ON THE CONVERGENCE OF DISCRETE-TIME STATIONARY REGULAR MULTIVARIATE MARKOV CHAINS

BY WYNN C. STIRLING[1,a] AND DON ROSS[2,b]

[1]*Electrical and Computer Engineering Department, Brigham Young University,* [a]wynnstirling@gmail.com

[2]*School of Society, Politics, and Ethics, University College Cork*
School of Economics, University of Cape Town
Center for Economic Analysis of Risk, Georgia State University , [b]don.ross931@gmail.com

This paper extends stationary Markov chain convergence theory to model discrete-time regular multivariate Markov chains with joint-conditioning transition probability functions that condition the future state of each process on the present states of all processes considered jointly. This result is a formal extension of currently modeled multivariate Markov chains discussed in the literature, which employ marginal-conditioning transition probability functions that condition the future state of each process separately on the present state of each of the complementary processes considered individually. The advantage of this generalization is that it accounts for interdependencies among the members of the complementary sets. We also establish that it is generally not possible to deduce equivalent marginal-conditioning transition probability functions from joint-conditioning probability functions.

Our methodology is to establish the existence of a *Markov equivalent* complementary graphical representation of a multivariate Markov process with the complementary subset of the process as vertices and the edges comprising joint-conditioning/joint-conditioned transition probabilities, to which the Markov chain convergence theorem may be directly applied. This generalization is particularly applicable to scenarios where interrelationships between conditioning sources of causal or informational influence cannot be modeled as individual conditioning functions, and potentially informs designs for domains such as economic experiments, mobile robotics, and other artificial intelligence applications.

1. Introduction.

Multivariate Markov chain theory is an important tool of applied probability and statistics for the study of multivariate phenomena such as biologics [35], finance [31], machine learning [13], robotic surveillance [8], and others, to represent correlations of variables in sequentially ordered data sets, for purposes of modeling. We begin by reviewing univariate Markov chain convergence theory. For a univariate finite-state first-order Markov chain with time-indexed increments $t = \{0, 1, 2, \dots\}$, the Markov Chain Convergence (MCC) theorem establishes conditions for the existence of a limiting stationary, or steady-state, probability distribution, which we state for reference and to establish notation.[1]

THEOREM 1.1 (Markov Chain Convergence). *Let T be a square matrix with nonnegative entries such that each column sums to unity and T is* regular, *meaning that there exists an integer m such that that all elements of T^m are strictly positive. Then there exists a unique mass vector $\overline{\mathbf{p}}$ such that*

- $T\overline{\mathbf{p}} = \overline{\mathbf{p}}$, *that is, $\overline{\mathbf{p}}$ is the eigenvector corresponding to the unique unit eigenvalue of T;*

MSC2020 subject classifications: Primary 60J10, 68R10; secondary 05C20, 05C50.

Keywords and phrases: Markov chain convergence, Markov equivalent graphs, Network theory.

[1]This paper will be restricted to stationary Markov processes (i.e., the transition probabilities are time-invariant).

- $\overline{T} = \lim_{t\to\infty} T^t = \left[\overline{\mathbf{p}} \cdots \overline{\mathbf{p}}\right]$*; and*
- $\overline{\mathbf{p}} = \overline{T}\mathbf{p}(0)$ *for every initial mass vector* $\mathbf{p}(0)$.

For a discussion and proof of this theorem, see, for example, [9, 18, 21]. The practical significance of this theorem is that it provides a closed-form expression for the steady-state mass vector $\overline{\mathbf{p}}$; namely, as the eigenvector corresponding to the unique unit eigenvalue of T.

The matrix T is termed the *state transition matrix*, whose elements are *transition probability functions* $p_{i|j}(a|a')$, which define the conditional probability that the process is in state a at time $t+1$ given that it is in state a' at time t. In matrix notation, this time iteration assumes the form

$$\mathbf{p}(t+1) = T\mathbf{p}(t)\,, \tag{1}$$

where $\mathbf{p}(t)$ is the unconditional probability vector for the process at time t. This model is directly applicable to the MCC theorem, resulting in a unique steady-state probability vector $\overline{\mathbf{p}} = \lim_{t\to\infty} T^t\mathbf{p}(0)$ that is independent of the initial conditions $\mathbf{p}(0)$.

A general n-dimensional time-indexed multivariate Markov process comprises a set of n Markov processes denoted $\{Y_i(t), i = 1,\ldots,n, t = 0,1,\ldots\}$, that are interconnected such that the transition probability of the future state a_i of each member $Y_i(t+1)$ is jointly conditioned on the present states, a'_j, $j = 1,\ldots,n$, of all members $\{Y_j(t),\, j = 1,\ldots,n\}$, considered simultaneously. To illustrate, consider the multivariate process $\{Y_1(t),\, Y_2(t),\, Y_3(t)\}$, where each $Y_i(t)$ corresponds to an agent who randomly chooses between two prizes, denoted $\mathcal{A}_i = \{y_{i_1}, y_{i_2}\}$, for $i \in \{1,2,3\}$, at time t, such that $Y_i(t+1)$'s probability of choosing $a_i \in \mathcal{A}_i$ is conditioned on $\{Y_j(t), Y_k(t)\}$ *jointly* choosing $(a'_j, a'_k) \in \mathcal{A}_j \times \mathcal{A}_k$; that is, according to the transition probability mass function $p_{i|jk}(a_i|a'_j, a'_k)$ for $a_i|a'_j, a'_k \in \mathcal{A}_i|\mathcal{A}_j \times \mathcal{A}_k$ for $i|jk \in \{1|23, 2|31, 3|12\}$. Expressed in matrix notation, this becomes

$$\mathbf{p}_i(t+1) = T_{i|jk}\mathbf{p}_{jk}(t)\,, \tag{2}$$

where the transition matrix $T_{i|jk}$ is populated by the transition functions $p_{i|jk}$, and $\mathbf{p}_{jk}(t)$ is the joint probability mass vector for $\{Y_j(t), Y_k(t)\}$.

Another example of a multivariate Markov chain is a stock market, where the probability of the value of the shares of a given company at a future time is modeled as conditionally dependent on the present values of the shares of several other companies. For such a scenario, real-time tractability requires conditioning on the joint share values of the other companies considered simultaneously, rather than attempting to determine the probability of a given company's state conditioned on the state of each influencing company considered in isolation.

These examples clearly satisfy the Markov condition in that the transition to the state of a given process is conditioned only on the present state (of all processes and not on the past states of any of the individual processes).

However, the iteration defined by (2) is *not* amenable to the direct application of the MCC theorem, since the conditioning is between each individual process $Y_i(t+1)$ at time $t+1$ and $\{Y_j(t), j \neq i\}$, its complementary subset of processes, at time t.

Several researchers have studied the issue of convergence for multivariate Markov chains. An example [17] from this literature discusses multivariate Markov chains convergence for specific population genetics models, but not results for general processes. An approach to modeling the conditional dependence of an individual process on its complementary set that has gained traction in the literature by Ching and coauthors [5, 6, 31, 33, 36] is to express the transition probabilities of each member as a convex combination of individual-process-to-individual-process transition probabilities. (This approach is further discussed in Section 2.1).[2] [4, 12] adopts this same approach to investigate stock exchanges.

[2]The approach draws its inspiration from [24] who developed a model for high-order Markov chains, termed the Raftery mixture transition distribution model, by defining the probability of the future state of a process

Since these models employ an individial-process-to-individual-process transition probability model, they are not as general as a multiple-process-to-individual-process transition probability model. Our approach is to determine the steady-state of a general stationary regular multivariate Markov process using the more general multiple-procsss-to-individual process model. To do so we invoke the mathematical machinery of network theory and graph theory.

The concept of a network applies to a vast diversity of contexts and applications. A network is interpreted as any collection of sources of influence that are interconnected and structured with respect to directions of influence and specifications of influence effects. Domains of study which are commonly modeled as networks include social and economic [10, 14], causal [23], probabilistic [7, 15, 16, 22], distributed control [20], multiagent systems [19, 30], neural [2], and others. Networks are typically represented as graphs. A network graph comprises a set of vertices, depicted as nodes, and a set of directed edges which define the influence linkages between the vertices. In this paper we focus on networks whose vertices are elements of a multivariate Markov chain and whose graphical representations identify the edges linking these vertices with transition probability mass functions.

A key motivation for the paper is to facilitate applications in which sources of influence are bi-directional. Such applications are particularly strongly motivated in economics and social science. For example, experiments studying transmission of economic advice reported in [29] are restricted to stylized settings in which no agent Y_i who influences an agent Y_j can be reciprocally influenced by Y_j. As emphasized by theorists of social influence [28, 34], this is not the standard ecology of interaction among people, or social animals generally; such influence typically involves mutually active convergence to dynamic equilibria.

2. Graphical Representations of Multivariate Markov Chains.

DEFINITION 2.1.

- Let $t \in \{0, 1, \ldots\}$ denote time indexed in unit increments.
- Let $\big\{\{Y_i,(t), i = 1, \ldots, n\}, t = 0, 1, \ldots\big\}$ denote a set of discrete finite-state Markov processes with state spaces $\mathcal{A}_i = \{y_{i_1}, \ldots, y_{i_{N_i}}\}$.
- Let $Y_{-i}(t) = \{Y_i,(t), i = 1, \ldots, n\} \backslash Y_i(t)$ denote Y_i's *complementary subset* and let $\mathcal{A}_{-i} = \bigtimes_{j \neq i} \mathcal{A}_j$ denote Y_i's *complementary state space*.
- Let $a_i \in \mathcal{A}_i$ denote an arbitrary element of $\mathcal{A}_i$ and let $a_{-i} \in \mathcal{A}_{-i}$ denote a vector of arbitrary elements of $\mathcal{A}_{-i}$. We use the "prime" superscript a_i' for states at time t to distinguish them from the "unprimed" notation a_i for states at time $t + 1$.
- We term Y_{-i} the subset of *conditioning* processes for Y_i, which is termed the *conditioned* process.
- Let $p_{i|j} \colon \mathcal{A}_i|\mathcal{A}_i \to [0, 1]$ denote a transition probability mass function such that $p_{i|j}(a_i|a_j')$ is the probability that $Y_i(t + 1)$ is in state $a_i \in \mathcal{A}_i$ at time $t + 1$ given that $Y_j(t)$ is in state $a_j' \in \mathcal{A}_j$ at time t. The set

$$\Big\{\big\{\{Y_i,(t), i = 1, \ldots, n\}, t = 0, 1, \ldots\big\}, \{p_{i|j}, i, j = 1, \ldots, n\}\Big\} \tag{3}$$

is termed a *marginal-conditioning multivariate Markov chain*. We use the term "marginal" in this context advisedly. The standard usage of such terminology applies to the *conditioned* elements of a probability function; that is, the terms on the left side of the conditioning symbol "|", with the *conditioning* elements on the right side of the conditioning

conditioned on the present and a finite number of past states of the same process as a convex combination of the marginal-conditioning transitoin functions for the future state given each of the present and past states considered individually.

symbol. By analogy with the conventional usage of the terms "joint" and "marginal", we say that the conditioning is "joint" if the conditioning side is populated by all members of the complementary set, and it is "marginal" if the conditioning side is populated by only one member of the complementary set at a t time.

- Let $p_{i|-i}$: $\mathcal{A}_i|\mathcal{A}_{-i} \to [0,1]$ denote a transition probability mass function such that $p_{i|-i}(a_i|a'_{-i})$ is the probability that $Y_i(t+1)$ is in state $a_i \in \mathcal{A}_i$ at time $t+1$ given that $Y_{-i}(t)$ is in state $a'_{-i} \in \mathcal{A}_{-i}$ at time t. The set

$$\Big\{ \big\{ \{Y_i,(t), i=1,\dots,n\}, t=0,1,\dots \big\}, \{p_{i|-i}, i=1,\dots,n\} \Big\} \tag{4}$$

is termed a *joint-conditioning non-self-influencing multivariate Markov chain*.

- Let $p_{i|i,-i}$: $\mathcal{A}_i|\mathcal{A}_i \times \mathcal{A}_{-i} \to [0,1]$ denote a transition probability mass function such that $p_{i|i,-i}(a_i|a'_i, a'_{-i})$ is the probability that $Y_i(t+1)$ is in state $a_i \in \mathcal{A}_i$ at time $t+1$ given that $\{Y_i(t), Y_{-i}(t)\}$ is in state $(a'_i, a'_{-i}) \in \mathcal{A}_i \times \mathcal{A}_{-i}$ at time t. The set

$$\Big\{ \big\{ \{Y_i,(t), i=1,\dots,n\}, t=0,1,\dots \big\}, \{p_{i|i,-i}, i=1,\dots,n\} \Big\} \tag{5}$$

is termed a *joint-conditioning self-influencing multivariate Markov chain*.

Although the non-self-influencing model is a special case of the self-influencing model, it is of sufficient practical importance to treat it separately, since it corresponds to the subclass of multivariate Markov processes where each member of the multivariate Markov chain is influenced by, and only by, other members.

A marginal-conditioning model is appropriate for scenarios where it is possible to isolate the probabilistic dependencies between each pair of processes Y_i and Y_j. The joint-conditioning models are appropriate for scenarios where the probabilistic dependency of each Y_i is expressed as the combined influence of all members of its complementary set Y_{-i}. As we shall establish, it is generally not possible to decompose the joint-conditioning transition probability functions $p_{i|-i}$ and $p_{i|i,-i}$ to generate a set of marginal-conditioning transition probability functions $\{p_{i|j}, i,j = 1,\dots,n\}$. Thus, the joint-conditioning models are more general.

Our usage of network theory and graph theory hinges on an important observation, namely, that a graph of a network is *not* the network; rather, it is a *representation* of the network, and representations are not unique. Our approach is to define a so-called *Markov equivalent* representation (i.e., a representation that preserves the conditioning structure defined by the transition probability functions) of a multivariate Markov chain that is amenable to the direct application of the MCC theorem. For ease of presentation we initially restrict our attention to non-self-influencing multivariate Markov chains, and we first consider the special case where the chain may be represented by a ring graph with individual members of the multivariate chain as vertices. We then extend to the general fully connected (but self-influence free) case by defining a *Markov equivalent* ring graph representation with the complementary subsets as vertices. Finally, we extend the theory to scenarios involving self-influence. This result provides a systematic methodology for determining the steady-state probability of an arbitrary stationary regular multivariate Markov chain.

To facilitate our development we introduce two types of graphs for multivariate Markov chains: transition graphs and network graphs. The following development applies to the non-self-influencing case, but obvious modifications can be made for the self-influencing case.

DEFINITION 2.2. Transition Graph: A *(state) transition graph* of a multivariate Markov chain comprises a set of *vertices* $\big\{\{y_{11},\ldots,y_{1N_1}\},\ldots,\{y_{n1},\ldots y_{nN_n}\}\big\}$ consisting of the states that the members of the chain may assume, with *edges* as the transition probability mass functions $p_{i|j}(a_i|a'_j)$ that define the transition probability that $Y_i(t+1)$ is in state a_i, denoted $Y_i(t+1) \vDash a_i$, given that $Y_j(t)$ is in state a'_i at time t, denoted $Y_j(t) \vDash a'_j$. The symbol $a'_j \longrightarrow a_i$ indicates that probability propagates in only one direction — a *directed edge* — from state a'_j to state a_i during the time increment from t to $t+1$.

Network Graph: A *network graph* of a multivariate Markov chain is a graph with the member processes $\{Y_1(t),\ldots,Y_n(t)\}$ as *vertices* with the set of incoming edges to $Y_i(t+1)$ originating from $Y_{-i}(t)$ collectively comprising the transition probability mass function $p_{i|-i}(a_i|a'_{-i})$ that determines the transition probability of $Y_i(t+1) \vDash a_i$ given the joint-conditioning process $Y_{-i}(t) \vDash a'_{-i}$. The symbol $Y_j(t) \longrightarrow Y_i(t+1)$ means that probability propagates in only one direction — a *directed edge* — from $Y_j(t)$ to $Y_i(t+1)$.

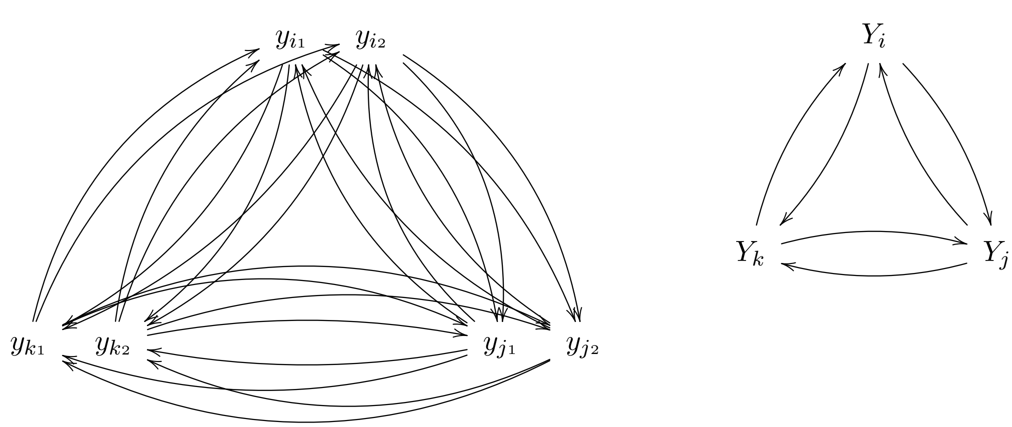
- A *path* from $Y_j(t)$ to $Y_i(t+1)$ is a sequence of directed edges from $Y_j(t)$ to $Y_i(t+1)$, denoted $Y_j(t) \mapsto Y_i(t+1)$.
- A *self-loop* for $Y_i(t)$ is an edge $Y_i(t) \to Y_i(t+1)$.
- A *fully connected network* is a network such that there is an edge between $\{Y_i(t), Y_j(t+1)\}$, $i,j = 1,\ldots,n$ (excluding self-loop edges for the non-self-influencing case).

For our treatment, all edges of both transition and network graphs are directed.

To keep our development as straightforward as possible, we initially do not allow self-influence, that is, transition probability mass functions of the form $p_{i|i,-i}$. Such a model would correspond to situations where the future probability of each member of the multivariate Markov chain would be modulated by its own present state as well as the present states of other members of the multivariate chain. However, the model $p_{i|-i}$ introduces sufficient complexity to identify and address the key issues with multivariate Markov chain convergence, and is simpler to deal with. Once that theory is established, we go on to consider self-influence.

To illustrate the distinction between a transition graph and a network graph, consider the two graphs for a 3×2 network (i.e., a network comprising three members with each being in one of two states):

(6)



representing a multivariate Markov chain network with the transition graph on the left and the network graph on the right (with time arguments suppressed).

2.1. *Transition Graphs.* For a transition graph representation of a multivariate Markov chain as in (6), the transition from state $Y_j(t) \models a'_j$ to state $Y_i(t+1) \models a_i$ and from state $Y_k(t) \models a'_k$ to state $Y_i(t+1) \models a_i$ is governed by marginal-conditioning transition probability functions $p_{i|j}(a_i|a'_j)$, and $p_{i|k}(a_i|a'_k)$, respectively, as indicated by the transition graph fragment

$$\begin{tikzcd} a'_j \arrow[dr, "p_{i|j}"'] & & a'_k \arrow[dl, "p_{i|k}"] \\ & a_i & \end{tikzcd} . \tag{7}$$

Restricting attention to the non-self-influencing case, the transition probability that $Y_i(t+1) \models a_i$ is conditioned on two sources: $Y_j(t)$ and $Y_k(t)$. To account for these multiple sources of conditioning, the approach taken by Ching and coauthors (see [5, 6, 31, 33, 36]) is to model $p_i(a_i, t+1)$ as a convex combination of the transitions from $Y_j(t)$ and $Y_k(t)$ to $Y_i(t+1)$, yielding

$$p_i(a_i, t+1) = \lambda_{ij} \sum_{a'_j} p_{i|j}(a_i|a'_j)\, p_j(a'_j, t) + \lambda_{ik} \sum_{a'_k} p_{i|k}(a_i|a'_k)\, p_k(a'_k, t) \tag{8}$$

with $\lambda_{ij}, \lambda_{ik} \geqslant 0$ and $\lambda_{ij} + \lambda_{ik} = 1$. Expressed using matrix notation, this becomes

$$\mathbf{p}_i(t+1) = \lambda_{ij} T_{i|j} \mathbf{p}_j(t) + \lambda_{ik} T_{i|k} \mathbf{p}_k(t) \; . \tag{9}$$

with

$$T_{i|j} = \begin{bmatrix} p_{i|j}(y_{i_1}|y_{j_1}) & p_{i|j}(y_{i_1}|y_{j_2}) \\ p_{i|j}(y_{i_2}|y_{j_1}) & p_{i|j}(y_{i_2}|y_{j_2}) \end{bmatrix} \qquad T_{i|k} = \begin{bmatrix} p_{i|k}(y_{i_1}|y_{k_1}) & p_{i|k}(y_{i_1}|y_{k_2}) \\ p_{i|k}(y_{i_2}|y_{k_1}) & p_{i|k}(y_{i_2}|y_{k_2}) \end{bmatrix} . \tag{10}$$

Combining equations of the form (9) for the three members of a 3×2 non-self-influencing network yields

$$\underbrace{\begin{bmatrix} \mathbf{p}_i(t+1) \\ \mathbf{p}_j(t+1) \\ \mathbf{p}_k(t+1) \end{bmatrix}}_{\boldsymbol{P}\ (t+1)} = \underbrace{\begin{bmatrix} 0 & \lambda_{ij} T_{i|j} & \lambda_{ik} T_{i|k} \\ \lambda_{ji} T_{j|i} & 0 & \lambda_{jk} T_{j|k} \\ \lambda_{ki} T_{k|i} & \lambda_{kj} T_{k|j} & 0 \end{bmatrix}}_{\mathbf{T}} \underbrace{\begin{bmatrix} \mathbf{p}_i(t) \\ \mathbf{p}_j(t) \\ \mathbf{p}_k(t) \end{bmatrix}}_{\boldsymbol{P}\ (t)} , \tag{11}$$

and the iteration model

$$\boldsymbol{P}\ (t+1) = \mathbf{T} \boldsymbol{P}\ (t) \tag{12}$$

is now amenable to the direct application of the MCC theorem. This approach provides a solution for the steady-state probability of a multivariate Markov chain by approximating the joint-conditioning transition probability of the chain with a convex combination of the marginal-conditioning transition probabilities of each member of the chain.

2.2. *Network Graphs.* Section 2.1 provides a solution for multivariate Markov chain applications with transitions defined by marginal-conditioning functions but, as we claim in Section 1, it is generally not possible to decompose a joint-conditioning transition probability function $p_{i|-i}$ into a set of equivalent marginal-conditioning transition probability functions $\{p_{i|j}, j = 1, \ldots, n\}$. The simplist way to establish this claim is to offer a counterexample.

Consider a three-individual community comprising Isabel (I), John (J), and Karl (K), who are at a dance. John and Karl are strangers to each other, and we may assume that their behaviors are independent. Both men wish to ask Isabel to dance, but they also wish to avoid conflict, so neither will ask her to dance if the other intends to do so. Their action sets are $\mathcal{A}_J = \mathcal{A}_K = \{a, na\}$, that is, to ask (a) or to not ask (na). Isabell's action set is $\mathcal{A}_I = \{w, nw\}$, that is, to be willing (w) or not willing (nw) to accept an invitation to dance. The objective of avoiding conflict between J and K is fixed by setting the joint-conditioning function $p_{J|KI}(a|a, w) = 0$. Now suppose it were also true that J and K are conditionally independent, given I, which would mean that joint-conditioning function $p_{J|KI}(a|a, w) = 0$ reduces to the marginal-conditioning function $p_{J|I}(a|w) = 0$. But this would mean that J would never dance with I, which would defeat J's reason to attend the dance. Thus, we may safely assume that $p_{J|I}(a|w) > 0$. which would then violate the assumption that J and K are conditionally independent given I. The reason for this result is that the marginal-conditioning model does not account for the existence of conflictual relationships between the objectives of J and K. This counterexample establishes that it is generally not possible to derive equivalent marginal-conditioning transition probability functions $p_{i|j}$ from the joint-conditioning transition probability function $p_{i|-i}$.

A joint-conditioning graphical representation of this scenario is

$$
\begin{tikzcd}
Y_{J}(t) \arrow[dr] & p_{I|JK} & Y_{K}(t) \arrow[dl] \\
& Y_{I}(t+1) &
\end{tikzcd}
\tag{13}
$$

with marginal probability mass function

$$
p_{I}(a_{I}, t+1) = \sum_{a'_J a'_K} p_{I|JK}(a_{I}|a'_{J}, a'_{K})\, p_{J}(a'_{J}, t)\, p_{K}(a'_{K}, t)\,, \tag{14}
$$

where $p_{J}(a'_{J}, t)$ and $p_{K}(a'_{K}, t)$ are marginal probability mass functions for J and K, respectively.

Extending to the multivarite-conditioning case for a fully connected three-member Markov chain with network graph illustrated in (6), a corresponding network fragment is

$$
\begin{tikzcd}
Y_j(t) \arrow[d, "p_{j|jk}"] \arrow[dr] & Y_k(t) \arrow[d, "p_{i|jk}"'] \arrow[dl] \\
Y_j(t+1) & Y_i(t+1)
\end{tikzcd}
\;. \tag{15}
$$

Since it is not generally possible to decompose a joint-conditioning transition probability function into an equivalent set of marginal-conditioning transition probability functions, we cannot decouple the influence $Y_j(t)$ exerts on $Y_j(t+1)$ and $Y_i(t+1)$ from the influence $Y_k(t)$ exerts on $Y_j(t+1)$ and $Y_i(t+1)$, and so must consider the joint influence that the subset $\{Y_j(t), Y_k(t)\}$ exerts on the subset $\{Y_j(t+1), Y_i(t+1)\}$. A natural way to do this is to treat $\{Y_j(t), Y_k(t)\}$ and $\{Y_i(t+1), Y_j(t+1)\}$ as *dyadic vertices* where we drop the braces and the separating comma and express these subgroups of processes as units denoted $Y_jY_k(t)$ and $Y_iY_j(t+1)$, that canot be separated into their constituant parts as far as conditioning is concerned. We thus re-express (15) as

$$
Y_jY_k(t) \xrightarrow[p_{ij|jk}]{\qquad\qquad} Y_iY_j(t+1) \tag{16}
$$

where $p_{ij|jk}$ is a *joint-conditioning/joint-conditioned* transition function, yet to be defined, that governs the transition to $Y_iY_j(t+1)$ from $Y_jY_k(t)$. The corresponding marginal probability mass function for the dyad $Y_iY_j(t+1)$ is

$$p_{ij}(a_i, a_j, t+1) = \sum_{a'_j, a'_k} p_{ij|jk}(a_i, a_j | a'_j, a'_k) p_{jk}(a'_j, a'_k, t)\,. \tag{17}$$

We may express (17) using matrix notation by defining a *complementary-state-to-complementary-state transition matrix* $T_{ij|jk}$, whose entries are the transition probability mass functions $p_{ij|jk}$, yielding

$$\mathbf{p}_{ij}(t+1) = T_{ij|jk}\mathbf{p}_{jk}(t) \tag{18}$$

with

$$\mathbf{p}_{ij}(t) = \begin{bmatrix} p_{ij}(y_{i_1}, y_{j_1}, t) \\ p_{ij}(y_{i_1}, y_{j_2}, t) \\ p_{ij}(y_{i_2}, y_{j_1}, t) \\ p_{ij}(y_{i_2}, y_{j_2}, t)) \end{bmatrix} \qquad \mathbf{p}_{jk}(t+1) = \begin{bmatrix} p_{jk}(y_{j_1}, y_{k_1}, t+1) \\ p_{jk}(y_{j_1}, y_{k_2}, t+1) \\ p_{jk}(y_{j_2}, y_{k_1}, t+1) \\ p_{jk}(y_{j_2}, y_{k_2}, t+1) \end{bmatrix} \tag{19}$$

and

$$T_{ij|jk} = \begin{bmatrix} p_{ij|jk}(y_{i1}y_{j1}|y_{j1}, y_{k1}) & p_{ij|jk}(y_{i1}, y_{j1}|y_{j1}, y_{k2}) & p_{ij|jk}(y_{i1}, y_{j1}|y_{j2}, y_{k1}) & p_{ij|jk}(y_{i1}, y_{j1}|y_{j2}, y_{k2}) \\ p_{ij|jk}(y_{i1}y_{j2}|y_{j1}, y_{k1}) & p_{ij|jk}(y_{i1}, y_{j2}|y_{j1}, y_{k2}) & p_{ij|jk}(y_{i1}, y_{j2}|y_{j2}, y_{k1}) & p_{ij|jk}(y_{i1}, y_{j2}|y_{j2}, y_{k2}) \\ p_{ij|jk}(y_{i2}y_{j1}|y_{j1}, y_{k1}) & p_{ij|jk}(y_{i2}, y_{j1}|y_{j1}, y_{k2}) & p_{ij|jk}(y_{i2}, y_{j1}|y_{j2}, y_{k1}) & p_{ij|jk}(y_{i2}, y_{j1}|y_{j2}, y_{k2}) \\ p_{ij|jk}(y_{i2}y_{j2}|y_{j1}, y_{k1}) & p_{ij|jk}(y_{i2}, y_{j2}|y_{j1}, y_{k2}) & p_{ij|jk}(y_{i2}, y_{j2}|y_{j2}, y_{k1}) & p_{ij|jk}(y_{i2}, y_{j2}|y_{j2}, y_{k2}) \end{bmatrix}. \tag{20}$$

Our task is to define $p_{ij|jk}$ in a way that preserves the conditioning struture. To do so, it is convenient to first consider ring graphs and then extend to the full-connected case.

3. Graphical Topologies. Unlike the iteration model defined by (1), the model defined by (18) is *not* in the form required for the direct application of the MCC theorem. As mentioned previously, our goal is to rectify this problem by defining a Markov equivalent representation to which the MCC theorem does apply. We first consider a special case, where the members of the chain form a ring, and then use those results to establish the fully connected case where the probability of the future state of each member is conditioned on the present states of all other members.

3.1. *Ring Graph.* A ring comprises an n-member multivariate Markov process $\{Y_1(t), \ldots, Y_n(t+n\delta)\}$, that transitions from state $Y_i(t+i\delta) \models a'_i$ to the state $Y_{i+1}(t+(i+1)\delta) \models a_{i+1} \mod n$, where $\delta = \frac{1}{n}$ is the fractional time increment to transition from one state to another, that form the path

$$Y_1(t) \to Y_2(t+\delta) \to Y_3(t+2\delta) \to \cdots \to Y_n(t+(n-1)\delta) \to Y_1(t+1), \tag{21}$$

yielding the network graph

$$\begin{tikzcd} Y_1(t) \arrow[r, "p_{2|1}"] & Y_2(t+\delta) \arrow[d, "p_{3|2}"] \\ Y_n(t+(n-1)\delta) \arrow[u, "p_{1|n}"] & Y_3(t+2\delta) \arrow[l, dashed] \end{tikzcd} \tag{22}$$

where the cycle is completed at time $t+1$ with $Y_1(t+1)$ replacing $Y_1(t)$ and the cycle continues in a clockwise orientation (the results will also apply to a counterclockwise orientating convention), where each member is influenced by only its predecessor according to the transition probability mass functions $p_{i|j}$ for $i|j \in \{2|1, 3|2, \ldots, n|n-1, 1|n\}$.

Starting at time $t = 0$ and considering the network fragment $Y_1(0) \to Y_2(\delta)$, where $Y_1(0) \models a'_1$ at time $t = 0$ and $Y_2(\delta) \models a_2$ at time $t = \delta$. The joint probability mass function for $\{Y_1(0), Y_2(\delta)\} \models (a'_1, a_2)$ is

$$p_{12}(a'_1, a_2, \delta) = p_{2|1}(a_2|a'_1)\, p_1(a'_1, 0)\,, \tag{23}$$

where $p_1(a'_1, 0)$ is the initial probability mass function for $Y_1(0) \models a'_1$. The marginal probability mass function for $Y_2(\delta) \models a_2$ is

$$p_2(a_2, \delta) = \sum_{a'_1} p_{12}(a'_1, a_2, \delta) = \sum_{a'_1} p_{2|1}(a_2|a'_1)\, p_1(a'_1, 0)\,. \tag{24}$$

Now considering the fragment $Y_2(\delta) \to Y_3(2\delta)$, the joint probability mass function is

$$p_{23}(a'_2, a_3, 2\delta) = p_{3|2}(a_3|a'_2)\, p_2(a'_2, \delta) \tag{25}$$

with the probability mass function for $Y_3(2\delta) \models a_3$ becoming

$$p_3(a_3, 2\delta) = \sum_{a'_2} p_{23}(a'_2, a_3, 2\delta) = \sum_{a'_2} p_{3|2}(a_3|a'_2)\, p_2(a'_2, \delta)\,. \tag{26}$$

Continuing this process, the joint mass function for the fragment $Y_i(i\delta) \to Y_{i+1}(i+1)\delta)$ is

$$p_{i\,i+1}(a'_i, a_{i+1}, (i+1)\delta) = p_{i+1|i}(a_{i+1}|a'_i)\, p_i(a'_i, i\delta)) \tag{27}$$

and the probability mass function function for $Y_{i+1}(i\delta)$ becomes

$$p_{i+1}(a_{i+1}, (i+1)\delta) = \sum_{a'_i} p_{i\,i+1}(a'_i, a_{i+1}, (i+1)\delta) = \sum_{a'_i} p_{i+1|i}(a_{i+1}|a'_i)\, p_i(a'_i, i\delta) \tag{28}$$

for $i = 1, 2, \ldots, \mod n$. Expressed in matrix form, this becomes

$$\mathbf{p}_{i+1}\big((i+1)\delta\big) = T_{i+1|i}\mathbf{p}_i\big(i\delta\big) \tag{29}$$

for $i = 0, 1, \ldots, \mod n$, where

$$\mathbf{p}_i(i\delta) = \begin{bmatrix} p_i\big(y_{i1}, i\delta\big) \\ p_i\big(y_{i2}, i\delta)\big) \\ \vdots \\ p_i\big(y_{iN_i}, i\delta\big) \end{bmatrix} \tag{30}$$

is the probability mass vector for $Y_i(i\delta)$ and

$$T_{i+1|i} = \begin{bmatrix} p_{i+1|i}\big(y_{(i+1)1}|y_{i1}\big) & \cdots & p_{i+1|i}\big(y_{(i+1)1}|y_{iN_i}\big) \\ \vdots & & \vdots \\ p_{i+1|i}\big(y_{(i+1)N_{i+1}}|y_{i1}\big) & \cdots & p_{i+1|i}\big(y_{(i+1)N_{i+1}}|y_{iN_i}\big) \end{bmatrix} \tag{31}$$

is the *network member-to-network member transition matrix*. Expressing this cycle with the linkages represented by the transition matrices yields

$$\begin{tikzcd} Y_1 \arrow[r, "T_{2|1}"] & Y_2 \arrow[d, "T_{3|2}"] \\ Y_n \arrow[u, "T_{1|n}"] & Y_3 \arrow[l, dashed] \end{tikzcd} \quad . \tag{32}$$

Now define the *closed-loop transition matrices*

$$T_i = T_{i|i+n-1}T_{i+n-1|i+n-2}\cdots T_{i+2|i+1}T_{i+1|i} \bmod n\,. \tag{33}$$

Thus, after one cycle, $\mathbf{p}_i(1) = T_i\mathbf{p}_i(0)$, after two cycles, $\mathbf{p}_i(2) = T_i\mathbf{p}_i(1)$, and so on. After t cycles,

$$\mathbf{p}_i(t+1) = T_i\mathbf{p}_i(t) = T_iT_i\mathbf{p}_i(t-1) = \cdots = T_i^{t+1}\mathbf{p}_i(0)\,. \tag{34}$$

We may now apply the Markov chain convergence theorem to generate the steady-state probability vectors, denoted

$$\overline{\mathbf{p}}_i = \lim_{t\to\infty}\mathbf{p}_i(t) = \lim_{t\to\infty}T_i^t\mathbf{p}_i(0)\,, \tag{35}$$

with

$$\overline{\mathbf{p}}_i = \begin{bmatrix} \overline{p}_i(y_{i_1}) \\ \overline{p}_i(y_{i_2}) \\ \vdots \\ \overline{p}_i(y_{i_{N_i}}) \end{bmatrix},\ i = 1,\ldots,n \tag{36}$$

as the eigenvectors corresponding to the unique unit eigenvalues of T_i, $i = 1,\ldots,n$.

3.2. *Fully Connected Graphs.*

3.2.1. *Non-Self-Influence Network Graphs.*

We now extend to the fully connected graph case without self-influence, where the probability of $Y_iY_j(t+1) \vDash (a_i, a_j)$ is conditioned on $Y_jY_k(t) \vDash (a_j', a_k')$ via

$$p_i(a_i, a_j, t+1) = \sum_{a_j'a_k'} p_{ij|jk}(a_i, a_j|a_j', a_k')p_{jk}(a_j', a_k', t) \tag{37}$$

which, expressed in matrix form, is

$$\mathbf{p}_{ij}(t+1) = T_{ij|jk}\mathbf{p}_{jk}(t)\,. \tag{38}$$

As stated earlier, this expression is not in the form as required for (35) and, therefore, the Markov chain convergence theorem cannot be directly applied.

To proceed, we now invoke the critical observation that a graph of a network is only a *representation* of the network, and representations are not unique. Our challenge is to identify a representation of the network to which the MCC theorem applies. Based on our observation that the convergence of a multivariate Markov chain involves the convergence of the probability of the complementary subset, we are motivated to create an alternative graphical representation of a multivariate Markov process as a ring graph with the vertices comprising the complementary subsets and edges as joint transition probability mass functions to be defined in such a way that the transformed graph preserves the conditionality structure of the original graph.

Consider the three-vertex network graph whose directed edges are transition probability mass functions of the form $p_{i|jk}$ for $i|jk \in \{1|23, 2|31, 3|12\}$, where j, k are ordered such that i precedes j precedes k (which now precedes i) with a clockwise rotation, as time progresses, yielding the network graph (termed the original graph)

(39)
$$\begin{tikzcd}
& Y_i \arrow[dl, shift left] \arrow[dr, shift left] & \\
Y_k \arrow[ur, shift left] \arrow[rr, shift left] & & Y_j \arrow[ul, shift left] \arrow[ll, shift left]
\end{tikzcd}$$

and introduce another graph according to the following definition.

DEFINITION 3.1. Given a fully connected network graph with vertices $\{Y_i,\ i = 1, \ldots, n\}$ and edges $\{p_{i|-i},\ i = 1, \ldots, n\}$, the *complementary network graph* is a hub-spoke graph with the hub comprising a ring with vertices $\{Y_{-i},\ i = 1, \ldots, n\}$ and edges $p_{-i|-(i+1)},\ i = 1, \ldots, n\}$ termed the *complementary transition probability mass functions*, and spokes with vertices Y_i and edges $\{p_{i|-i},\ i = 1, \ldots, n\}$.

The complementary network graph corresponding to (39) is[3]

(40)
$$
\begin{tikzcd}
 & & Y_i & & \\
 & & Y_jY_k \arrow[u, "p_{i|jk}"'] \arrow[dr, "p_{ij|jk}"] & & \\
 & Y_kY_i \arrow[ur, "p_{jk|ki}"] \arrow[dl, "p_{j|ki}"'] & & Y_iY_j \arrow[ll, "p_{ki|ij}"] \arrow[dr, "p_{k|ij}"] & \\
Y_j & & & & Y_k
\end{tikzcd}
$$

for $ij|jk \in \{12|23, 23|31, 31|12\}$, with the transitions assuming a clockwise orientation $Y_iY_j(t) \to Y_kY_i(t+\delta) \to Y_jY_k(t+2\delta) \to Y_iY_j(t+1)$ (with $\delta = \frac{1}{3}$).

We have now generated two network graphs (39) and (40). Analogous to the ring graph displayed by (22), the dyadic vertices of the complementary network graph (40) are connected by complementary joint-conditioning/joint-conditioned transition probability mass functions $p_{ij|jk}$, yet to be defined. To proceed, let us focus on the graph fragments (15) and (16), which we repeat here, since this is a critical part of our development. The fragment from the original graph is

(41)
$$
\begin{tikzcd}
Y_j(t) \arrow[d, "p_{j|jk}"] \arrow[dr] & Y_k(t) \arrow[d, "p_{i|jk}"'] \arrow[dl] \\
Y_j(t+1) & Y_i(t+1)
\end{tikzcd}
$$

and the corresponding complementary graph fragment is

(42)
$$
Y_jY_k(t) \xrightarrow[p_{ij|jk}]{} Y_iY_j(t+1) \ .
$$

Our goal is to define $p_{ij|jk}$ such that the conditioning relationships for graph (42) are equivalent to the relationships for the graph (41). Such an equivalence is referred to as *Markov equivalence* [1, 3, 7, 13], meaning that the graphs have the same conditioning structure.

[3] In graph-theoretic parlance, a complementary graph or, more specifically, a complementary edge graph, is a new graph with the same vertices as the original graph, where the edges are complementary, meaning that edges between vertex pairs appear in the complementary graph if and only if there is no edge between the same vertex pairs on the original graph. By contrast, a complementary vertex graph is a new graph whose vertices are the complementary subsets Y_{-i}, rather than Y_i, as with the original graph. Markov equivalent networks and complementary vertex graphs have been employed by [25–27, 32] for application of multiagent Markov chain theory to social influence networks.

To establish Markov equivalence, we proceed by applying the chain rule to obtain the factorization

$$p_{ij|jk}(a_i, a_j|a'_j, a'_k) = p_{j|jki}(a_j|a'_j, a'_k, a_i)\, p_{i|jk}(a_i|a'_j, a'_k)\,. \tag{43}$$

Now consider the probability mass function $p_{j|jki}(a_j|a'_j, a'_k, a_i)$, which we must define. We first stipulate that $Y_j(t+1)$ must not depend on $Y_i(t+1)$, which may be achieved by requiring $p_{j|jki}(a_j|a'_j, a'_k, a_i) = p_{j|jk}(a_j|a'_j, a'_k)$. However, $p_{j|jk}(a_j|a'_j, a'_k)$ still presents a problem since it requires the existence of a self-loop edge $Y_j(t) \rightarrow Y_j(t+1)$ such that $Y_j(t+1 \models a_j$ is influenced by $Y_j(t) \models a'_j$. But this model does not allow self-influence and, therefore, no such self-loop edge can exist. A natural way to deal with this situation is to insert a *degenerate* self-loop edge $Y_j(t) \rightarrow Y_j(t+1)$ into the network that has no net effect, which we may achieve by defining $p_{j|jk}$ as a *degenerate* transition probability mass function that ascribes its entire mass to $a_j = a'_j$, yielding

$$p_{j|jk}(a_j|a'_j, a'_k) = \begin{cases} 1 & \text{if } a_j = a'_j \\ 0 & \text{otherwise.} \end{cases} \tag{44}$$

Substituting (44) into (43) yields

$$p_{ij|jk}(a_i, a_j|a'_j, a'_k) = \begin{cases} p_{i|jk}(a_i|a'_j, a'_k) & \text{if } a_j = a'_j \\ 0 & \text{otherwise,} \end{cases} \tag{45}$$

which ensures that the conditionality structure is preserved and, therefore, the complementary network representation is Markov equivalent.

The *complementary transition matrix* from the dyad $Y_jY_k(t)$ to the dyad $Y_iY_j(t+1)$ is

$$\begin{aligned} T_{ij|jk} = &\begin{bmatrix} p_{ij|jk}(y_{i_1}, y_{j_1}|y_{j_1}, y_{k_1})\; p_{ij|jk}(y_{i_1}, y_{j_1}|y_{j_1}, y_{k_2}) \\ p_{ij|jk}(y_{i_1}, y_{j_2}|y_{j_1}, y_{k_1})\; p_{ij|jk}(y_{i_1}, y_{j_2}|y_{j_1}, y_{k_2}) \\ p_{ij|jk}(y_{i_2}, y_{j_1}|y_{j_1}, y_{k_1})\; p_{ij|jk}(y_{i_2}, y_{j_1}|y_{j_1}, y_{k_2}) \\ p_{ij|jk}(y_{i_2}, y_{j_2}|y_{j_1}, y_{k_1})\; p_{ij|jk}(y_{i_2}, y_{j_2}|y_{j_1}, y_{k_2}) \end{bmatrix}. \\ &\quad \begin{matrix} p_{ij|jk}(y_{i_1}, y_{j_1}|y_{j_2}, y_{k_1})\; p_{ij|jk}(y_{i_1}, y_{j_1}|y_{j_2}, y_{k_2}) \\ p_{ij|jk}(y_{i_1}, y_{j_2}|y_{j_2}, y_{k_1})\; p_{ij|jk}(y_{i_1}, y_{j_2}|y_{j_2}, y_{k_2}) \\ p_{ij|jk}(y_{i_2}, y_{j_1}|y_{j_2}, y_{k_1})\; p_{ij|jk}(y_{i_2}, y_{j_1}|y_{j_2}, y_{k_2}) \\ p_{ij|jk}(y_{i_2}, y_{j_2}|y_{j_2}, y_{k_1})\; p_{ij|jk}(y_{i_2}, y_{j_2}|y_{j_2}, y_{k_2}) \end{matrix} \Bigg] . \end{aligned} \tag{46}$$

Substituting (44) into (65) yields

$$T_{ij|jk} = \begin{bmatrix} p_{i|jk}(y_{i_1}|y_{j_1}, y_{k_1}) & p_{i|jk}(y_{i_1}|y_{j_1}, y_{k_2}) & 0 & 0 \\ 0 & 0 & p_{i|jk}(y_{i_1}|y_{j_2}, y_{k_1}) & p_{i|jk}(y_{i_1}|y_{j_2}, y_{k_2}) \\ p_{i|jk}(y_{i_2}|y_{j_1}, y_{k_1}) & p_{i|jk}(y_{i_2}|y_{j_1}, y_{k_2}) & 0 & 0 \\ 0 & 0 & p_{i|jk}(y_{i_2}|y_{j_2}, y_{k_1}) & p_{i|jk}(y_{i_2}|y_{j_2}, y_{k_2}) \end{bmatrix}. \tag{47}$$

The *closed-loop complementary transition matrices* are

$$T_{ij} = T_{ij|jk}T_{jk|ki}T_{ki|ij}\,. \tag{48}$$

After t cycles,

$$\mathbf{p}_{ij}(t) = T_{ij}^t \mathbf{p}_{ij}(0)\,. \tag{49}$$

Extending this development to the general fully connected case, (49) becomes

$$\mathbf{p}_{-i}(t) = T_{-i}^t \mathbf{p}_{-i}(0)\,, \tag{50}$$

where

$$T_{-i} = T_{-i|-(i+1)} T_{-(i+1)|-(i+2)} \cdots T_{-(i-1)|-i} \bmod\ n\,, \tag{51}$$

which is now in the form required for the application of the MCC theorem to the ring with vertices Y_{-i}, yielding steady-state complementary probability mass functions

$$\overline{\mathbf{p}}_{-i} = \begin{bmatrix} \overline{p}_{-i}(y_{(i+1)1}, \ldots, y_{(i+n)1}) \\ \vdots \\ \overline{p}_{-i}(y_{(i+1)N_{i+1}}, \ldots, y_{(i+n)N_{i+n}}) \end{bmatrix}. \tag{52}$$

Finally, the steady-state probability mass functions are computed via (18), yielding

$$\overline{\mathbf{p}}_i = T_{i|-i} \overline{\mathbf{p}}_{-i}\,. \tag{53}$$

The following special cases illustrate the flexibility and power of the complementary network representation.

Case I: Consider a three-member chain with graphical representation

$$\begin{tikzcd} & Y_1 \arrow[dl, bend left, "p_{2|31}"] & \\ Y_2 \arrow[ur, bend left, "p_{1|2}"] \arrow[rr, bend right, "p_{3|2}"'] & & Y_3 \arrow[ll, bend right] \end{tikzcd} \tag{54}$$

yielding

$$\begin{aligned}
T_{1|23} &= \begin{bmatrix} p_{1|2}(y_{11}|y_{21}) & p_{1|2}(y_{11}|y_{21}) & p_{1|2}(y_{11}|y_{22}) & p_{1|2}(y_{11}|y_{22}) \\ p_{1|2}(y_{12}|y_{21}), & p_{1|2}(y_{12}|y_{21}) & p_{1|2}(y_{11}|y_{22}) & p_{1|2}(y_{22}|y_{22}) \end{bmatrix} \\
T_{2|31} &= \begin{bmatrix} p_{2|31}(y_{21}|y_{31}, y_{11}) & p_{2|31}(y_{21}|y_{31}, y_{12}) & p_{2|31}(y_{21}|y_{32}, y_{11}) & p_{2|31}(y_{21}|y_{32}, y_{12}) \\ p_{2|31}(y_{22}|y_{31}, y_{11}) & p_{2|31}(y_{22}|y_{31}, y_{12}) & p_{2|31}(y_{22}|y_{32}, y_{11}) & p_{2|32}(y_{21}|y_{32}, y_{12}) \end{bmatrix} \\
T_{3|12} &= \begin{bmatrix} p_{3|2}(y_{31}|y_{21}) & p_{3|2}(y_{31}|y_{21}) & p_{3|2}(y_{31}|y_{22}) & p_{3|2}(y_{31}|y_{22}) \\ p_{3|2}(y_{32}|y_{21}) & p_{3|2}(y_{32}|y_{21}) & p_{3|2}(y_{31}|y_{22}) & p_{3|2}(y_{32}|y_{22}) \end{bmatrix}
\end{aligned} \tag{55}$$

and

$$\begin{aligned}
T_{12|23} &= \begin{bmatrix} p_{1|2}(y_{11}|y_{21}) & p_{1|2}(y_{11}|y_{21}) & 0 & 0 \\ 0 & 0 & p_{1|2}(y_{11}|y_{22}) & p_{1|2}(y_{11}|y_{22}) \\ p_{1|2}(y_{12}|y_{21}) & p_{1|2}(y_{12}|y_{21}) & 0 & 0 \\ 0 & 0 & p_{1|2}(y_{12}|y_{22}) & p_{1|2}(y_{12}|y_{22}) \end{bmatrix} \\
T_{23|31} &= \begin{bmatrix} p_{2|31}(y_{21}|y_{31}, y_{11}) & p_{2|31}(y_{21}|y_{31}, y_{12}) & 0 & 0 \\ 0 & 0 & p_{2|31}(y_{21}|y_{32}, y_{11}) & p_{2|31}(y_{21}|y_{32}, y_{12}) \\ p_{2|31}(y_{22}|y_{31}, y_{11}) & p_{2|31}(y_{22}|y_{31}, y_{12}) & 0 & 0 \\ 0 & 0 & p_{2|31}(y_{22}|y_{32}, y_{11}) & p_{2|31}(y_{22}|y_{32}, y_{12}) \end{bmatrix} \\
T_{31|12} &= \begin{bmatrix} p_{3|2}(y_{31}|y_{21}) & p_{3|21}(y_{31}|y_{21}) & 0 & 0 \\ 0 & 0 & p_{3|2}(y_{31}|y_{22}) & p_{3|22}(y_{31}|y_{22}) \\ p_{3|2}(y_{32}|y_{21}) & p_{3|2}(y_{32}|y_{21}) & 0 & 0 \\ 0 & 0 & p_{3|2}(y_{32}|y_{22}) & p_{3|2}(y_{32}|y_{22}) \end{bmatrix}.
\end{aligned} \tag{56}$$

Case II: Consider the three-member network with acyclic network graph

$$
\begin{tikzcd}
 & Y_1 \arrow[dl, "p_{2|1}"'] \arrow[dr] & \\
Y_2 \arrow[rr, "p_{3|12}"] & & Y_3
\end{tikzcd}
\tag{57}
$$

where Y_1 is a root vertex governed by the unconditional probability mass function p_1. Although this network contains no cycles, we may still identify the complementary network graph of the form (40) with transition matrices

$$
\begin{aligned}
T_{1|23} &= \begin{bmatrix} p_1(y_{11})\; p_1(y_{11})\; p_1(y_{11})\; p_1(y_{11}) \\ p_1(y_{12})\; p_1(y_{12})\; p_1(y_{12})\; p_1(y_{12}) \end{bmatrix} \\
T_{2|31} &= \begin{bmatrix} p_{2|1}(y_{21}|y_{11})\; p_{2|1}(y_{21}|y_{12})\; p_{2|1}(y_{21},y_{11})\; p_{2|1}(y_{21}|y_{12}) \\ p_{2|1}(y_{22}|y_{11})\; p_{2|1}(y_{22}|y_{12})\; p_{2|1}(y_{22},y_{11})\; p_{2|1}(y_{22}|y_{12}) \end{bmatrix} \\
T_{3|12} &= \begin{bmatrix} p_{3|12}(y_{31}|y_{11},y_{21})\; p_{3|12}(y_{31}|y_{11},y_{22})\; p_{3|12}(y_{31}|y_{12},y_{21})\; p_{3|2}(y_{31}|y_{12},y_{22}) \\ p_{3|12}(y_{32}|y_{11},y_{21})\; p_{3|12}(y_{32}|y_{11},y_{22})\; p_{3|12}(y_{32}|y_{12},y_{21})\; p_{3|2}(y_{32}|y_{12},y_{22}) \end{bmatrix}
\end{aligned}
\tag{58}
$$

and

$$
\begin{aligned}
T_{12|23} &= \begin{bmatrix} p_1(y_{11}) & p_1(y_{11}) & 0 & 0 \\ 0 & 0 & p_1(y_{11}) & p_1(y_{11}) \\ p_1(y_{12}) & p_1(y_{12}) & 0 & 0 \\ 0 & 0 & p_1(y_{12}) & p_1(y_{12}) \end{bmatrix} \\
T_{23|31} &= \begin{bmatrix} p_{2|1}(y_{21}|y_{11}) & p_{2|1}(y_{21}|y_{11}) & 0 & 0 \\ 0 & 0 & p_{2|1}(y_{21}|y_{12}) & p_{2|1}(y_{21}|y_{12}) \\ p_{2|1}(y_{22}|y_{11}) & p_{2|1}(y_{22}|y_{12}) & 0 & 0 \\ 0 & 0 & p_{2|1}(y_{22}|y_{12}) & p_{2|1}(y_{22}|y_{12}) \end{bmatrix} \\
T_{31|12} &= \begin{bmatrix} p_{3|12}(y_{31}|y_{11},y_{21}) & p_{3|12}(y_{31}|y_{11},y_{22}) & 0 & 0 \\ 0 & 0 & p_{3|12}(y_{31}|y_{12},y_{21}) & p_{31|2}(y_{31}|y_{12},y_{22}) \\ p_{3|12}(y_{32}|y_{11},y_{21}) & p_{3|12}(y_{32}|y_{11},y_{22}) & 0 & 0 \\ 0 & 0 & p_{3|12}(y_{32}|y_{12},y_{21}) & p_{3|12}(y_{32}|y_{12},y_{22}) \end{bmatrix} .
\end{aligned}
\tag{59}
$$

This special case reveals an important feature of complementary network graphs, namely, that the complementarity does not depend on the existence of cycles. Thus, even an acyclic network can be represented with a complementary network graph with appropriately defined linkages.

Case III: Consider the three-agent transition graph in (6) (left graph) with marginal-conditioning transition probabilities $p_{i|j}$ for $i|j \in \{1|2, 1|3, 2|1, 2|3, 3|1, 3|2\}$ rather than the joint-conditioning transition probabilities $p_{i|jk}$ for $i|jk \in \{1|23, 2|31, 3|12\}$. Of course, we can certainly fall back on Ching's approach as defined in Section 2.1, but a precise theoretical justification for that methodology has not yet been produced. An alternative is to incorporate the marginal-conditioning functions into our more general methodology. The general form for the time-updated marginal probability is

$$
p_i(a_i, t+1) = \sum_{a'_j a'_k} p_{i|jk}(a_i|a'_j, a'_k)\, p_{jk}(a'_j, a'_k, t)\,. \tag{60}
$$

However, this is not the model we are given for this case. A possible way to deal with this situation is to introduce the notion of *conditioning independence*, and to stipulate that

Y_j and Y_k are *conditioning independent with respect to* Y_i, which implies that the joint-conditioning function $p_{i|jk}(a_i|a'_j, a'_k)$ can now be factored into the form

$$p_{i|jk}(a_i|a'_j, a'_k) = p_{i|j}(a_i|a'_j)\, p_{i|k}(a_i|a'_k)\,. \tag{61}$$

Conditioning independence is the converse of the conventional notion of conditional independence of Y_j and Y_k given Y_i, which would yield $p_{jk|i}(a_j, a_k|a_i) = p_{j|i}(a_j|a_i)\, p_{k|i}(a_k|a_i)$. The concept of conditioning independence is not found in conventional treatments of probability theory, but it enjoys an intuitive appeal similar to the intuitive appeal of conditional independence, namely, that the conditional probability given a joint event is the product of the conditional probabilities given the marginal events. The notion of conditioning independence of Y_j and Y_k with respect to Y_i requires no assumptions regarding either the independence of Y_j and Y_k or of the conditional independence of Y_j and Y_k given Y_i.

Substituting (61) into (60) yields

$$p_i(a_i, t+1) = \sum_{a_j, a'_k} p_{i|j}(a_i|a'_j)\, p_{i|k}(a_i|a'_k)\, p_{jk}(a'_j, a'_k, t) \tag{62}$$

which, in matrix notation, becomes

$$\mathbf{p}_i(t+1) = T_{i|jk}\mathbf{p}_{jk}\,, \tag{63}$$

which is the same model as (2) with

$$T_{i|jk} = \begin{bmatrix} p_{i|j}(y_{i_1}|y_{j_1})\, p_{i|k}(y_{i_1}|y_{k_1}) & p_{i|j}(y_{i_1}|y_{j_1})\, p_{i|k}(y_{i_1}|y_{k_2}) & p_{i|j}(y_{i_1}|y_{j_2})\, p_{i|k}(y_{i_1}|y_{k_1}) & p_{i|j}(y_{i_1}|y_{j_2})\, p_{i|k}(y_{i_1}|y_{k_2}) \\ p_{i|j}(y_{i_2}|y_{j_1})\, p_{i|k}(y_{i_2}|y_{k_1}) & p_{i|j}(y_{i_2}|y_{j_1})\, p_{i|k}(y_{i_2}|y_{k_2}) & p_{i|j}(y_{i_2}|y_{j_2})\, p_{i|k}(y_{i_2}|y_{k_1}) & p_{i|j}(y_{i_2}|y_{j_2})\, p_{i|k}(y_{i_2}|y_{k_2}) \end{bmatrix}. \tag{64}$$

The complementary transition matrix thus becomes

$$T_{ij|jk} = \begin{bmatrix} p_{i|j}(y_{i_1}|y_{j_1})\, p_{i|k}(y_{i_1}|y_{k_1}) & p_{i|j}(y_{i_1}|y_{j_1})\, p_{i|k}(y_{i_1}|y_{k_2}) & 0 & 0 \\ 0 & 0 & p_{i|j}(y_{i_1}|y_{j_2})\, p_{i|k}(y_{i_1}|y_{k_1}) & p_{i|j}(y_{i_1}|y_{j_2})\, p_{i|k}(y_{i_1}|y_{k_2}) \\ p_{i|j}(y_{i_2}|y_{j_1})\, p_{i|k}(y_{i_2}|y_{k_1}) & p_{i|j}(y_{i_2}|y_{j_1})\, p_{i|k}(y_{i_2}|y_{k_2}) & 0 & 0 \\ 0 & 0 & p_{i|j}(y_{i_2}|y_{j_2})\, p_{i|k}(y_{i_2}|y_{k_1}) & p_{i|j}(y_{i_2}|y_{j_2})\, p_{i|k}(y_{i_2}|y_{k_2}) \end{bmatrix}, \tag{65}$$

which conforms with (48). Thus, the MCC theorem can be applied to (49).

3.2.2. *Self-Influencing Network Graphs.* When self-influence exist, the transition probability mass functions are of the form $p_{i|i,-i}(a_i|a'_i, a'_{-i})$. The corresponding network graph for a three-member network is

$$\begin{tikzcd} & Y_i \arrow[loop above] \arrow[dl, shift left] \arrow[dr, shift left] & \\ Y_k \arrow[loop left] \arrow[ur, shift left] \arrow[rr, shift left] & & Y_j \arrow[loop right] \arrow[ul, shift left] \arrow[ll, shift left] \end{tikzcd}\,. \tag{66}$$

With this model, the transition function for each $Y_i(t+1)$ is conditioned on the states of all processes, including itself, at time t. To keep the discussion as simple as possible, we continue to focus on a three-member fully connected (including self-influence) Markov chain and extend the analysis by considering the hub-spoke ring network graph with the vertices comprising the triads

$$Y_iY_jY_k(t) \to Y_kY_iY_j(t+\delta) \to Y_jY_kY_i(t+2\delta) \to Y_iY_jY_k(t+1) \tag{67}$$

and edges $p_{ijk|jki}$, $p_{kji|ijk}$ and $p_{jki|kij}$, as illustrated by the following graph:

$$\begin{tikzcd}
 & & Y_i & & \\
 & & Y_iY_jY_k \arrow[u, "p_{i|ijk}"'] \arrow[dr, "p_{kij|ijk}"] & & \\
 & Y_jY_kY_i \arrow[ur, "p_{ijk|jki}"] \arrow[dl, "p_{j|jki}"'] & & Y_kY_iY_j \arrow[ll, "p_{jki|kij}"] \arrow[dr, "p_{k|kij}"] & \\
Y_j & & & & Y_k
\end{tikzcd} . \tag{68}$$

We follow the non-self-influencing procedure by applying the chain rule to obtain

$$p_{ijk|jki}(a_i, a_j, a_k|a'_j, a'_k, a'_i) = p_{k|jkiij}(a_k|a'_j, a'_k, a'_i, a_i, a_j)\, p_{ij|jki}(a_i, a_j|a'_j, a'_k, a'_i)\,. \tag{69}$$

A second application of the chain rule yields

$$p_{ij|jki}(a_i, a_j|a'_j, a'_k, a'_i) = p_{j|jkii}(a_j|a'_j, a'_k, a'_i, a_i)\, p_{i|jki}(a_i|a'_j, a'_k, a'_i)\,. \tag{70}$$

To achieve Markov equivalence we require $Y_k(t+1) \vDash a_k$ given $\{Y_j(t), Y_k(t), Y_i(t)\} \vDash (a'_j, a'_k, a'_i)$ to be independent from $\{Y_i(t+1), Y_k(t+1)\} \vDash (a_i, a_j)$, thereby yielding

$$p_{k|jkiij}(a_k|a'_j, a'_k, a'_i, a_i, a_j) = p_{k|jki}(a_k|a'_j, a'_k, a'_i) \tag{71}$$

and, by a similar argument,

$$p_{j|jkii}(a_j|a'_j, a'_k, a'_i, a_i) = p_{j|jki}(a_j|a'_j, a'_k, a'_i)\,. \tag{72}$$

Thus, Markov equivalence is achieved by substituting (71) and (72) into (69), yielding

$$p_{ijk|jki}(a_i, a_j, a_k|a'_j, a'_k, a'_i) = p_{k|jki}(a_k|a'_j, a'_k, a'_i)\, p_{j|jki}(a_j|a'_j, a'_k, a'_k)\, p_{i|j'k'i'}(a_i|a'_j, a'_k, a'_i)\,. \tag{73}$$

Let $\mathbf{p}_{ijk}(t)$ denote the unconditional probability vector at time t for this scenario and define the transition matrix

$$T_{ijk|jki} = \begin{bmatrix} p_{ijk|jki}(y_{i^1}, y_{j^1}, y_{k^1}|y_{j^1}, y_{k^1}, y_{i^1}) & \cdots & p_{ijk|jki}(y_{i^1}, y_{j^1}, y_{k^1}|y_{j^{N_j}}, y_{k^{N_k}}, y_{i^{N_i}}) \\ \vdots & \vdots & \vdots \\ p_{ijk|jki}(y_{i^{N_i}}, y_{j^{N_j}}, y_{k^{N_k}}|y_{j^1}, y_{k^1}, y_{i^1}) & \cdots & p_{ijk|jki}(y_{i^{N_i}}, y_{j^{N_j}}, y_{k^{N_k}}|y_{j^{N_j}}, y_{k^{N_k}}, y_{i^{N_i}}) \end{bmatrix}. \tag{74}$$

for $ijk|jki \in \{123|231, 231|312, 312|123\}$. Then

$$\mathbf{p}_{ijk}(t+1) = \underbrace{T_{ijk|jki}T_{jki|kij}T_{kij|ijk}}_{T_{ijk}}\,\mathbf{p}_{ijk}(t) = T_{ijk}\mathbf{p}_{ijk}(t)\,, \tag{75}$$

where T_{ijk} is the closed-loop transition matrix. We can now apply the MCC theorem to obtain

$$\overline{\mathbf{p}}_{ijk} = \lim_{t\to\infty} T_{ijk}^t \mathbf{p}_{ijk}(0)\,, \tag{76}$$

from which we can deduce the steady-state individual probability vectors

$$\overline{\mathbf{p}}_i = T_{i|ijk}\overline{\mathbf{p}}_{ijk}\,, \tag{77}$$

where $T_{i|ijk}$ is the transition matrix populated by $p_{i|ijk}$. for $i|ijk \in \{1|123, 2|231, 3|312\}$.

4. General Algorithm. The application of network theory to Markov chains requires the creation of a complementary network graph to represent the network. The graph defined by (40) illustrates the role of the complementary subsets, which form a ring to which the Markov chain convergence theorem can be applied to generate steady-state complementary probability mass functions, which can then be used to generate steady-state individual probability mass functions. This methodology for the non self-influence case is summarized by theorem 4.1 below. The methodology for the self-influencing case can be established with the appropriate modifications.

THEOREM 4.1 (Markov equivalence of complementary network representations). *Consider the multivariate Markov process $\{Y_1(t), \ldots, Y_n(t)\}$ originally represented by a graph with the $Y_i(t)$'s as vertices and the transition probability mass functions $p_{i|-i}$ as edges. Let us define the complementary network graph with vertices comprising the complementary subsets and edges defined as*

$$p_{-n|-1}(a_{-n}|a'_{-1}) = p_{1:n-1|2:n}(a_1, \ldots, a_{n-1}|a'_2, \ldots, a'_n) = \begin{cases} p_{1|-1}(a_1|a'_2, \ldots, a'_n) & \text{if } (a_2, \ldots, a_{n-1}) = (a'_2, \ldots, a'_{n-1}) \\ 0 & \text{otherwise,} \end{cases} \tag{78}$$

$$p_{-1|-2}(a_{-1}|a'_{-2}) = p_{2:n|3:1}(a_2, \ldots, a_n|a'_3, \ldots, a'_n, a'_1) = \begin{cases} p_{2|-2}(a_2|a'_3, \ldots, a'_n, a'_1) & \text{if } (a_3, \ldots, a_n) = (a'_3, \ldots, a'_n) \\ 0 & \text{otherwise,} \end{cases} \tag{79}$$

continuing,

$$p_{-(n-1)|-n}(a_{-(n-1)}|a'_{-n}) = p_{n:(n-2)|1:(n-1)}(a_n, a_1, \ldots, a_{n-2}|a'_1, \ldots, a'_{n-2}, a'_{n-1}) = \begin{cases} p_{n|-n}(a_n|a'_1, \ldots, a'_{n-1}) & \text{if } (a_1, \ldots, a_{n-2}) = (a'_1, \ldots, a'_{n-2}) \\ 0 & \text{otherwise,} \end{cases} \tag{80}$$

where indexing is $\bmod n$*; that is,* $3:(n+1) \bmod n \equiv 3:1$ *and so forth. Then the original representation and the complementary representation are Markov equivalent.*

The proof of this theorem is by construction. We begin by establishing the result for $p_{-n|-1} \equiv p_{1:n-1|2:n}$ displayed in (78), with the results for $p_{-1|-2}$ through $p_{-(n-1)|-n}$ established similarly. We first factor the complementary transition probability mass function $p_{-n|-1}$ as

$$p_{1:n-1|2:n}(a_1, \ldots, a_{n-1}|a'_2, \ldots, a'_{n-1}, a'_n) = p_{n-1|1:n-2,2:n}(a_{n-1}|a_1, \ldots, a_{n-2}, a'_2, \ldots, a'_{n-1}, a'_n) \\ p_{1:n-2,2:n}(a_1, \ldots, a_{n-2}, a'_2, \ldots, a'_{n-1}, a'_n) \tag{81}$$

and notice that $p_{n-1|1:n-2,2:n}(a_{n-1}|a_1,\ldots,a_{n-2},a'_2,\ldots,a'_{n-1},a'_n)$ is a degenerate mass function, since it involves the conditioning component a'_{n-1} and the conditioned component a_{n-1}. Thus,

$$p_{n-1|1:n-2,2:n}(a_{n-1}|a_1,\ldots,a_{n-2},a'_2,\ldots,a'_{n-1},a'_n) = \begin{cases} 1 & \text{if } a_{n-1} = a'_{n-1} \\ 0 & \text{otherwise.} \end{cases} \tag{82}$$

We next turn our attention to the complementary transition probability mass function $p_{1:n-2,2:n}$, which may be factored via the chain rule to yield

$$\begin{aligned} &p_{1:n-2,2:n}(a_1,\ldots,a_{n-2},a'_2,\ldots,a'_n) = \\ &\qquad p_{n-2|1:n-3,2:n}(a_{n-2}|a_1,\ldots,a_{n-3},a'_2,\ldots,a'_n)p_{1:n-3,2:n}(a_1,\ldots,a_{n-3},a'_2,\ldots,a'_n)\,, \end{aligned} \tag{83}$$

and we immediately observe that $p_{n-2|1:n-3,2:n}(a_{n-2}|a_1,\ldots,a_{n-3},a'_2,\ldots,a'_n)$ is degenerate since it involves a_{n-2} as a conditioned state and a'_{n-2} as a conditioning state, yielding

$$p_{n-2|1:n-3,2:n}(a_{n-2}|a_1,\ldots,a_{n-3},a'_2,\ldots,a'_n) = \begin{cases} 1 & \text{if } a_{n-2} = a'_{n-2} \\ 0 & \text{otherwise.} \end{cases} \tag{84}$$

Continuing to the final factorization,

$$p_{12}(a_1,a_2|a'_2,\ldots,a'_n) = p_{2|1:n}(a_2|a_1,a'_2\ldots,a'_n)p_{1|2:n}(a_1|a'_2,\ldots,a'_n)\,, \tag{85}$$

where $p_{2|1:n}(a_2|a_1,a'_2\ldots,a'_n)$ is degenerate, yielding

$$p_{2|1:n}(a_2|a_1,a'_2\ldots,a'_n) = \begin{cases} 1 & \text{if } a_2 = a'_2 \\ 0 & \text{otherwise.} \end{cases} \tag{86}$$

Combining all of these factors (omitting arguments),

$$p_{1:n-1|2:n} = p_{n-1|1:n-2,2:n}\,p_{n-2|1:n-3,2:n}\cdots p_{2|1:n}p_{1|2:n} \tag{87}$$

becomes

$$p_{1:n-1|2:n}(a_1,\ldots,a_{n-1}|a'_2,\ldots,a'_n) = \begin{cases} p_{1|2:n}(a_1|a'_2,\ldots,a'_n) & \text{if } (a_2,\ldots,a_{n-1}) = (a'_2,\ldots,a'_{n-1}) \\ 0 & \text{otherwise,} \end{cases} \tag{88}$$

which establishes (78). Establishing (79) and (80) is obtained by similar arguments. The complementary transition matrices are obtained via Theorem 4.1, yielding

$$T_{1:n-1|2:n} = \begin{bmatrix} p_{1:n-1|2:n}(y_{11},\ldots,y_{(n-1)1}|y_{21},\ldots y_{n1}) & \cdots & p_{1:n-1|2:n}(y_{11},\ldots,y_{(n-1)1}|y_{2N_2},\ldots y_{nN_n}) \\ p_{1:n-1|2:n}(y_{11},\ldots,y_{(n-1)2}|y_{21},\ldots y_{n1}) & \cdots & p_{1:n-1|2:n}(y_{11},\ldots,y_{(n-1)2}|y_{2N_2},\ldots y_{nN_n}) \\ \vdots & & \vdots \\ p_{1:n-1|2:n}(y_{1N_1},\ldots,y_{(n-1)N_{n-1}}|y_{21},\ldots y_{n1}) & \cdots & p_{1:n-1|2:n}(y_{1N_1},\ldots,y_{(n-1)N_{n-1}}|y_{2N_2},\ldots y_{nN_n}) \end{bmatrix}, \tag{89}$$

$$T_{2:n|3:1} = \begin{bmatrix} p_{2:n|3:1}(y_{21},\ldots,y_{n1}|y_{31},\ldots,y_{11}) & \cdots & p_{2:n|3:1}(y_{21},\ldots,y_{n1}|y_{3N_3},\ldots,y_{1N_1}) \\ p_{2:n|3:1}(y_{21},\ldots,y_{n2}|y_{31},\ldots,y_{11}) & \cdots & p_{2:n|3:1}(y_{21},\ldots,y_{n2}|y_{3N_3},\ldots,y_{1N_1}) \\ \vdots & & \vdots \\ p_{2:n|3:1}(y_{2N_2},\ldots,y_{nN_n}|y_{31},\ldots,y_{11}) & \cdots & p_{2:n|3:1}(y_{2N_2},\ldots,y_{nN_n}|y_{3N_3},\ldots,y_{1N_1}) \end{bmatrix} \tag{90}$$

and, continuing,

(91)

$$T_{n:(n-2)|1:(n-1)} = \begin{bmatrix} p_{n:(n-2)|1:(n-1)}(y_{n1},\ldots,y_{(n-2)1}|y_{11},\ldots,y_{(n-1)|1}) & \cdots & p_{n:(n-2)|1:(n-1)}(y_{n1},\ldots,y_{(n-2)1}|y_{1N_1},\ldots,y_{(n-1)N_{n-1}}) \\ p_{n:(n-2)|1:(n-1)}(y_{n1},\ldots,y_{(n-2)2}|y_{11},\ldots,y_{(n-1)|1}) & \cdots & p_{n:(n-2)|1:(n-1)}(y_{n1},\ldots,y_{(n-2)2}|y_{1N_1},\ldots,y_{(n-1)N_{n-1}}) \\ \vdots & & \vdots \\ p_{n:(n-2)|1:(n-1)}(y_{nN_n},\ldots,y_{(n-2)N_{n-2}}|y_{11},\ldots,y_{(n-1)|1}) & \cdots & p_{n:(n-2)|1:(n-1)}(y_{nN_n},\ldots,y_{(n-2)N_{n-2}}|y_{1N_1},\ldots,y_{(n-1)N_{n-1}}) \end{bmatrix}.$$

The notation for these matrices is unavoidably complex, and the reader is invited to examine the subsequently introduced four- and five-member networks as an aid for following the generation of the general n-member network. The closed-loop transition matrices are

$$T_{-i} = T_{-i|-(i+1)}T_{-(i+1)|-(i+2)}\cdots T_{-(i-1)|-i} \bmod n. \tag{92}$$

Thus, as time progresses, the complementary mass functions evolve as

$$\mathbf{p}_{-i}(t) = T_{-i}\mathbf{p}_{-i}(t-1) \tag{93}$$

and applying the Markov chain convergence theorem generates the *steady-state complementary probability mass functions* $\overline{\mathbf{p}}_{-i}$ as the eigenvectors corresponding to the unique unit eigenvalues of T_{-i}, denoted

$$\overline{\mathbf{p}}_{-1} = \begin{bmatrix} \overline{p}_{-1}(y_{21},\ldots,y_{n1}) \\ \vdots \\ \overline{p}_{-1}(y_{2N_2},\ldots,y_{nN_n}) \end{bmatrix} \tag{94}$$

$$\overline{\mathbf{p}}_{-2} = \begin{bmatrix} \overline{p}_{-2}(y_{31},\ldots,y_{11}) \\ \vdots \\ \overline{p}_{-2}(y_{3N_3},\ldots,y_{1N_1}) \end{bmatrix} \tag{95}$$

and, continuing,

$$\overline{\mathbf{p}}_{-n} = \begin{bmatrix} \overline{p}_{-n}(y_{11},\ldots,y_{(n-1)1}) \\ \vdots \\ \overline{p}_{-n}(y_{1N_1},\ldots,y_{(n-1)N_{n-1}}) \end{bmatrix}. \tag{96}$$

The steady-state utility mass functions are computed as

$$\overline{\mathbf{p}}_i = \begin{bmatrix} \overline{p}_i(y_{i1}) \\ \vdots \\ \overline{p}_i(y_{iN_i}) \end{bmatrix} = T_{i|-i}\overline{\mathbf{p}}_{-i}\,, \tag{97}$$

with

(98)

$$T_{i|-i} = \begin{bmatrix} p_{i|-i}(y_{i1}|y_{(i+1)1},\ldots,y_{(i+n)1}) & \cdots & p_{i|-i}(y_{i1}|y_{(i+1)N_{i+1}},\ldots,y_{(i+n)N_{i+n}}) \\ p_{i|-i}(y_{i2}|y_{(i+1)1},\ldots,y_{(i+n)1}) & \cdots & p_{i|-i}(y_{i2}|y_{(i+1)N_{i+1}},\ldots,y_{(i+n)N_{i+n}}) \\ \vdots & & \vdots \\ p_{i|-i}(y_{iN_i}|y_{(i+1)1},\ldots,y_{(i+n)1}) & \cdots & p_{i|-i}(y_{iN_i}|y_{(i+1)N_{i+1}},\ldots,y_{(i+n)N_{i+n}}) \end{bmatrix} \bmod n.$$

4.1. *Four-Member Networks.* Consider a four-member network with original graphical representation

$$\begin{tikzcd}
Y_1 \arrow[r, bend left] \arrow[d, bend left] \arrow[dr, shift left] & Y_2 \arrow[l, bend left] \arrow[d, bend left] \arrow[dl, shift left] \\
Y_4 \arrow[u, bend left] \arrow[r, bend left] \arrow[ur, shift left] & Y_3 \arrow[l, bend left] \arrow[u, bend left] \arrow[ul, shift left]
\end{tikzcd} \tag{99}$$

where the transition probabilities are defined by $p_{1|234}$, $p_{2|341}$, $p_{3|412}$, and $p_{4|123}$.

The complementary network representation for the triads

$$\{Y_i(t)Y_j(t)Y_k(t)\} \text{ for } (ijk) \in \{(123),\ (234),\ (341),\ (412)\}\} \tag{100}$$

is

$$\begin{tikzcd}
Y_4 & & Y_3 \\
Y_1Y_2Y_3 \arrow[u, "T_{4|123}"] \arrow[rr, "T_{412|123}"] & & Y_4Y_1Y_2 \arrow[u, "T_{3|412}"'] \arrow[dd, "T_{341|412}"'] \\
& & \\
Y_2Y_3Y_4 \arrow[uu, "T_{123|234}"] \arrow[d, "T_{1|234}"'] & & Y_3Y_4Y_1 \arrow[ll, "T_{234|341}"] \arrow[d, "T_{2|341}"] \\
Y_1 & & Y_2
\end{tikzcd} \tag{101}$$

.

Successively applying the chain rule yields (suppressing arguments)

$$\begin{aligned} p_{ijk|jkl} &= p_{k|ijjkl}\, p_{ij|jkl} \\ &= p_{k|ijjkl}\, p_{j|ijkl}\, p_{i|jkl}\,. \end{aligned} \tag{102}$$

Eliminating redundant conditioning terms yields

$$\begin{aligned} p_{ijk|jkl}(a_i, a_j, a_k|a_j', a_k', a_l') = p_{k|ijkl}(a_k|a_i, a_j', a_k', a_l')\, p_{j|ijkl}(a_j|a_i, a_j', a_k', a_l') \\ p_{i|jkl}(a_i|a_j', a_k', a_l')\,, \end{aligned} \tag{103}$$

Thus, the transition mass functions $p_{k|ijkl}$ and $p_{j|ijkl}$ are degenerate, and must place all of their mass on the conditioning states, yielding

$$p_{k|ijkl}(a_k|a_i, a_j', a_k', a_l') = \begin{cases} 1 & \text{if } a_k = a_k' \\ 0 & \text{otherwise} \end{cases} \tag{104}$$

and

$$p_{j|ijkl}(a_j|a_i, a_j', a_k', a_l') = \begin{cases} 1 & \text{if } a_j = a_j' \\ 0 & \text{otherwise,} \end{cases} \tag{105}$$

which ensures Markov equivalence. Thus,

(106)

$$T_{ijk|jkl} =$$

$$\begin{bmatrix}
p_{i|jkl}(y_{i1}|y_{j1},y_{k1},y_{l1}) & p_{i|jkl}(y_{i1}|y_{j1},y_{k1},y_{l2}) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & p_{i|jkl}(y_{i1}|y_{j1},y_{k2},y_{l1}) & p_{i|jkl}(y_{i1}|y_{j1},y_{k2},y_{l2}) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & p_{i|jkl}(y_{i1}|y_{j2},y_{k1},y_{l1}) & p_{i|jkl}(y_{i1}|y_{j2},y_{k1},y_{l2}) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & p_{i|jkl}(y_{i1}|y_{j2},y_{k2},y_{l1}) & p_{i|jkl}(y_{i1}\|y_{j2},y_{k2},y_{l2}) \\
p_{i|jkl}(y_{i2}|y_{j1},y_{k2},y_{l1}) & p_{i|jkl}(y_{i2}|y_{j1},y_{k2},y_{l2})) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & p_{i|jkl}(y_{i2}|y_{j1},y_{k2},y_{l1})) & p_{i|jkl}(y_{i2}|y_{j1},y_{k2},y_{l2}) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & p_{i|jkl}(y_{i2}|y_{j2},y_{k1},y_{l1})) & p_{i|jkl}(y_{i2}|y_{j2},y_{k1},y_{l2}) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & p_{i|jkl}(y_{i2}|y_{j2},y_{k2},y_{l1}) & p_{i|jkl}(y_{i2}|y_{j2},y_{k2},y_{l2})
\end{bmatrix}.$$

The closed-loop transition matrices are

$$T_{ijk} = T_{ijk|jkl}T_{jkl|kli}T_{kli|lij}T_{lij|ijk} \tag{107}$$

for $(ijkl) \in \{(1234), (2341), (3412), (4123)\}$. The complementary probability mass functions are the eigenvectors $\overline{\mathbf{p}}_{ijk}$ of the unit eigenvalues of T_{ijk} for $ijk \in \{234,\ 341,\ 412,\ 123\}$, from which the steady-state utility functions are obtained, for a 4×2 network, via

(108)

$$T_{i|jkl} = \begin{bmatrix}
p_{i|jkl}(y_{i1}|y_{j1},y_{k1},y_{l1}) & p_{i|jkl}(y_{i1}|y_{j1},y_{k1},y_{l2}) & p_{i|jkl}(y_{i1}|y_{j1},y_{k2},y_{l1}) & p_{i|jkl}(y_{i1}|y_{j1},y_{k2},y_{l2}) & p_{i|jkl}(y_{i1}|y_{j2},y_{k1},y_{l1}) & p_{i|jkl}(y_{i1}|y_{j2},y_{k1},y_{l2}) & p_{i|jkl}(y_{i1}|y_{j2},y_{k2},y_{l1}) & p_{i|jkl}(y_{i1}|y_{j2},y_{k2},y_{l2}) \\
p_{i|jkl}(y_{i2}|y_{j1},y_{k1},y_{l1}) & p_{i|jkl}(y_{i2}|y_{j1},y_{k1},y_{l2}) & p_{i|jkl}(y_{i2}|y_{j1},y_{k2},y_{l1}) & p_{i|jkl}(y_{i2}|y_{j1},y_{k2},y_{l2}) & p_{i|jkl}(y_{i2}|y_{j2},y_{k1},y_{l1}) & p_{i|jkl}(y_{i2}|y_{j2},y_{k1},y_{l2}) & p_{i|jkl}(y_{i2}|y_{j2},y_{k2},y_{l1}) & p_{i|jkl}(y_{i2}|y_{j2},y_{k2},y_{l2})
\end{bmatrix},$$

yielding

$$\overline{\mathbf{p}}_i = T_{i|jkl}\overline{\mathbf{p}}_{jkl} \tag{109}$$

with

$$\overline{\mathbf{p}}_i = \begin{bmatrix} \overline{p}_i(y_{i_1}) \\ \overline{p}_i(y_{i_2}) \end{bmatrix} s \tag{110}$$

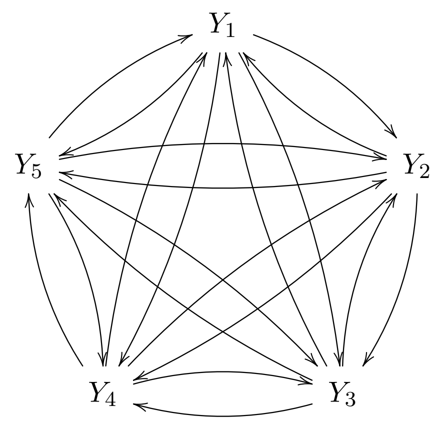
and xx

$$\overline{\mathbf{p}}_{jkl} = \begin{bmatrix}
\overline{p}_{jkl}(y_{j_1}, y_{k_1}, y_{l_1}) \\
\overline{p}_{jkl}(y_{j_1}, y_{k_1}, y_{l_2}) \\
\overline{p}_{jkl}(y_{j_1}, y_{k_2}, y_{l_1}) \\
\overline{p}_{jkl}(y_{j_1}, y_{k_2}, y_{l_2}) \\
\overline{p}_{jkl}(y_{j_2}, y_{k_1}, y_{l_1}) \\
\overline{p}_{jkl}(y_{j_2}, y_{k_1}, y_{l_2}) \\
\overline{p}_{jkl}(y_{j_2}, y_{k_2}, y_{l_1}) \\
\overline{p}_{jkl}(y_{j_2}, y_{k_2}, y_{l_2})
\end{bmatrix} \tag{111}$$

for $i|jkl \in \{1|234,\ 2|341,\ 3|412,\ 4|123\}$.

4.2. *Five-Member Networks.* Consider a five-member network comprising $\{Y_1, Y_2, Y_3, Y_4, Y_5\}$ with original-form graphical representation

(112)



and transition probability mass functions $p_{1|2345}$, $p_{2|3451}$, $p_{3|4512}$, $p_{4|5123}$, and $p_{5|1234}$. The components of $Y_{-i}(t)$ are the quadric elements $Y_i(t)Y_j(t)Y_k(t)Y_l(t) \in \{(1234), (2345), (3451), (4512), (5123)\}$, and the complementary network graph is

$$\begin{tikzcd}
 & & & Y_5 & & & \\
 & & & Y_1Y_2Y_3Y_4 \arrow[u, "T_{5|1234}"] \arrow[drr, "T_{5123|1234}"] & & & \\
Y_1 & Y_2Y_3Y_4Y_5 \arrow[l, "T_{1|2345}"'] \arrow[urr, "T_{1234|2345}"] & & & & Y_5Y_1Y_2Y_3 \arrow[r, "T_{4|5123}"] \arrow[dl, "T_{4512|5123}"] & Y_4 \\
 & & Y_3Y_4Y_5Y_1 \arrow[ul, "T_{2345|3451}"] \arrow[dl, "T_{2|3451}"'] & & Y_4Y_5Y_1Y_2 \arrow[ll, "T_{3451|4512}"] \arrow[dr, "T_{3|4512}"] & & \\
 & Y_2 & & & & Y_3 &
\end{tikzcd} \tag{113}$$

.

Successively applying the chain rule (suppressing arguments) yields

$$\begin{aligned} p_{ijkl|jklm} &= p_{l|ijkjklm}\, p_{ijk|jklm} \\ &= p_{l|ijkjklm}\, p_{k|ijjklm}\, p_{ij|jklm} \\ &= p_{l|ijkjklm}\, p_{k|ijjklm}\, p_{j|ijklm}\, p_{i|jklm}\,. \end{aligned} \tag{114}$$

The functions $p_{l|ijklm}$, $p_{k|ijklm}$, and $p_{j|ijklm}$ are degenerate mass functions, thus

$$p_{ijkl|jklm}(a_i, a_j, a_k, a_l | a'_j, a'_k, a'_l, a_m) = \begin{cases} p_{i|jklm}(a_i|a_j, a_k, a_l, a_m) & \text{if } (a_j, a_k, a_l) = (a'_j, a'_k, a'_l) \\ 0 & \text{otherwise,} \end{cases} \tag{115}$$

which preserves Markov equivalence.

The transition from $Y_2(t)Y_3(t)Y_4(t)Y_5(t)$ to $Y_1(t+1)Y_2(t+1)Y_3(t+1)Y_4(t+1)$ for a 5×2 network is the 16×16 matrix $T_{1234|2345} = [t_{ijkl,mnpq}]$ where

$$\begin{aligned} t_{ijkl,mnpq} &= p_{1234|2345}(y_{1i}, y_{2j}, y_{3k}, y_{4l}|y_{2m}, y_{3n}, y_{4p}, y_{5q}) \\ &= \begin{cases} p_{1|2345}(y_{1i}|y_{2j}, y_{3k}, y_{4l}, y_{5m}) & \text{if } (j,k,l) = (m,n,p) \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{116}$$

is the entry in the $ijkl$th row, $mnpq$th column, with row indexing convention

$$(ijkl) \in \{(1111), (1112), (1121), (1122), (1211), (1212), (1221), (1222), (2111), (2112), (2121), (2122), (2211), (2212), (2221), (2222)\}$$

and column indexing convention

$$(mnpq) \in \{(1111), (1112), (1121), (1122), (1211), (1212), (1221), (1222), (2111), (2112), (2121), (2122), (2211), (2212), (2221), (2222)\}.$$

By similar arguments, the remaining transition matrices are
$T_{2345|4512} = [t_{ijkl,mnpq}]$ with

$$\begin{aligned} t_{ijkl,mnpq} &= p_{2345|4512}(y_{2i}, y_{3j}, y_{4k}, y_{5l}|y_{4m}, y_{5n}, y_{1p}, y_{2q}) \\ &= \begin{cases} p_{2|3451}(y_{2i}|y_{3j}, y_{4k}, y_{5l}, y_{1m}) & \text{if } (j,k,l) = (m,n,p) \\ 0 & \text{otherwise}\,; \end{cases} \end{aligned} \tag{117}$$

$T_{3451|4512} = [t_{ijkl,mnpq}]$ with

$$\begin{aligned} t_{ijkl,mnpq} &= p_{3451|4512}(y_{3i}, y_{4j}, y_{5k}, y_{1l}|y_{4m}, y_{5n}, y_{1p}, y_{2q}) \\ &= \begin{cases} p_{3|4512}(y_{3i}|y_{4m}, y_{5n}, y_{1p}, y_{2q}) & \text{if } (j,k,l) = (m,n,p) \\ 0 & \text{otherwise}\,; \end{cases} \end{aligned} \tag{118}$$

$T_{4512|5123} = [t_{ijkl,mnpq}]$ with

$$\begin{aligned} t_{ijkl,mnpq} &= p_{4512|5123}(y_{4i}, y_{5j}, y_{1k}, y_{2k}|y_{5m}, y_{1n}, y_{2p}, y_{3q}) \\ &= \begin{cases} p_{4|5123}(y_{4i}|y_{5m}, y_{1n}, y_{2p}, y_{3q}) & \text{if } (j,k,l) = (m,n,p) \\ 0 & \text{otherwise}\,; \end{cases} \end{aligned} \tag{119}$$

and $T_{5123|1234} = [t_{ijkl,mnpq}]$ with

$$t_{ijkl,mnpq} = p_{5123|1234}(y_{5i}, y_{1j} y_{2k}, y_{3k} | y_{1m}, y_{2n}, y_{3p}, y_{4q}) \tag{120}$$
$$= \begin{cases} p_{5|1234}(y_{5i} | y_{1m}, y_{2n}, y_{3p}, y_{4q}) & \text{if } (j,k,l) = (m,n,p) \\ 0 & \text{otherwise}\,. \end{cases}$$

The closed-loop transition matrices are

$$T_{ijkl} = T_{ijkl|jklm} T_{jklm|klmi} T_{klmi|lmij} T_{lmij|mijk} T_{mijk|ijkl} \tag{121}$$

for $ijkl \in \{1234, 2345, 3451, 4512, 5123\}$.

The complementary probability mass functions are the eigenvectors $\overline{\mathbf{p}}_{ijkl}$ of the unit eigenvalues of T_{ijkl}, from which the steady-state probability mass functions are

$$\overline{\mathbf{p}}_i = T_{i|jklm} \overline{\mathbf{p}}_{jklm} \tag{122}$$

with

$$T_{i|jklm} = \begin{bmatrix} p_{i|jklm}(y_{i1} | y_{j1}, y_{k1}, y_{l1}, y_{m1}) \cdots p_{i|jklm}(y_{i1} | y_{j2}, y_{k2}, y_{l2}, y_{m2}) \\ p_{i|jklm}(y_{i2} | y_{j1}, y_{k1}, y_{l1}, y_{m1}) \cdots p_{i|jklm}(y_{i2} | y_{j2}, y_{k2}, y_{l2}, y_{m2}) \end{bmatrix} \tag{123}$$

and

$$\overline{\mathbf{p}}_{jklm} = \begin{bmatrix} \overline{p}_{jklm}(y_{j1}, y_{j1}, y_{k1}, y_{m1}) \\ \vdots \\ \overline{p}_{jklm}(y_{j2}, y_{j2}, y_{k2}, y_{m2}) \end{bmatrix}, \tag{124}$$

with entries descending in lexicographical order for $i|jklm \in \{1|2345,\ 2|3451,\ 3|4512,\ 4|5123,\ 5|1234\}$.

5. Summary and Conclusions. We establish a methodology for determining the steady state of a stationary multivariate Markov chain with joint-conditioning transition probabilities $p_{i|-i}\colon \mathcal{A}_i | \mathcal{A}_{-i} \to [0,1]$, rather than marginal-conditioning transition transitions $p_{i|j}\colon \mathcal{A}_i | \mathcal{A}_j \to [0,1]$ which in general, do not account for structural interrelationships that may exist between members of Y_{-i}. The key result of this approach is to establish the existence of a Markov equivalent graphical representation of the original representation of a non-self-influence network that admits the direct application of the Markov chain convergence theorem, which is then extended to general multivariate Markov chains that admit self-influence (and,by direct implication, reciprocal influence).

The distinguishing property of the complementary network graphical representation is that, rather than the vertices comprising the individual members of the network with edges comprising joint-conditioning/marginal-conditioned transition probability mass functions, the vertices of the complementary representation comprise the complementary subsets and the edges comprise joint-conditioning/joint conditioned transition mass functions that are defined to be Markov equivalent to the original conditional mass functions.

This result has numerous potential applications, including to economic, sociological, political, genetic, ecological, robotics, and image modeling, machine learning, and neural network engineering. For such applications it is most useful to model the probabilistic behavior of each influence source in terms of the joint states of all sources, rather than to try to isolate the dependency of each source separately. Indeed, it is not generally possible to deduce equivalent individual conditional dependency relationships from joint conditional dependency relationships.

A significant operational principle that emerges from this analysis is that the statistical intra-relationships of individual influence generators within a group are determined by the statistical interrelationships between their complementary subgroups.This principle may inform the analysis and design of multiagent systems such as robots and of games involving multiple subjects in an economics laboratory or in behavioral ecology experiments. Such projects are facilitated by the design of experiments that generate joint-conditioning transition probabilities rather than focusing on marginal-conditioning transition probabilities (for examples see [26, 27]

It is important to identify limitations of the approach. It cannot be applied in cases where analysis depends on the strict directedness of graphs, for example in solving for extensive-form equilibria of games by backward induction, even when such games involve learning characterized by Markov processes [11, chapters 5 and 6]. However, even in such cases, our approach can narrow the set of available equilibria by being used to model processes by which game players arrive at shared priors for Bayesian updating during subsequent play.

The key practical achievement of the analysis is to extend the reach of Markov-chain analysis to a wide potential range of contexts in which it has hitherto been inapplicable, by relaxing the requirement of strict directedness of associated graphs.

REFERENCES

[1] ANDERSSON, S. A., MADIGAN, D. and PERLMAN, M. D. (1997). A characterization of Markov equivalence classes for acyclic digraphs. *The Annals of Statistics* **25** 505 – 541. https://doi.org/10.1214/aos/1031833662

[2] ANTHONY, M. and BARTLETT, P. (1999). *Neural Network Learning: Theoretical Foundations*. Cambridge University Press.

[3] CASTELLETTI, F., CONSONNI, G., VEDOVA, M. L. D. and PELUSO, S. (2018). Learning Markov Equivalence Classes of Directed Acyclic Graphs: An Objective Bayes Approach. *Bayesian Analysis* **13** 1235–1260.

[4] CECHIN, R. B. and CORSO, L. L. (2019). Higher-Order Multivariate Markov Chain Applied in Dow Jones and Ibovespa Indexes. *Pesquisa Operacional* **39** 205-223. https://doi.org/10.1590/0101-7438.2019.039.01.0205

[5] CHING, W., FUNG, E. and NG, M. (2002). A multivariate Markov chain model for categorical data Sequences and its applications in demand predictions. *IMA Journal of Management Mathematics* **13** 180-199.

[6] CHING, W., FUNG, E. and NG, M. (2008). Higher-order multivariate Markov chains and their applications. *Lin. Alg. Appl* **3** 492-507.

[7] COWELL, R. G., DAWID, A. P., LAURITZEN, S. L. and SPIEGELHALTER, D. J. (1999). *Probabilistic Networks and Expert Systems*. Springer Verlag, New York, NY.

[8] DIAZ-GARCIA, G., BULLO, F. and MARDEN, J. R. (2023). Distributed Markov Chain-Based Strategies for Multi-Agent Robotics Surveillance. *IEEE Control Systems Letters* **7** 2527-2532.

[9] DOOB, J. L. (1953). *Stochastic Processes*. John Wiley & Sons, New York, NY.

[10] EASLEY, D. and KLEINBERG, J. (2010). *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, Cambridge.

[11] FUDENBERG, D. and LEVINE, D. K. (1998). *The Theory of Learning in Games*. MIT Press, Cambridge, MA.

[12] GÜL, M. and ÖZ, E. (2018). Multivariate Markov Chain Model: An Application to S&P500 and FTSE-100 Stock Exchanges. *Marmara Journal of Economics* **2** 75-88.

[13] HE, Y., JIA, J. and YU, B. (2015). Counting and exploring sizes of Markov equivalence classes of directed acyclic graphs. *The Journal of Machine Learning Research* **16** 2589–2609.

[14] JACKSON, M. O. (2008). *Social and Economic Networks*. Princeton University Press, Princeton, NJ.

[15] JENSEN, F. V. (2001). *Bayesian Networks and Decision Graphs*. Springer Verlag, New York, NY.

[16] JORDAN, M. I., ed. (2001). *Learning in Graphical Models*. MIT Press.

[17] KHARE, K. and MUKHERJEE, N. (2013). Convergence Analysis of Some Multivariate Markov Chains Using Stochastic Monotonicity. *The Annals of Applied Probability* **23** 811-833.

[18] LUENBERGER, D. G. (1979). *Introduction to Dynamic Systems*. John Wiley, New York.

[19] MESBAHI, M. and EGERSTEDT, M. (2021). *Graphical Theoretic Methods in Multiagent Networks*. Princeton University Press.

[20] MEYN, S. (2008). *Control Techniques for Complex Networks*. Cambridge University Press, Cambridge, UK.

[21] NORRIS, J. R. (1997). *Markov Chains*. Cambridge University Press.

[22] PEARL, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, CA.

[23] PEARL, J. (2009). *Causality: Models, Reasoning, and Inference*, second ed. Cambridge University Press, Cambridge, UK.

[24] RAFTERY, A. E. (1985). A Model for High-order Markov Chains. *Journal Royal Statistical Society* **47** 528-539.

[25] ROSS, D. and STIRLING, W. C. (2021). Economics, Social Neuroscience, and Mindshaping. In *The Brain and the Social–Methods and Philosophy of Integrating Neuroscience and Economics* (J. Harbecke and C. Hermann-Pillath, eds.) 10, 179-201. Routledge.

[26] ROSS, D., STIRLING, W. C. and TUMMOLINI, L. (2023). Strategic Theory of Norms for Empirical Applications in Political Science and Political Economy. In *Oxford Handbook of Philosophy of Political Science* (H. Kincaid and J. V. Bouwel, eds.) 5, 86 - 121. Oxford University Press.

[27] ROSS, D., STIRLING, W. C. and TUMMOLINI, L. (2024). Modeling norm-governed communities with conditional games: Sociological game-determination and economic equilibria. *Oeconomia* **forthcoming**.

[28] SAMMUT, G. and BAUER, M. (2021). *The Psychology of Social Influence*. Cambridge University Press.

[29] SCHOTTER, A. (2023). *Advice, Social Learning, and the Evolution of Conventions*. Cambridge University Press.

[30] SHOHAM, Y. and LEYTON-BROWN, K. (2009). *Multiagent Systems*. Cambridge University Press, Cambridge, UK.

[31] SIU, T., CHING, W., NG, M. and FUNG, E. (2005). On a multivariate Markov chain model for credit risk measurement. *Quant. Fin.* **7** 543-556.

[32] STIRLING, W. C. (2019). Conditional Coordination Games on Cyclic Social Influence Networks. *IEEE Transactions on Computational Social Systems* **6** 250-267.

[33] WANG, C., HUANG, T.-Z. and CHING, W.-K. (2014). A New Multivariate Markov Chain Model for Adding a New Categorical Data Sequence. *Mathematical Problems in Engineering* **2014**. https://doi.org/10.1155/2014/502808

[34] ZAWIDZKI, T. W. (2013). *Mindshaping: A New Framework for Understanding Human Social Cognition*. MIT Press.

[35] ZHANG, S. Q., CHING, W. K., JIAO, Y., WU, L. Y. and CHAN, R. H. (2008). *A Simplified Multivariate Markov Chain Model for the Construction and Control of Genetic Regulatory Networks*. 2008 2nd International Conference on Bioinformatics and Biomedical Engineering, Shanghai, China. https://doi.org/10.1109/ICBBE.2008.138

[36] ZHU, D.-M. and CHING, W.-K. (2011). A note on the stationary property of high-dimensional Markov chain models. *International Journal of Pure and Applied Mathematics* **66** 321-330.