightarrow \tilde{G}^t(\theta) \), maps into the finite state space \( \mathcal{G} \), we have

\[
 \{ G^0 = G_k \} := \{ \theta \in \Theta : \tilde{G}^0(\theta) = G_k \in \mathcal{G} \} \in \mathcal{F}.
\]
Also, we will denote by
\[
[0_h(P)_{-h}] := [p_1, \ldots, p_{(h-1)}, 0_h, p_{(h+1)}, \ldots, p_H],
\]
where \([0_h(P)_{-h}]\) is the Markov transition matrix \(P\) with the \(h^{th}\) column replaced by the zero vector and let \([0_h(P)_{-h}]^\tau\) be the matrix \([0_h(P)_{-h}]\) raised to the \(\tau^{th}\) power. Then we have
\[
([0_h(P)_{-h}]^\tau)^l \text{ the } l^{th} \text{ row of } [0_h(P)_{-h}]^\tau,
\]
\[
([0_h(P)_{-h}])_{h'} \text{ the } h'^{th} (\neq h) \text{ column of } [0_h(P)_{-h}].
\]

The following result a variation on a result proved by Hunter (1983) - also, see Qianying Wu (2008).

**Theorem 1 (First Passage Probabilities)**

Let, \(\{G^n\}_n\), be the Markov process of network formation governed by Markov transition \(P\) with state space
\[
G := \{G_1, G_2, \ldots, G_H\} = \{G_l : l \in H\},
\]
and Markov supernetwork \(M\).

1. The systemic risk of the network formation process, \(\{\tilde{G}^n\}_n\), departing network \(G_l\) and arriving at failed network \(G_h\) exactly at time \(t = \tau\) in supernetwork \(M\) is given by
\[
SR_\tau(G_l, G_h) := \left< ([0_h(P)_{-h}]^\tau)^l, (P)_{\tau h} \right>.
\]
We will call \(SR_\tau(G_l, G_h)\) the systemic \(l\tau\)-risk.

2. If \(G_{\mathcal{F}} := \{G_l : l \in \mathcal{F}\}\) is the collection of all failed networks, then departing network \(G_l\) in supernetwork \(M\), the systemic risk of failure (at any level) at time \(\tau\) is given by
\[
SR_\tau(G_l, G_{\mathcal{F}}) := \sum_{h \in \mathcal{F}} SR_\tau(G_l, G_h)
\]
\[
:= \sum_{h \in \mathcal{F}} \left< ([0_h(P)_{-h}]^\tau)^l, (P)_{h} \right>.
\]
We will call \(SR_\tau(G_l, G_{\mathcal{F}})\) the systemic \(l\mathcal{F}\tau\)-risk

Thus, departing from network \(G_l\) and traveling along the transportation system provided by supernetwork \(M\), the risk of arriving at a failed network (i.e., the risk of arriving at some network contained in \(G_{\mathcal{F}}\)) after \(\tau\) periods (or moves) is given by \(SR_\tau(G_l, G_{\mathcal{F}})\) - while the risk of arrival on or before \(\tau\) is given by
\[
SR_{[1,\tau]}(G_l, G_{\mathcal{F}}) := \sum_{t=1}^{\tau} SR_t(G_l, G_{\mathcal{F}}) := \sum_{t=1}^{\tau} \sum_{h \in \mathcal{F}} \left< ([0_h(P)_{-h}]^\tau)^l, (P)_{h} \right>.
\]
The risk of arriving at some failed network in finite time, departing from network \( G_t \) in supernetwork \( M \), is given by

\[
SR_{[1, \infty]}(G_t, G_F) := \sum_{t=1}^{\infty} SR_t(G_t, G_F) = \sum_{t=1}^{\infty} \left[ \sum_{h \in F} \left( (0_h(P) - h)_i^T \right)^t, (P)_h \right].
\] (18)

Thus, given the stationary Markov process driving network formation, \( SR_{[1, \infty]}(G_t, G_F) \) is the systemic risk of ever experiencing any level of failure as the network moves through the supernetwork \( M \) after departing network \( G_t \in G \). We will call \( SR_{[1, \infty]}(G_t, G_F) \) the systemic \( lF \)-risk.

### 4.1.2 Computing Systemic Risk Schedules

Let \( \{\tilde{G}^n\}_n \) be the Markov process of network formation governed by the stationary Markov transitions given by Markov matrix, \( P \). Let state space of networks \( G \) be given by

\[
G = \{G_1, G_2, G_3, G_4, G_5, G_6\} = \{G_2, G_4\} \cup \{G_1, G_2, G_5, G_6\},
\]

with Markov transition matrix, \( P \), given by

\[
P = \begin{bmatrix}
.5 & .5 & 0 & 0 & 0 & 0 \\
.13447 & .13447 & 0 & 0 & .36553 & .36553 \\
0 & 0 & .5 & .5 & 0 & 0 \\
0 & 0 & .0596 & .0596 & .4404 & .4404 \\
.4404 & .4404 & 0 & 0 & .0596 & .0596 \\
0 & 0 & 0 & 0 & .5 & .5
\end{bmatrix}.
\]

The corresponding Markov supernetwork \( M \) (with only active connections shown) is depicted in Figure 2.
Using Theorem 2 and our definition of systemic risk, we have the following systemic risk schedules for arrival times \( \tau = 1-6 \), departing network \( G_3 \) arriving failed network \( G_1 \) in Column 1, and departing network \( G_3 \) arriving failed network \( G_5 \) in Column 2.

<table>
<thead>
<tr>
<th>( \tau = 0 )</th>
<th>Column 1</th>
<th>Systemic Risk Schedule</th>
<th></th>
<th>Column 2</th>
<th>Systemic Risk Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Departing</td>
<td>Departing Network 3,</td>
<td></td>
<td></td>
<td>Departing</td>
<td>Departing Network 3,</td>
</tr>
<tr>
<td>Arriving</td>
<td>Arriving Network 1</td>
<td></td>
<td></td>
<td>Arriving</td>
<td>Arriving Network 5</td>
</tr>
<tr>
<td>( \tau = 1 )</td>
<td>( SR_1(3, 1) = 0 )</td>
<td></td>
<td></td>
<td>( SR_1(3, 5) = 0 )</td>
<td></td>
</tr>
<tr>
<td>( \tau = 2 )</td>
<td>( SR_2(3, 1) = 0 )</td>
<td></td>
<td></td>
<td>( SR_2(3, 5) = 0.2202 )</td>
<td></td>
</tr>
<tr>
<td>( \tau = 3 )</td>
<td>( SR_3(3, 1) = 0.096976 )</td>
<td></td>
<td></td>
<td>( SR_3(3, 5) = 0.233324 )</td>
<td></td>
</tr>
<tr>
<td>( \tau = 4 )</td>
<td>( SR_4(3, 1) = 0.121576 )</td>
<td></td>
<td></td>
<td>( SR_4(3, 5) = 0.185618 )</td>
<td></td>
</tr>
<tr>
<td>( \tau = 5 )</td>
<td>( SR_5(3, 1) = 0.123064 )</td>
<td></td>
<td></td>
<td>( SR_5(3, 5) = 0.131397 )</td>
<td></td>
</tr>
<tr>
<td>( \tau = 6 )</td>
<td>( SR_6(3, 1) = 0.112832 )</td>
<td></td>
<td></td>
<td>( SR_6(3, 5) = 0.087292 )</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Systemic Risk Schedules from network 3 to networks 1 and 5 given supernetwork, \( M \), in Figure 2

We note that if the current network is \( G_3 \), then there is a probability of 0.185618 that at time \( \tau = 4 \) the prevailing network will be failed network \( G_5 \).

**4.1.3 Systemic Risk: Nodes and Sets of Nodes**

Let \( P(N) \) denote the collection of all nonempty subsets of nodes with typical element \( E \), and define the subset of states \( G_\mathcal{F}(E) \) to be the collection of networks having failed nodes given by the set \( E \subset N \). Thus, we have

\[
G_\mathcal{F}(E) := \{ G_l \in G_\mathcal{F} : N_f(G_l) = E \}.
\]  

The risk of ever arriving at a network having failed nodes \( E \) starting at network \( G_l \) is given by

\[
SR_{[1, \infty]}(G_l, G_\mathcal{F}(E)) := \begin{cases} 
0 & \text{if } G_\mathcal{F}(E) = \emptyset \\
\sum_{t=1}^{\infty} \left[ \sum_{h \in G_\mathcal{F}(E)} \langle ([0_h(P)_-h]^t)_h, (P)_h \rangle \right] & \text{otherwise.}
\end{cases}
\]

\[
(20)
\]

Suppose the current network is \( G_l \in G \) and suppose also that in network \( G_l \) node \( i \) is a continuation node - so that \( i \in N_c(G_l) \) (i.e., \( i \) is not a failed node). The probability that node \( i \) ever becomes a failed node is given by

\[
SR_{[1, \infty]}(G_l, G_\mathcal{F} \{i\}),
\]
while the probability that node \( i \) becomes a failed node at any time \( t \in \{1, 2, \ldots, \tau\} \) is
\[
SR_{\{1, \tau\}}(G_I, \mathcal{G}_F(\{i\})),
\]
or the probability that node \( i \) becomes a failed node exactly at time \( t = \tau \) is
\[
SR_{\tau}(G_I, \mathcal{G}_F(\{i\})).
\]
In the next section, we will analyze the effect that the failure of any particular node has on the entire network. If this effect negative, we will call \( node \ i \) systemic, if it is catastrophic we call \( node \ i \) killer. We will be helped in making this determination by the fact that all stationary Markov process of network formation with finite or compact state spaces generate a finite number of basins of attraction.

5 The Network Ruin Problem

The Markov process of network formation, \( \{\widetilde{G}^n\}_{n=0}^{\infty} \), governed by Markov transition matrix \( P \) with state space of networks,
\[
\mathcal{G} := \{G_1, G_2, \ldots, G_H\},
\]
generates a unique partition of \( \mathcal{G} \) into a finite number of basins of attraction, given by
\[
\mathcal{A} := \{\mathcal{A}^1, \mathcal{A}^2, \ldots, \mathcal{A}^L\},
\]
and a transient set \( T \). In particular,
\[
\mathcal{G} = \left( \bigcup_{l=1}^{L} \mathcal{A}^l \right) \cup T,
\]
where each \( \mathcal{A}^l \) is a basin of attraction\(^7\), and \( T \) is transient. Moreover, the process, \( \{\widetilde{G}^n\}_{n=0}^{\infty} \), possesses a unique finite set of ergodic probability measures (i.e., long run equilibrium probability measures) over networks, one for each basin of attraction, and in general each invariant probability measure is a convex combination of these ergodic measures (see Appendix B, for proofs of the statements above).

The presence of basins of attraction has major implications for the structure and magnitude of the systemic risk schedules determined by the supernetwork. These implications underscore the importance of understanding network dynamics in being able to properly assess systemic risk, identify tipping points, and design policies which increase the likelihood that the process of network formation will follow the safer paths at tipping points in the supernetwork. In this section we will discuss the implications of the presence of basins of attraction for systemic risk schedules and we will give precise characterizations of tipping points and systemic nodes.

\(^7\)Basins of attractions are the largest absorbing sets (see Definition 7).
5.1 Basin of Attraction, Tipping Points and Spheres of Influence

Because the Markov process of network formation partitions the state space into a transient set, \( T \), together with finitely many basins of attraction, \( A := \{A^1, A^2, \ldots, A^L\} \), we can say much more about systemic risk than before.

Because failed nodes persist (see assumption \[A-3\]) and because the process of network formation, once it enters a basin of attraction stays in the basin visiting each state (and hence each network) in the basin infinitely often, it follows that if the basin, say basin \( A^\lambda \), contains even one failed network, \( G^f_h \in A^\lambda \), having failed node set, \( F^f \), consisting of \( k \) failed nodes (i.e., \( |F^f| = k \)), then all networks in basin \( A^\lambda \) have failed networks with failed node set, \( F^f \). We say that such a basin, \( A^\lambda \) is a \( k \)-level basin. We summarize these observations in the following Theorem.

**Theorem 2 (Homogeneity of Failure Levels of Basins of Attraction)**

Let \( \{G^n\}_{n=0}^\infty \) be the Markov process of network formation governed by Markov transition matrix, \( P \), with path structure given by supernetwork \( M \) having basins of attraction

\[
A := \{A^1, A^2, \ldots, A^L\}.
\]

For each basin, \( A^\lambda \), there is a unique subset of failed nodes, \( E^\lambda \), such that

\[
A^\lambda \subseteq \mathbb{G}_F(E^\lambda).
\]  

(23)

Again let \( \mathbb{G}_F \subseteq \mathbb{G} \) denote the collection of failed networks (i.e., the set of networks having a failed nodes), and let \( B := \{1, 2, \ldots, L\} \), be index set for the set of basins of attraction, \( A \). Consider subsets of networks given by

\[
A_f := \bigcup_{l \in B_f} A^l \quad \text{and} \quad A_c := \bigcup_{l \in B_c} A^l,
\]

(24)

where

\[
B_f := \{l \in B : \mathbb{G}_F \cap A^l \neq \emptyset\},
\]

and

\[
B_c := \{l \in B : \mathbb{G}_F \cap A^l = \emptyset\}.
\]

If \( l \in B_f \), then basin \( A^l \) consists of failed networks each having a failed nodes, \( E^l \), consisting of \( l \) failed nodes. Of course, \( E^l \) can differ from basin to basin. Alternatively, if \( l \in B_c \), then \( A^l \) is free of failed networks. Thus, the state space of networks can be partitioned as follows:

\[
\mathbb{G} = T \cup A_c \cup A_f,
\]

where for each transient network \( G^T_h \in T \),

\[
SR_{[1,\infty)}(G^T_h, A_c) + SR_{[1,\infty)}(G^T_h, A_f) = 1
\]

If at \( G^T_h \in T \),

\[
SR_{[1,\infty)}(G^T_h, A_c) = 1,
\]

15
then failure is no longer possible because in finite time with probability 1 the process of network formation will enter a basin of attraction containing no failed networks - and will remain there for all future periods. However, if at $G_h \in \mathbb{T}$,

$$SR_{(1,\infty)}(G_h, A_f) = 1,$$

then failure is inevitable because in finite time with probability 1 the process of network formation will enter a basin of attraction containing failed networks.

But we can say more. Each basin of attraction, $A^l \in \mathcal{A}$, has a sphere of influence defined as

$$SI(A^l) := \left\{ G_h \in \mathbb{G} : SR_{(1,\infty)}(G_h, A^l) = 1 \right\}.$$  \hspace{2cm} (25)

Thus, if the process reaches network, $G_h$, contained in the sphere of influence, $SI(A^l)$, of basin $A^l$, then in finite time with probability 1 the process will enter basin $A^l$ - and if $A^l$ is a $k$-level basin, $k > 0$, then $k$-level failure is inevitable once network $G_h \in SI(A^l)$ is reached. The sphere of influence of any basin can be small. In fact, it is even possible that a basin is its own sphere of influence. Thus for some basins, it is possible that

$$SI(A^l) = A^l.$$

Now let

$$SI(A_f) := \bigcup_{l \in B_f} SI(A^l)$$

and

$$SI(A_c) := \bigcup_{l \in B_c} SI(A^l).$$

The subset of states $SI(A_f)$ is the sphere of influence of the set of all failed networks, while $SI(A_c)$ is the sphere of influence of safe networks.

### 5.2 Systemic and Killer Nodes

Again, we begin with a definition.

**Definition 5** (Endogenous Tipping Points, Systemic Nodes, and Killer Nodes)

Let $\{G^n\}_{n=0}^{\infty}$ be the Markov process of network formation governed by Markov transition matrix, $P$, with path structure given by supernetwork $M$ having basins of attraction

$$\mathcal{A} := \{ A^1, A^2, \ldots, A^L \}.$$  \hspace{2cm} (1)

1. (Tipping Points) A network $G_h \in \mathbb{G}$ is a tipping point if $SR_1(G_h, SI(A_c)) > 0$ and $SR_1(G_h, SI(A_f)) > 0$ and

$$SR_1(G_h, SI(A_c)) + SR_1(G_h, SI(A_f)) = 1.$$  \hspace{2cm} (26)

2. (Systemic Nodes and Killer Nodes) If for tipping point network, $G_h \in \mathbb{G}$, with node $i$ not a failed node (i.e., with $i \in N_c(G_h)$) there is a successor network, $G_h' \in \mathbb{G}$ such that

$$N_f(G_h') = N_f(G_h) \cup \{i\}$$
and

\[ SR_1(G_{h'}, SI(\Lambda_f)) = 1, \]

then we say that node \( i \) is systemic. If there is a basin, \( A^{tip} \), containing a network \( G_{h'} \), with all failed nodes, and if node \( i \) is such that for some tipping point network, \( G_h \in \mathbb{G} \), with \( i \in N_c(G_h) \), there is a successor network, \( G_{h'} \in \mathbb{G} \) such that

\[ N_f(G_{h'}) = N_f(G_h) \cup \{i\} \]

and

\[ SR_1(G_{h'}, SI(A^{tip})) = 1, \]

then we say that node \( i \) is killer.

Thus, a killer node \( i^{tip} \) is a node in a tipping point network whose failure propels the network formation process into the sphere of influence of a basin consisting of totally failed networks. The failure of such a node is catastrophic.

We close this section by noting that if each of the finitely many basins of attraction contain a failed network, then \( \Lambda_c = \emptyset \), and

\[ SR_{[1,\infty)}(G_h, \Lambda_f) = 1 \text{ for all } G_h \in \mathbb{T}. \]

In this case, failure (but perhaps not catastrophic failure) is inevitable because in finite time with probability 1 the process of network formation will enter a basin of attraction containing failed networks.

6 Strategic Foundations of Systemic Risk

We mentioned in the introduction that our approach to systemic risk allows us to bring into the analysis the strategic underpinnings of systemic risk in a very direct and simple way - directly through the supernetwork. Let’s see how this works. If we model network formation as a discounted stochastic game (as we will do immediately below), then the players’ Nash equilibrium in stationary Markov strategies of network formation will give rise to an equilibrium stationary Markov process of network formation - and therefore, to an equilibrium Markov supernetwork. Using the structure and properties of this equilibrium Markov supernetwork, we are able to compute the schedule of endogenous systemic risk measures - that is, the schedule of systemic risks that arises from the dynamic, strategic behavior of the players in forming networks.

6.1 Discounted Stochastic Games of Network Formation

Consider a discounted stochastic game (DSG) of network formation having the following primitives:

A finite set of players,

\[ N := \{i_1, i_2, \ldots, i_n\}.^8 \]

^8Here we are assuming that the set of players and the set of nodes, \( N \), are one and the same.
A finite state space of networks,

\[ \mathcal{G} := \{G_1, G_2, \ldots, G_H\} := \{G_h : h \in \mathcal{H}\}. \]

A law of motion,

\[ (G, G_N) \rightarrow q(\cdot|G, G_N) \in \Delta(\mathcal{G}(G)), \]

where

\[ G_N := (G^i)_{i \in N} \in \mathcal{G}_N \]

and

\[ G^i \in \mathcal{G} := \{G_h : h \in \mathcal{H}\} \]

is player \( i \)'s network proposal,

and where in harmony with assumption \([A-2]\),

\[ \mathcal{G}(G) := \{G' \in \mathcal{G} : \mathcal{D}_f(G) \subseteq \mathcal{D}_f(G')\}. \]

A set of player value functions, payoff functions, discount rates, and constraint correspondences,

\[ \{v_i(\cdot), r_i(\cdot, \cdot), \beta_i, \Phi_i(\cdot)\}_{i \in N}. \]

where

\[ v_i(\cdot) : \mathcal{G} \rightarrow [-M, M] \text{ is player } i \text{'s value function,} \]

\[ r_i(\cdot, \cdot) : \mathcal{G} \times \mathcal{G}_N \rightarrow [-M, M] \text{ is player } i \text{'s value function,} \]

\[ \beta_i \in [0, 1] \text{ is player } i \text{'s discount rate,} \]

and

\[ \Phi_i(\cdot) : \mathcal{G} \rightarrow P(\mathcal{G}) \text{, is player } i \text{'s constraint correspondence.} \]

In our game each player’s action takes the form of a network recommendation or network proposal. In particular, given current state, \( G \in \mathcal{G} \), each player \( i \in N \) has available a nonempty subset of network proposals \( \Phi_i(G) \subset \mathcal{G} \) that can be put forth by player \( i \). Player \( i \)'s proposal together with the proposals of other players are then discussed, bargained over, and compromised leading to a risky final outcome. Here we will not model this bargaining process explicitly, but instead we will represent the process using a controlled (action-dependent) Markov transition kernel (we will provide details below). We will call this controlled Markov transition kernel, the law of motion.

We can think of each player’s action (network proposal) constraint correspondence as representing the rules of network formation. In particular, if in current state, \( G \), network \( G' \in \Phi_i(G) \) is proposed by player \( i \in N \), then the proposed network \( G' \)
must be such that under the rules of network formation it is possible for player \( i \), working with other players, to change the status quo network \( \Gamma \) to network \( \Gamma_0 \). We will assume the following concerning players’ constraint correspondences:

[A-4] (properties of constraint mappings) 
For each player \( i \in \mathbb{N} \), if the current state is \( \Gamma \in \mathbb{G} \), then (1) for each continuing player \( i \in N_c(G) \), \( G \in \Phi_i(G) \subset \mathbb{G}(G) \), and (2) for each failed player \( i \in N_f(G) \), \( \Phi_i(G) = \{G\} \).

Under [A-4](1), each continuing player can propose that the status quo be maintained or can propose a new network in accordance with the rules of network formation. Under [A-4](2) each failed player can propose only that the status quo network be maintained. The aggregate constraint correspondence \( \Phi(\cdot) \) is given by

\[
G \rightarrow \Phi(G) := \Pi_{i \in \mathbb{N}} \Phi_i(G). \tag{27}
\]

The probabilistic movement of the process of network formation through the supernetwork from one network to another is governed by the law of motion,

\[
q(\cdot, \cdot): Gr\Phi(\cdot) \rightarrow \Delta(\mathbb{G}), \tag{28}
\]

a mapping defined on the graph of the feasible action profile correspondence, \( Gr\Phi(\cdot) \), taking values in the set of probability measures on the state space, \( \Delta(\mathbb{G}) \), such that under the law of motion, \( q(\cdot, \cdot) \), for each state-action profile pair, \( (G, G_N) \in Gr\Phi(\cdot) \),

\[
q(\cdot|G, G_N) \in \Delta(\mathbb{G}(G)),
\]

where \( \Delta(\mathbb{G}(G)) \) is the set of all probability measures with support contained in \( \mathbb{G}(G) \).

By Blackwell’s Theorem (1962, 1965) extended to discounted stochastic games (i.e., DSGs), a conditional product measure formed from a profile of player stationary Markov strategies,

\[
\sigma^*(\cdot|\cdot) := (\sigma^*_1(\cdot|\cdot) \times \cdots \times \sigma^*_n(\cdot|\cdot))
\]

together with a profile of player value functions

\[
v^*(\cdot) = (v^*_1(\cdot), \ldots, v^*_n(\cdot)),
\]

are an equilibrium for the discounted stochastic game of network formation specified above if and only if for each player \( d \) and for each initial state \( G \), the following two conditions (the Bellman condition and the Nash condition) are satisfied:

Bellman Condition

\[
v^*_i(G) = U_i(G, v^*_i, (\sigma^*_i(\cdot|G), \sigma^*_{-i}(\cdot|G))), \tag{Bellman Condition}
\]

Nash condition

\[
U_i(G, v^*_i, (\sigma^*_i(\cdot|G), \sigma^*_{-i}(\cdot|G))) = \max_{\sigma_{-i} \in \Delta(\mathbb{G}(G))} U_i(G, v^*_i, (\sigma^*_i, \sigma^*_{-i}(\cdot|G))). \tag{Nash condition}
\]
where

\[ U_i(G, v_i^*, (\sigma_i^*(*|G), \sigma_{-i}^*(*|G))) = (1 - \beta_i) \sum_{G_N \in G_N} r_d(G, G_N)\sigma^*(G_N|G) + \beta_i \sum_{G_N \in G_N} v_i^*(G')q(G'|G, G_N)\sigma^*(G_N|G). \]

By Federgruen (1978), we know that our discounted stochastic game of network formation has an equilibrium, \((\sigma^*(*|\cdot), v^*(*))\) (see Filar et al, 1991, for a full characterization of the equilibrium as well as Herings and Peeters, 2004).

Given the strategy part of the equilibrium, and therefore given the stationary Markov equilibrium strategy profile, \(\sigma^*(*|\cdot) \in \prod_i \Sigma(\Phi_i(*))\), the induced equilibrium Markov supernetwork of the equilibrium network formation dynamics is given by \(M(p^*(*|\cdot)) := M^*\) with typical connection

\[ G_h \xrightarrow{p^*(G_h|G_i)} G_i \]

Figure 2: The directed connection from network \(G_i\) to \(G_h\) in Markov equilibrium supernetwork \(M^*\)

where \(p^*_h := p^*(G_h|G_i)\) is the conditional probability that for any \(t\) the equilibrium process departs network \(G_i\) at \(t\) and arrives network \(G_h\) at \(t + 1\). This conditional probability is given by

\[ p^*(G_h|G_i) := \sum_{G_N \in \Phi_N(G_i)} q(G_h|G_i, G_N)\sigma^*(G_N|G_i) := q(G_h|G_i, \sigma^*(G_i)). \]

### 6.2 From Systemic Risk to Endogenous Systemic Risk

Using the equilibrium Markov supernetwork, \(M^*\), in computing systemic risk schedules, we obtain *endogenous systemic risk* measures. We use this terminology because by basing our computation upon the strategically informed supernetwork, \(M^*\), we are taking into account the feedback between network structure, strategic behavior, and risk in computing endogenous systemic risk. These feedback effects are the determinants of the stochastic process of network formation which emerges in equilibrium.

### 6.3 How Restrictive are Game Theoretic Models with Finite State and Action Spaces

First, if our objective is to compute endogenous systemic risk schedules or to find tipping point networks and systemic nodes - or to find killer nodes, then we must restrict our discounted stochastic game models of network formation to finite state and finite action space models. In financial applications, this might seem very restrictive because such models by their very nature require uncountable many states and actions. For example, if the networks we are analyzing are interbank lending networks,
then connections between bank pairs take the form of state-contingent loan contracts with payoffs depending on - for example - the banks underlying asset returns. Often the stochastic nature of these returns is best captured using a continuous joint density function - and the set of loan contracts is best represented by a function space. Thus, even if there are only finitely many node pairs (or pre-connections), the set of arcs is given by a function space of contracts where each contract is a function of an underlying vector of continuous asset returns. In such an application, the state space of networks is uncountable. However, often times, even in such banking applications the state space of networks, while uncountable, is compact or locally compact. If for example the state space of networks is compact metric space (see Gong, Page, and Wooders, 2015, for many such examples) then for each $\varepsilon > 0$, there is a finite subset of networks such that each network in the uncountable state space is within $\varepsilon$ distance of at least one network contained in the approximating finite subset of networks.\(^9\) Thus, game theoretic models with finite state and action spaces are good approximations of the uncountable-compact models - and moreover, as mentioned above, finite models are potentially computable.

### References


\(^9\) Of course in order for the desirable approximating properties of finite state-finite action discounted stochastic games to be useful in providing approximate measures of endogenous systemic risk, it must be the case that the corresponding uncountable-compact discounted stochastic game of network formation has a stationary Markov equilibrium in the first place. This turns out to be a very difficult existence question - but see Page and Fu (2019) for just such existence results.


Appendix A: Classical Results on Finite Markov Chains

7.1 Hitting Times and Systemic Risk

Closely related to systemic risk (a probability) is the notion of a hitting time. Often we will be interested in determining the probability with which the network formation process reaches or hits in finite time a particular network $G_h$ after leaving a particular network $G_l$. Here the issue is, does the process reach network $G_h$ at all starting from $G_l$.

Let $\{\tilde{G}^n\}_{n=0}^{\infty}$ be the Markov process of network formation governed by Markov transition $P$ with state space of networks $G := \{G_1, G_2, \ldots, G_H\}$, and Markov supernetwork $M$. Consider the integer-valued random variable, $\tau_{G_h}(\cdot) : \Theta \rightarrow \{1, 2, \ldots\}$, given by

$$\tau_{G_h} := \inf \left\{ n \geq 1 : \tilde{G}^n = G_h \right\}. \quad (29)$$

We call $\tau_{G_h}$ network $G_h$’s hitting time under the network formation process $\{\tilde{G}^n\}_{n=0}^{\infty}$.

Next, let $\rho_{lh} := \Pi \left\{ \tau_{G_h} < \infty | \tilde{G}^0 = G_l \right\}$ be the probability that the process $\{\tilde{G}^n\}_{n=0}^{\infty}$ reaches state $G_h$ in finite time after leaving network $G_l$ at time zero. It is easy to see that if $G_h$ is a failed state, then the hitting probability, $\rho_{lh}$, is equal to the systemic $lh$-risk, $SR_{[1,\infty)}(G_l, G_h)$. Thus, we have

$$\rho_{lh} := \Pi \left\{ \tau_{G_h} < \infty | \tilde{G}^0 = G_l \right\} = SR_{[1,\infty)}(G_l, G_h),$$

and the probability that the process $\{\tilde{G}^n\}_{n=0}^{\infty}$ reaches network $G_h$ in finite time after leaving network $G_l$ at time zero is equal to the systemic $lh$-risk in supernetwork $M$.

The expected recurrence time (expected number of moves) for the process $\{\tilde{G}^n\}_{n=0}^{\infty}$ to reach $G_l$ again after leaving $G_l$ at time zero is

$$E_{G_l, \tau_{G_h}} := \mathbb{E}_t[\tilde{G}^0 = G_l, \tilde{G}^s \neq G_l, \forall 1 < s < t] = \sum_{t=1}^{\infty} t f_{li} \quad (31)$$

Next consider the sequence of hitting times $\{\tau_{G_h}^l\}_{l=0}^{\infty}$ defined recursively as follows:

$\tau_{G_h}^0 := 0$, $\tau_{G_h}^1 := \tau_{G_h}$, and for $l \geq 2,$

$$\tau_{G_h}^l := \inf \left\{ n > \tau_{G_h}^{l-1} : \tilde{G}^n = G_h, \tilde{G}^{\tau_{G_h}^{l-1}} = G_h \right\}. \quad (32)$$

$\tau_{G_h}^l$ is the number of moves required for the $l^{th}$ hit on $G_h$. The following two results on hitting times can be found in Durrett (2005) section 5.3.
Theorem 3 (Hitting Times)
Let \( \{ \tilde{G}^n \}_{n=0}^{\infty} \) be the Markov process of network formation governed by Markov transition \( P \) with state space of networks
\[ \mathcal{G} := \{ G_1, G_2, \ldots, G_H \}. \]
The following statements are true:
1. For all \( l \) and all networks \( G_h \) and \( G_{h'} \), \( \Pi \left\{ \tau_{G_h}^l < \infty | \tilde{G}^0 = G_h \right\} = \rho_{hh'} \rho_{h'}^{l-1}. \)
2. For all networks \( G_h \), \( G_h' \) and \( G_{h''} \), \( \rho_{hh''} \geq \rho_{hh} \rho_{h'h''} \).

7.2 Recurrence, Transience, and Absorbing Sets
A network \( G_h \) is said to be recurrent if \( \rho_{hh} = 1 \) and transient if \( \rho_{hh} < 1 \). By part (1) of Theorem 3, if \( G_h \) is recurrent, then
\[ \Pi \left\{ \tau_{G_h}^l < \infty | \tilde{G}^0 = G_h \right\} = 1 \text{ for all } l. \]
Given any network \( G_{h'} \in \mathcal{G} \), the number of visitations to \( G_{h'} \) by the process \( \{ \tilde{G}^n \}_{n=0}^{\infty} \) after time zero is given by
\[ \eta_{G_{h'}} := \sum_{n=1}^{\infty} I_{\{ \tilde{G}^n = G_{h'} \}}. \tag{33} \]
If network \( G_{h'} \) is transient, then the expected number of visitations to \( G_{h'} \) starting from network \( G_h \) is given by
\[ g_{hh'} := \mathbb{E}_{G_h} [\eta_{G_{h'}}] = \sum_{l=1}^{\infty} \Pi_{G_h} \{ \eta_{G_{h'}} \geq l \} = \sum_{l=1}^{\infty} \Pi \left\{ \tau_{G_{h'}}^l < \infty | \tilde{G}^0 = G_h \right\} = \sum_{l=1}^{\infty} \rho_{hh'} \rho_{h'h'}^{l-1} \text{ (by Theorem 3 (1))} = \frac{\rho_{hh'}}{1 - \rho_{h'h'}} < \infty. \tag{34} \]
We can conclude from (20) that in fact \( G_{h'} \) is recurrent if and only if
\[ g_{hh'} := \mathbb{E}_{G_{h'}} [\eta_{G_{h'}]} = \infty. \]
The following classical result on recurrent states can be found in Durrett (2005), for example.

Theorem 4 (Recurrence is Contagious)
Let \( \{ \tilde{G}^n \}_{n=0}^{\infty} \) be the Markov process of network formation governed by Markov transition \( P \) with state space of networks
\[ \mathcal{G} := \{ G_1, G_2, \ldots, G_H \}. \]
If network \( G_h \) is recurrent and network \( G_{h'} \) is reachable, that is, if \( \rho_{hh'} > 0 \), then \( G_{h'} \) is recurrent and \( \rho_{h'h} = 1 \).
Given Markov transition kernel, \( p(G' | G) \), a set of networks, \( \mathcal{A} \), is said to be absorbing if \( p(\mathcal{A} | G) = 1 \) for all \( G \in \mathcal{A} \). Because the number of failed nodes, \( |N_f(G^n)| \), is nondecreasing across time (see expression 15) if \( G^n \in \mathcal{A} \subseteq \mathcal{G}_\mathcal{F} \) is a failed network, having failed nodes, \( N_f(G^n) \), then all networks \( G'^n \) in absorbing set \( \mathcal{A} \) will be failed networks having exactly the same set of failed nodes. In particular,

\[
N_f(G'^n) = N_f(G^n) \text{ for all } n' > n.
\]

### 7.3 Functions, Transition Probabilities, and Expectations

Let \( f(\cdot) \) be a real-valued function defined on the state space \( \mathcal{G} := \{G_1, G_2, \ldots, G_H\} \) of networks and let \( f = (f_1, \ldots, f_H) \in \mathbb{R}^H \) be the corresponding vector representation. Thinking of \( f \) as a column vector and the initial probability measure \( \pi = (\pi_1, \ldots, \pi_H) \in \Delta(\mathcal{G}) \) on the state space of networks as a row vector, the expected value of \( f(\cdot) \) given initial probability measure \( \pi \) is given by

\[
\pi f = \sum_{G_0} \pi(G_0) \left( \sum_{G_{0,n} \in \mathcal{G}} f(G_{0,n}) \Pi \{ G_{0,n} = G_0 \} \right)
\]

and

\[
\mathbb{E}(f(G^n) | G^0 = G_0) = \sum_{G_{0,n} \in \mathcal{G}} f(G_{0,n}) \Pi \{ G^n_{0,n} = G_0, G^0 = G_0 \}
\]

\[
= \sum_{G_{0,n} \in \mathcal{G}} (P^n)_{G_{0,n}G_0} f_{G_{0,n}} = (P^n f)_{G_0}.
\]

Therefore, \( \mathbb{E}(f(G^n)) = \sum_{G_0} \gamma_{G_0} \mathbb{E}(f(G^n), G^0 = G_0) = \gamma P^n f \).

### 8 Appendix B: Basins of Attraction and Ergodic Probability Measures

#### 8.1 Basins of Attraction

Our primary objective in this section is to show that the Markov process of network formation, \( \{G^n\}_{n=0}^\infty \), governed by Markov transition \( P \) with state space \( \mathcal{G} := \{G_1, G_2, \ldots, G_H\} \), generates a unique partition of the state space of networks \( \mathcal{G} \)

\[
\mathcal{G} = \left( \bigcup \mathcal{A}^l \right) \cup \mathcal{T},
\]

where each \( \mathcal{A}^l \) is a basin of attraction\(^{10} \), and \( \mathcal{T} \) is transient. We will also show that \( \{G^n\}_{n=0}^\infty \) possesses a unique finite set of ergodic probability measures (i.e., long run equilibrium probability measures) over networks one for each basin of attraction, and that each invariant probability measure is a convex combination of these ergodic measures.

\(^{10}\)Basins of attractions are the largest absorbing sets (see Definition 7).
We say that there is a path from network $G_l$ to network $G_h$ if $\rho_{lh} > 0$. If, in addition, $\rho_{hl} > 0$, so there is a path back, then we say that networks $G_l$ and $G_h$ are on the same circuit. In particular, if $\rho_{ll} > 0$, then there is a path from $G_l$ to $G_l$.

We say that a set of states $G_E \subseteq G$ is irreducible if for each pair of network $G_l$ and $G_h$ contained in $G_E$, there is a path from $G_l$ to $G_h$ and back. Thus, $G_E$ is irreducible if and only if for every pair of networks $G_l$ and $G_h$ contained in $G_E$, $\rho_{lh} > 0$ and $\rho_{hl} > 0$ - and thus, if $G_E$ is irreducible, then $\rho_{ll} > 0$ for all $G_l \in G_E$.\footnote{If the entire state space of networks $G$ is irreducible, we say that the process $\{\tilde{G}^n\}_{n=0}^{\infty}$ governed by Markov transition $P$ is irreducible.} In fact, by the Contagious Theorem 4, if $G_E$ is irreducible, then all states in $G_E$ are either recurrent ($\rho_{ll} = 1$ for all $G_l \in G_E$) or transient ($\rho_{ll} < 1$ for all $G_l \in G_E$). Finally, we say that a set of states $G_E \subseteq G$ is closed if for all $G_l \in G_E$, $\rho_{lh} > 0$ implies that $G_h \in G_E$.

**Definition 7 (Basins of Attraction)**

A set of networks $A \subseteq G$ is said to be a basin of attraction for the process $\{\tilde{G}^n\}_{n=0}^{\infty}$ governed by Markov transition $P$ if $A$ is closed and irreducible.

The following classical results tell us the precise relationship between irreducibility, recurrence, and closedness (see, for example, Durrett, 2005).

**Theorem 5 (Closedness Implies Recurrence Under Irreducibility in Finite State Markov Processes)**

Let $\{\tilde{G}^n\}_{n=0}^{\infty}$ be the Markov process of network formation governed by Markov transition $P$ with state space $G := \{G_1, G_2, \ldots, G_H\}$.

The following statements are true:

(1) If $A \subset G$ is closed, then it contains at least one recurrent network $G_l$, that is, $\rho_{ll} = 1$ (or equivalently, $g_{ll} = \infty$) for some $G_l \in A$. Moreover, $A \subset G$ is closed if and only if for all $G_l \in A$,

$$\Pi \{\tilde{G}^n \in A | \tilde{G}^0 = G_l\} = 1$$

for all $n$. ($\star$)

(2) If $A \subset G$ is closed and irreducible, then it is recurrent.

**Proof.** (1) The proof of ($\star$) is straightforward. Now suppose $A$ is closed but contains no recurrent states, so that $g_{hh} < \infty$ for all $G_h \in A$. But now we have a contradiction because by (34) and ($\star$) in part (1) we have

$$\infty > \sum_{G_h \in A} g_{lh} = \sum_{G_h \in A} \mathbb{E} G_l [\eta_{G_h}]$$

$$= \sum_{G_h \in A} \sum_{n=1}^{\infty} [P^n]_{lh} = \sum_{n=1}^{\infty} \sum_{G_h \in A} [P^n]_{lh}$$

$$= \sum_{n=1}^{\infty} \Pi \{\tilde{G}^n \in A | \tilde{G}^0 = G_l\}$$

$$= \lim_{n \to \infty} \Pi \{\tilde{G}^n = G_l\}$$

If the entire state space of networks $G$ is irreducible, we say that the process $\{\tilde{G}^n\}_{n=0}^{\infty}$ governed by Markov transition $P$ is irreducible.
Suppose $A$ is closed and irreducible. By closedness we know by part (1) that there is at least one recurrent network in $A$. By the Contagious Theorem 4 we know that any network reachable from this recurrent network is also recurrent. By irreducibility, we know that all networks in $A$ are reachable. Therefore, all networks in $A$ are recurrent. 

8.2 Markov Dominance and Systemic Risk

Let $M$ be the Markov supernetwork corresponding to the Markov transition matrix $P$. Thus, 

$$M \subseteq [0, 1] \times (\mathbb{G} \times \mathbb{G}),$$

with typical connection, 

$$(p_{lh}, (G_l, G_h)) \in [0, 1] \times (\mathbb{G} \times \mathbb{G}),$$

where $p_{lh}$ is the $lh^{th}$ entry in the Markov transition matrix $P$ specifying the probability with which the network formation process moves from network $G_l$ to network $G_h$. The connection $(p_{lh}, (G_l, G_h)) \in M$ is active if and only if the process of network formation \(\{G^n\}_{n=0}^\infty\) governed by Markov transition $P$ is such that for all $n = 1, 2, \ldots,$ 

$$\Pi\{G^n = G_l|G^{n-1} = G_l\} = p_{lh} > 0.$$

Network $G_h \in \mathbb{G}$ Markov dominates network $G_l \in \mathbb{G}$, written, $G_h \succeq_M G_l$, if there is an active path from network $G_l$ to network $G_h$, or if $G_l = G_h$. Formally, we have the following definition:

**Definition 8 (The Markov Dominance Relation)**

*Let $M \subseteq [0, 1] \times (\mathbb{G} \times \mathbb{G})$ be a Markov supernetwork for Markov transition $P$. For any two networks, $G_l$ and $G_h$ in $\mathbb{G}$ define*

$$G_h \succeq_M G_l \text{ if and only if } \begin{cases} \rho_{lh} > 0, & \text{or} \\ G_h = G_l, \end{cases}$$

*where*

$$\rho_{lh} := \Pi\{\tau_{G_h} < \infty|G^0 = G_l\} = SR_{\{1, \infty\}}(G_l, G_h),$$

*The Markov dominance relation $\succeq_M$ is a weak ordering on the set of networks $\mathbb{G}$. In particular, $\succeq_M$ is reflexive ($G_l \succeq_M G_l$) and $\succeq_M$ is transitive ($G_{h'} \succeq_M G_h$ and $G_h \succeq_M G_l$ implies that $G_{h'} \succeq_M G_l$).*

*If $G_h \succeq_M G_l$ and $G_l \succeq_M G_h$, we say that states $G_l$ and $G_h$ are equivalent and we write $G_l \equiv_M G_h$. If states $G_l$ and $G_h$ are equivalent this means that either states $G_l$ and $G_h$ coincide or that $G_l$ and $G_h$ are on the same circuit in Markov supernetwork $M$. If states $G_l$ and $G_h$ are such that $G_h \succeq_M G_l$ but $G_l$ and $G_h$ are not equivalent (i.e., but not $G_l \equiv_M G_h$), we say that network $G_h$ is a descendant of network $G_l$ and we write*

$$G_h \succ_M G_l.$$
We say that a network \( G_{h'} \in \mathbb{G} \) is maximal in supernetwork \( \mathbf{M} \) if for any \( G_l \in \mathbb{G} \)

\[
G_l \succeq_{\mathbf{M}} G_{h'} \implies G_l \equiv_{\mathbf{M}} G_{h'},
\]
that is, if \( G_{h'} \) is maximal then \( G_l \succeq_{\mathbf{M}} G_{h'} \) implies that \( G_l \) and \( G_{h'} \) coincide or lie on the same circuit in supernetwork \( \mathbf{M} \). Thus, given the definition of descendants, maximal networks are precisely those networks without descendants in supernetwork \( \mathbf{M} \). Letting

\[
P_{>\mathbf{M}}(G_{h'}) := \{ G_l \in \mathbb{G} : G_l \succeq_{\mathbf{M}} G_{h'} \},
\]
a network \( G_{h'} \in \mathbb{G} \) is without descendants or is maximal in the supernetwork \( \mathbf{M} \) if and only if

\[
P_{>\mathbf{M}}(G_{h'}) = \emptyset.
\]

Here is our main result concerning networks without descendants.

**Theorem 6 (All Finite Markov Supernetworks Have Networks Without Descendants)**

Let \( \mathbf{M} \subset [0, 1] \times (\mathbb{G} \times \mathbb{G}) \) be a Markov supernetwork for Markov transition \( \mathbf{P} \). For every network \( G_l \in \mathbb{G} \) there exists a network \( G_h \in \mathbb{G} \) such that,

1. \( G_h \succeq_{\mathbf{M}} G_l \), and
2. \( P_{>\mathbf{M}}(G_h) = \emptyset \).

**Proof.** Let \( G^0 \) be any network in \( \mathbb{G} \). If \( P_{>\mathbf{M}}(G^0) = \emptyset \), we are done. If not, choose \( G^1 \in P_{>\mathbf{M}}(G^0) \). If \( P_{>\mathbf{M}}(G^1) = \emptyset \), we are done. If not, continue by choosing \( G^2 \in P_{>\mathbf{M}}(G^1) \). Proceeding iteratively in this way, we can generate a sequence of networks, \( G^0, G^1, G^2, \ldots \), of networks in \( \mathbb{G} \). Now observe that in a finite number of iterations we must come to a network \( G'' \) such that \( P_{>\mathbf{M}}(G'') = \emptyset \). Otherwise, we could generate an infinite sequence of states, \( \{G^l\}_l \) in \( \mathbb{G} \) such that for all \( l \),

\[
G^l \succeq_{\mathbf{M}} G^{l-1}.
\]

However, because \( \mathbb{G} \) is finite, this sequence would contain at least one state, say \( G'' \), which is repeated an infinite number of times. Thus, all the networks in the sequence lying between any two consecutive repetitions of \( G'' \) would be on the same circuit in supernetwork \( \mathbf{M} \), contradicting the fact that for all \( l \), \( G^l \) is a descendant of \( G^{l-1} \) (i.e., \( G^l \succeq_{\mathbf{M}} G^{l-1} \)).

By Theorem 6, in any finite Markov supernetwork \( \mathbf{M} \), corresponding to any network \( G_l \in \mathbb{G} \) there is a network \( G_h \in \mathbb{G} \) without descendants which is reachable from \( G_l \). Thus, in any finite Markov supernetwork \( \mathbf{M} \) the set of networks without descendants given by

\[
Z := \{ G' \in \mathbb{G} : P_{>\mathbf{M}}(G') = \emptyset \}
\]

is nonempty.

We show that all Markov transitions governing network and coalition formation generate a unique set of basins of attraction and these basins are determined by the set of states without descendants with respect to Markov dominance.
Theorem 7 (Basins of Attraction and States Without Descendants)

Let \( M \subset [0, 1] \times (\mathbb{G} \times \mathbb{G}) \) be a Markov supernetwork for Markov transition matrix \( P \). The following statements are equivalent:

1. \( \mathcal{A} \subset \mathbb{G} \) is a basin of attraction (a closed and irreducible set of networks).
2. There exists a network without descendants, that is a network \( \mathcal{A} \ni \mathcal{M} \in \mathbb{Z} \), such that

\[
\mathcal{A} = \{ G_h \in \mathbb{G} : G_h \equiv_{\mathcal{M}} G_l \}.
\]

Proof. (1) implies (2): Suppose \( \mathcal{A} \) is a basin of attraction but that

\[
\mathcal{A} \neq \{ G_h \in \mathbb{G} : G_h \equiv_{\mathcal{M}} G_l \} \text{ for all } G_l \in \mathbb{Z}.
\]

Then there is some state \( G_{l'} \in \mathcal{A} \) with a descendant \( G_{h'} \). Thus, \( \rho_{h'l'} > 0 \), but \( \rho_{l'l'} = 0 \). But now we have a contradiction. By the closedness of \( \mathcal{A} \), \( G_h \in \mathcal{A} \) and by irreducibility \( \rho_{h'l'} > 0 \).

(2) implies (1): Suppose that \( \mathcal{A} = \{ G_h \in \mathbb{G} : G_h \equiv_{\mathcal{M}} G_l \} \) for some state \( \mathcal{M} \in \mathbb{Z} \). Then for all pairs of networks \( G_h \) and \( G_l \) in \( \mathcal{A} \), \( \rho_{hl} > 0 \) and \( \rho_{lh} > 0 \), so \( \mathcal{A} \) is irreducible. And for all \( G_h \in \mathcal{A} \) and \( G_l \notin \mathcal{A} \), \( \rho_{hl} = 0 \). In particular, if \( \rho_{hl} > 0 \), then it must be true that \( \rho_{lh} > 0 \), because otherwise \( G_l \) would be a descendant of \( G_h \), which leads to a contradiction. But \( \rho_{hl} > 0 \) and \( \rho_{lh} > 0 \) implies that \( G_l \in \mathcal{A} \), also a contradiction. Thus, \( \mathcal{A} \) is also closed.

First, note that it is easy to show any two basins of attraction are either equal or disjoint. Second, note that for each network without descendants \( G_l \in \mathbb{Z} \),

\[
\{ G_h \in \mathbb{G} : G_h \equiv_{\mathcal{M}} G_l \} \subset \mathbb{Z}.
\]

Finally, note that in light of Theorems 6 and 7, we conclude that the Markov transition \( P \), with corresponding Markov supernetwork \( \mathcal{M} \), generates a unique, finite, disjoint collection of basins of attraction, say \( \{ \mathcal{A}^1, \mathcal{A}^2, \ldots, \mathcal{A}^{L} \} \), where for each \( l = 1, 2, \ldots, L \), there exists \( G_l \in \mathbb{Z} \) such that

\[
\mathcal{A}^l := \{ G_h \in \mathbb{G} : G_h \equiv_{\mathcal{M}} G_l \} \subset \mathbb{Z}.
\]

Note that for networks \( G_l' \) and \( G_l \) in \( \mathbb{Z} \) such that \( G_l' \equiv_{\mathcal{M}} G_l \),

\[
\{ G_h \in \mathbb{G} : G_h \equiv_{\mathcal{M}} G_l' \} = \{ G_h \in \mathbb{G} : G_h \equiv_{\mathcal{M}} G_l \} = \mathcal{A}^l
\]

for some \( l \).

8.3 Invariant and Ergodic Probability Measures: Existence, Characterization and Computation

8.3.1 Invariant and Ergodic Probability Measures

A probability measure \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_L) \in \Delta(\mathbb{G}) \) on the state space of feasible networks

\[
\mathbb{G} := \{ G_1, G_2, \ldots, G_H \}.
\]
is invariant for Markov transition \( P \) (i.e., is \( P \)-invariant) if
\[
\lambda P = \sum_{i \in \mathcal{H}} \langle \lambda_i, (P)_{i} \rangle = \sum_{i \in \mathcal{H}} \langle \lambda_i, p_i \rangle = \lambda. \tag{36}
\]

Thus, if probability measure \( \lambda \) is \( P \)-invariant, then for any set of networks \( G_S \subset G \), if the current status quo network \( G_{t_n} \) is chosen according to probability measure \( \lambda \) so that the probability that \( G_{t_n} \) lies in \( G_S \) is \( \lambda(G_S) := \sum_{t_n \in S} \lambda_{t_n} \), then the probability that any future period’s network \( G_{t_n+m} \) lies in \( G_S \) is also \( \lambda(G_S) := \sum_{t_n+m \in S} \lambda_{t_n+m} \). Denote by \( \mathcal{I} \) the collection of all \( P \)-invariant measure.

Let \( \mathcal{A} \) denote the collection of all basins of attraction (i.e., all closed, irreducible sets). A \( P \)-invariant measure \( \lambda \) is said to be \( P \)-ergodic if \( \lambda(\mathcal{A}) = 0 \) or \( \lambda(\mathcal{A}) = 1 \) for all basins of attraction \( \mathcal{A} \in \mathcal{A} \). Denote by \( \mathcal{E} \) the collection of all \( P \)-ergodic measures. Because the \( P \)-ergodic probability measures are the extreme points of the (possibly empty) convex set \( \mathcal{I} \) of \( P \)-invariant measures (see Theorem 19.25 in Aliprantis and Border, 2006), each measure \( \lambda \) in \( \mathcal{I} \) can be written as a convex combination of the measures in \( \mathcal{E} \).

The set of invariant probability measures \( \mathcal{I} \) for equilibrium Markov transition \( P \) with state space \( G := \{G_1, G_2, \ldots, G_l\} \) is given by the set
\[
\mathcal{I} = \{ \lambda \in \Delta(G) : \lambda P = \lambda \}.
\]

We will show that if
\[
\mathcal{A} := \{ \mathcal{A}^1, A^2, \ldots, A^L \}
\]
is the unique, finite, disjoint collection of basins of attraction generated by \( P \), then corresponding to each basin of attraction \( A^l \) there is unique ergodic probability measure \( \alpha^l(\cdot) \) concentrated on \( A^l \) such that \( \alpha^l(G) > 0 \) for all \( G \in A^l \). Moreover, we will show that the set of all ergodic probability measures for transition \( P \) is given by
\[
\mathcal{E} = \{ \alpha^1(\cdot), \ldots, \alpha^L(\cdot) \}.
\]

and that \( \mathcal{I} = co \mathcal{E} \). Thus, each \( P \)-invariant probability measure \( \lambda \in \mathcal{I} \) is a convex combination of the ergodic measures in \( \mathcal{E} \). Finally, we show how to compute these ergodic measures.

Each network contained in a basin of attraction is recurrent, that is,
\[
\rho_{il} := \Pi \left\{ \tau_{G_i} < \infty | \tilde{G}^0 = G_l \right\} = 1.
\]

We say that a recurrent network \( G_l \) is **positive recurrent** if \( E_{G_l} \tau_{G_l} < \infty \) and null recurrent if \( E_{G_l} \tau_{G_l} = \infty \). It is easy to show that all recurrent states of a finite state Markov chain are positive recurrent.

The following results are variations on classical results for finite state Markov chains (see for example, Durrett 2005, Kemeny, Snell, and Knapp 1976, or Norris 1997).
Theorem 8 (Ergodicity and Invariance Results for Probability Measures and Functions for Finite State Markov Processes)

Let $P$ be the Markov transition on state space $\mathbb{G} := \{G_1, G_2, \ldots, G_H\}$ with basins of attraction

$$\mathcal{A} = \{A_1, A_2, \ldots, A_L\}.$$ 

The following statements are true:

1. For each basin of attraction $A_i$ there is unique ergodic probability measure $\alpha_i(\cdot)$ with support contained in $A_i$ such that $\alpha_i(G_i) > 0$ for all $G_i \in A_i$. Moreover, for each $G_i \in A_i$,

$$\alpha_i(G_i) = \frac{1}{\mathbb{E}_{G_i} \tau_{G_i}}.$$ 

2. The set of all $P$-ergodic probability measures is given by $\mathcal{E} = \{\alpha_1(\cdot), \ldots, \alpha_L(\cdot)\}$ and $\mathcal{I} = \text{co} \mathcal{E}$, where $\text{co}$ denotes convex hull.

3. For each basin of attraction $A_i$ and each initial state $G_i_0 \in A_i$

$$\Pi \left\{ \lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} f(\tilde{G}^m) = f_{A_i} | \tilde{G}^0 = G_i_0 \right\} = 1,$$

and

$$\Pi \left\{ \lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} I\{\tilde{G}^m = G_i_0\} = \alpha_i(G_i_0) | \tilde{G}^0 = G_i_0 \right\} = 1.$$

Here, $f_{A_i} := \sum_{G_h \in A_i} f(G_h) \alpha^*_i(G_h)$ is the expected value of the function (random variable) $f(\cdot)$ on the basin of attraction $A_i$ with respect to the ergodic probability measure $\alpha^*_i(\cdot)$ concentrated on $A^*_i$, while the random variable

$$\frac{1}{n} \sum_{m=0}^{n-1} I\{\tilde{G}^m = G_i_0\}$$

is the average amount of time (average number of moves) the processes spends in state $G_i_0 \in A_i$ before $n$.

Proof. (1) Suppose basin of attraction $A_i$ is given by

$$A_i = \{G_i_1, \ldots, G_i_{N_i} \} \subset \mathbb{G}.$$ 

Let $P_i$ be the $N_i \times N_i$ submatrix of the Markov transition matrix $P$. The matrix $P_i$ has typical entry $(P_i)_{i,j} = p(G_i_j | G_i_l)$, i.e., $(P_i)_{i,j}$ is the probability that nature moves from network $G_i_l$ to network $G_i_j$. Because the basin of attraction $A_i$ is closed and irreducible, $P_i$ is also a Markov transition matrix for the process confined to $A_i$.

31
Recall that if the process begins in $A^l$ it will stay in $A^l$. In particular, by part (1) of Theorem 5 that for all $G_{l_t} \in A^l$ and for all $n$,

$$\Pi \{ \tilde{G}^n \in A^l | \tilde{G}^0 = G_{l_t} \} = 1.$$ 

Hence the process on $A^l$ is irreducible and every state in $A^l$ is positive recurrent. Thus, part (1) follows from Theorem 1.7.7 in Norris (1997).

(2) By Theorem 3.2.10 in Strook (2005), because $G$ contains positive recurrent states (for example any state in a basin of attraction), the set of invariant probability measures $I$ is nonempty and $I$ is clearly convex and compact. Thus by the finite-dimensional Krein-Milman Theorem (Aliprantis and Border 2009, p 297) $I$ is the convex hull of its extreme points. It only remains to show that each ergodic probability measure in $E = \{ \alpha^1(\cdot), \ldots, \alpha^L(\cdot) \}$ is an extreme point of $I$ and that $E$ contains all the extreme points of $I$. But these conclusions are an immediate consequence of Theorem 3.2.10 in (Strook 2005).

(3) Because the process on $A^l$ is irreducible and every state in $A^l$ is positive recurrent, part (3) follows from Theorem 1.10.2 in Norris (1997).

### 8.3.2 Computing Ergodic Probability Measures

We conclude this section by showing how to compute the unique ergodic probability measure $\alpha^l(\cdot)$ corresponding to any basin of attraction $A^l$. Because each basin is closed and irreducible, consisting entirely of positive recurrent states, the process confined to $A^l$ and governed by Markov transition matrix $P_l$ is ergodic.

Here we follow the approach introduced in the classic book by Kemeny, Snell, and Knapp 1976 on denumerable Markov chains. To begin, let $e_l$ be a vector of ones in $\mathbb{R}^{N_l}$, that is, let

$$e_l = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{N_l}.$$ 

Also, let $\mu_l$ be any probability vector in $\mathbb{R}^{N_l}$, that is, let

$$\mu_l = (\mu_1, \ldots, \mu_{N_l})$$

where $\mu_j \geq 0$ for all $j \in \{1, \ldots, N_l\}$ and $\sum_{j=1}^{N_l} \mu_j = 1$. Finally, let $Z_{\mu_l}$ be the $N_l \times N_l$ matrix given by

$$Z_{\mu_l} = (I - P_l + e_l \mu_l)^{-1}$$

where

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \mu_1 & \ldots & \mu_{N_l} \end{pmatrix} = \begin{pmatrix} \mu_1 & \ldots & \mu_{N_l} \\ \vdots \\ \mu_1 & \ldots & \mu_{N_l} \end{pmatrix}_{N_l \times N_l}.$$ 

Then, the ergodic probability measure

$$\alpha^l := (\alpha^l_1, \ldots, \alpha^l_{N_l}) := (\alpha^l(G_{l_1}), \ldots, \alpha^l(G_{l_{N_l}}))$$

is
is given by \[ \mu_l Z_{\mu_l} = \alpha^l. \]

What is interesting is that \( \mu_l Z_{\mu_l} = \alpha^l \) for all probability vectors \( \mu_l \).

Suppose the Markov transition matrix is

\[
P = \begin{pmatrix}
.5 & .5 & 0 & 0 & 0 & 0 \\
.13447 & .13447 & 0 & 0 & .36553 & .36553 \\
0 & 0 & .5 & .5 & 0 & 0 \\
0 & 0 & .0596 & .0596 & .4404 & .4404 \\
.4404 & .4404 & 0 & 0 & .0596 & .0596 \\
0 & 0 & 0 & 0 & .5 & .5
\end{pmatrix},
\]

and generates one basin of attraction, \( A^l = \{G_1, G_2, G_5, G_6\} \). The states in \( T = \{G_3, G_4\} \) are transient.

First, we will compute \( \mu := (\alpha_1, \alpha_2, \alpha_5, \alpha_6) := (\alpha(G_1), \alpha(G_2), \alpha(G_5), \alpha(G_6)) \).

Arbitrarily choosing \( \mu_l = (1, 0, 0, 0) \), we have

\[
Z_{\mu_l} = (I - P_l + e\mu)^{-1}
\]

\[
= \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} - \begin{pmatrix}
.5 & .5 & 0 & 0 \\
.13447 & .13447 & .36553 & .36553 \\
.4404 & .4404 & .0596 & .0596 \\
0 & 0 & .5 & .5
\end{pmatrix} + \begin{pmatrix}
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}
\]

\[
= \begin{pmatrix}
.27322 & .27322 & .22678 & .22678 \\
-1.18033 & .81967 & .68033 & .68033 \\
-.80073 & .19927 & 1.30073 & .30073 \\
-1.34718 & -.34718 & .84718 & 1.84718
\end{pmatrix},
\]

where \( P_l \) is the submatrix, gotten from deleting the third and fourth rows and columns of \( P \). Therefore,

\[
\mu Z_{\mu_l} = \begin{pmatrix}
1 & 0 & 0 & 0
\end{pmatrix} \begin{pmatrix}
.27322 & .27322 & .22678 & .22678 \\
-1.18033 & .81967 & .68033 & .68033 \\
-.80073 & .19927 & 1.30073 & .30073 \\
-1.34718 & -.34718 & .84718 & 1.84718
\end{pmatrix}
\]

\[
= \begin{pmatrix}
.27322 & .27322 & .22678 & .22678
\end{pmatrix}
\]

\[
= (\alpha_1, \alpha_2, \alpha_5, \alpha_6).
\]

Checking, we have

\[
\alpha P_l = \begin{pmatrix}
.27322 & .27322 & .22678 & .22678
\end{pmatrix} \begin{pmatrix}
.5 & .5 & 0 & 0 \\
.13447 & .13447 & .36553 & .36553 \\
.4404 & .4404 & .0596 & .0596 \\
0 & 0 & .5 & .5
\end{pmatrix}
\]

\[
= \begin{pmatrix}
.27322 & .27322 & .22678 & .22678
\end{pmatrix}.
\]