Proceedings of the 2011 Winter Simulation Conference
S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu, eds.

A FRAMEWORK FOR MODELING STOCHASTIC FLOW AND SYNCHRONIZATION NETWORKS

Mengran Xue

Washington State University
School of EECS, PO Box 642752
Pullman, WA 99164, USA

Sandip Roy

Washington State University
School of EECS, PO Box 642752
Pullman, WA 99164, USA

ABSTRACT

Motivated mainly by infrastructure-network management problems, our group has been pursuing analysis and design of various models for network dynamics, which vary in their specifics but broadly can be viewed as either stochastic flow or synchronization processes defined on a graph. So as to obtain a common framework for these models, here we introduce broad and complementary models for linear stochastic flow and synchronization dynamics in networks, that are structured only in that the network's state evolution is Markov and conditionally linear. We first provide mathematical and graphical formulations for each model, and then verify that the models are broad enough to capture several common synchronization/flow networks. As a first analysis, graph-theoretic characterizations of these models' asymptotics are given; these results generalize and enhance known graphical characterizations of existing synchronization/flow models. A comparison of the stochasticity of different flow network models within the framework is also included.

1 INTRODUCTION

During the last 25 years, scientists in several disciplines have advanced the perspective that many complex phenomena occurring in networks can be abstractly but fruitfully represented using simple dynamical models defined on graphs. A number of simple graphical dynamical models have been introduced within the scope of this *science of networks* (e.g., Strogatz (2001), Li and Chen (2003), Wan, Roy, and Saberi (2008), Krajci and Mrafko (1984), Roy, Sridhar, and Verghese (2003), Xue et al. (2010)). While these network models vary in their specifics, two common themes thread together many of the models:
1) **Synchronization** phenomena, wherein states or opinions of networked autonomous agents come to a common value through interaction and communication, are widely captured.
2) Alternately, many models capture **network flow dynamics**, that is, the movement of material or items among an infrastructural network's components.

The many network models for synchronization and flow have been motivated by applications from different research communities, and have significant differences in their dynamics: for instance, they may have continuous-valued or discrete-valued (or hybrid) states, may have stochastic or deterministic updates, and may exhibit complex state dependencies in their evolution (e.g., Roy, Sridhar, and Verghese (2003), Ribeiro et al. (2007), Xue et al. (2010), Xue et al. (2010), Xue et al. (2011)). Yet, the models have in common that the graph topology plays a critical role in the model's dynamics. For some of these models, connections between the graph topology and both asymptotic and transient characteristics of the dynamics have been obtained, and in turn classes of graphs that yield desirable or undesirable dynamical properties have been found. Very recently, for a very few network models, tools for estimation and design of dynamics that exploit the graph topology also have been developed (e.g., Wan, Roy, and Saberi (2008), Xue et al. (2011), Xue and Roy (2011)). However, in many of these domains, graph-based analysis and estimation/design of network dynamics remain areas of active research, and graph-theoretic characterizations remain limited and incomplete. The need for further analysis and design tools for such common network models is a key motivation for development of a common modeling framework for flow and synchronization problems, and consequent development of an integrated toolbox for analysis and design (and comparison among models). The research presented in this article is a first step in this direction.

In our ongoing research, which is largely focused on automated decision-management in infrastructure networks, we are using several different synchronization and flow network models to represent both physical network dynamics and algorithmic processes. These include both discrete-valued and continuous-valued synchronization models for consensus dynamics and clustering processes, queueing-network and other stochastic-flow models for traffic, and Markov-chain models, among others (e.g. Xue et al. (2011)). In studying these various network models, we are encountering common problems in network parameterization, simulation/analysis of dynamics, and state estimation for each model, that seem to admit similar but not identical solutions. For some of the models, parts of the desired analysis at least are well-established and studied, but these results often do not directly transfer to the other network models, while other analyses perhaps have not been studied thoroughly for any of the models. A more general network modeling framework will permit us to better understand essential properties of flow and synchronization processes, and hence allow us to extend graph-based analysis and estimation/design tools to a much broader family of networks. At the same time, a useful modeling framework must be sufficiently structured to permit efficient simulation and statistical characterization of the dynamics, so as to overcome the "curse of dimensionality" that is ever-present in complex stochastic network models.

The purpose of this article is to introduce dual stochastic modeling frameworks for flow and synchronization networks that can capture numerous common models in the literature, and yet permit extensive graph-theoretic characterization of simulated dynamics. To this end, we begin by defining the broad stochastic modeling frameworks for flow and synchronization networks (Sections 2 and 3), and show that numerous commonly-studied models fit within these classes. We then pursue an introductory graph-theoretic analysis of the two models (Section 4), focusing here on full characterization of the networks' asymptotics and briefly on model comparison (Sections 4.2 and 4.3). In introducing the stochastic-modeling frameworks for flow and synchronization, we pursue two key outcomes:

1) We illustrate a broad family of networked flow- and synchronization- processes, including stochastic and deterministic models, discrete- and continuous-valued dynamics, and even certain apparently-nonlinear processes, admit a common representation and common graphical analyses. This common representation, which only enforces Markovianity and a certain conditional linearity in the state, also permits us to compare the dynamics of different flow- and synchronization- models.

2) We aim to generalize and enhance the graph-theoretic analyses of particular flow- and synchronization-processes in the literature. Our methodology demonstrates that graphical analysis methods that have been developed for a particular model can provide new insights into general flow and synchronization networks. In turn, new graph-theoretic insights can be obtained for other examples in the literature, and for new models. For instance, our work illustrates that the graphical asymptotic analysis of Markov chains translates to the broad flow and synchronization networks introduced here, and gives interesting new insights into e.g. distributed averaging (synchronization) algorithms.

A Notation for Graphs: At several points, our model formulation requires definition of graphs from state matrices describing linear dynamics on networks. The following notation will be used for such graph definitions: for an $n \times n$ state matrix A, the notation $\Gamma(A)$ will be used to describe a weighted and directed graph with n vertices, labeled $1, \ldots, n$. An edge will be drawn from vertex i to vertex j if and only if $a_{ji} > 0$, with the weight of the edge equal to a_{ji}.

2 THE LINEAR STOCHASTIC FLOW NETWORK MODEL

Flow network models are concerned with tracking the movement of items or material among network components or nodes. These models have found wide application in fields ranging from traffic engineering to cell biology (e.g., Roy, Sridhar, and Verghese (2003), Ribeiro et al. (2007)). In this section, let us introduce a tractable framework for modeling flow-network dynamics defined on a graph (Section 2.1), and confirm that it encompasses several widely-used stochastic and deterministic models for flows (Section 2.2).

2.1 Mathematical Formulation

A network with n **components/nodes**, labeled $1,\ldots,n$, is considered. We are concerned with tracking in discrete time a scalar state associated with each network component, that represents an amount of material or number of items (of a single type) at the component. Specifically, we use the notation $r_i[k]$ for the scalar state of component i ($i=1,\ldots,n$) at discrete time k ($k=0,1,2,\ldots$), and refer to this state variable as the **local quantity** of component i at time k.

The local quantities of the network components are modeled as evolving in discrete time, due to conservative flows between the components. We first posit a general stochastic model for the flows, and discuss a linear-algebraic- and graph-theoretic- representations of the model dynamics. Subsequently, we impose a weak conditional linearity condition on the model dynamics, that facilitates analysis.

First, let us describe the general (possibly non-linear) **stochastic flow network** model. In this model, we view the local quantities of the components as being updated by a two-stage process at each time step, namely a **flow-determination** stage followed by a **flow-combination** stage. First, in the flow-determination stage, each local quantity $r_i[k]$ is represented as forming **flows** $f_{ij}[k]$, $j=1,\ldots,n$, to the components in the network (including the component i). Each flow is assumed non-negative ($f_{ij}[k]\geq 0$), and the total of the flows equals the local quantity ($\sum_{j=1}^{n} f_{ij}[k]=r_i[k]$). Flows determination from local quantities is assumed general: it may be either deterministic or stochastic, and the flows $f_{ij}[k]$ may depend on $r_i[k]$ in an arbitrary way. We permit correlation between stochastic flow determinations originating from different components, but do assume that flow-determinations at time k are independent of the system's past history given the time-k local quantities. Second, in the flow-combination stage, the incoming flows to each component are summed to determine the local quantity at the component at the next time-step. That is, we compute $r_i[k+1]$ as follows: $r_i[k+1]=\sum_{j=1}^{n} f_{ji}[k]$. We have thus specified the stochastic flow network update.

We develop a matrix-theoretic formulation of the stochastic flow network, to facilitate graphical analysis of the model. To do so, let us define a **quantity vector** as $\mathbf{r}[k]=\begin{bmatrix} r_1[k] & \cdots & r_n[k]\end{bmatrix}^T$, and a **flow vector** for each component i as $\mathbf{f}_i[k]=\begin{bmatrix} f_{i1}[k] & \cdots & f_{in}[k]\end{bmatrix}^T$. To continue, we note that the flow-determination stage of the flow-network's update enforces that a fraction of each local quantity is directed to each component as a flow. Thus, it is automatic that the flow vector $\mathbf{f}_i[k]$ can be written as $\mathbf{f}_i[k]=\mathbf{p}_i[k]r_i[k]$, where the $n\times 1$ **flow fraction vector** $\mathbf{p}_i[k]$ has entries that are non-negative and sum to 1 (formally, $\mathbf{p}_i[k]\geq 0$ and $\mathbf{1}^T\mathbf{p}_i[k]=1$). Here, the vectors $\mathbf{p}_i[k]$ may be stochastically determined, and further the $\mathbf{p}_i[k]$ may depend on $r_i[k]$. Next, using this expression for flows together with the flow-combination update, the quantity vector at time $k+1$ can be expressed in terms of the vector at time k as $\mathbf{r}[k+1]=P[k]\mathbf{r}[k]$, where the **flow state matrix** $P[k]\overset{\triangle}{=}\begin{bmatrix}\mathbf{p}_1[k] & \ldots & \mathbf{p}_n[k]\end{bmatrix}$ is a column-stochastic matrix that (in general) is stochastically-determined and dependent on $\mathbf{r}[k]$. Thus, the matrix representation has been achieved.

The above matrix notation suggests one graphical representation of the flow-network dynamics. The (possibly randomly-generated) matrix $P[k]$ indicates flows of material/items at time k, and hence naturally admits a graphical interpretation. Thus, viewing $P[k]$ as an instantiation of the flow-network dynamics, we draw a corresponding **flow instantiation graph** $\Phi[k]$ as $\Phi[k]=\Gamma(P[k])$ for time k. The instantiation graph captures the particular splitting of local quantities that occur to form flows at time k, i.e. an edge is drawn from vertex i to j if material/items flow from component i to j at that time, and the weight captures the fraction of the quantity at component i that flows in this direction. We stress that a flow network may have many possible instantiations, and so instantiation graphs, at each time k.

Next, let us introduce a notion of linearity in the flow-network dynamics, that facilitates many graph-theoretic characterizations yet allows representation of several interesting dynamics. To introduce this notion of linearity, we first note that the flow-network dynamics described above are Markovian, in the sense that

the quantity vector at time $k+1$ can be determined only from the quantity vector at time k, given the whole past history of the network. Based on this Markovian structure, we can specify the model in terms of the conditional distribution for the quantity vector at time $k+1$ given the quantity vector at time k, for each k. In turn, conditional statistics for the time-$(k+1)$ quantity vector given the time-k quantity vector can be envisioned. Here, let us define linearity of the model in terms of the first-moment conditional statistics, i.e. the conditional mean for next quantity vector given the current one, or $E(\mathbf{r}[k+1] \mid \mathbf{r}[k])$. Specifically, we will view the stochastic flow network as **linear**, if $E(\mathbf{r}[k+1] \mid \mathbf{r}[k])$ is a purely linear function of $\mathbf{r}[k]$ for all k, or equivalently if $E(\mathbf{r}[k+1] \mid \mathbf{r}[k]) = Q[k]\mathbf{r}[k]$ for a fixed matrix $Q[k]$ (that does not depend on $\mathbf{r}[k]$) for each k. Some remarks on linear stochastic flow networks are needed:
1) The condition for linearity does not require that the $P[k]$ be independent of $\mathbf{r}[k]$, only that the conditional mean of the next quantity vector is a linear function of the current one. One example in Section 2.2 shows that even some dynamics with such state dependences may be linear.
2) For linear stochastic flow networks, the matrix $Q[k]$ permits us to specify another graphical representation. First, since the matrix $Q[k]$ maps the quantity vector at time k to expected flows and hence the expected quantity vector at time $k+1$, we refer to $Q[k]$ as the **flow expectation matrix**. Second, we define a weighted and directed **flow expectation graph** $\overline{\Phi}[k] = \Gamma(Q[k])$. The flow expectation graph captures whether or not, on average, there is a flow between each pair of vertices at each time.
3) Some of the results that we obtain depend on time-invariance in addition to linearity of the quantity vector's conditional expectation. If $E(\mathbf{r}[k+1] \mid \mathbf{r}[k]) = Q\mathbf{r}[k]$ for some fixed Q for all k, we will refer to the stochastic flow network model as a linear time-invariant or LTI one. A single flow expectation graph $\overline{\Phi} = \Gamma(Q)$ can be defined for an LTI stochastic flow network.

2.2 Examples

To illustrate the scope of our modeling framework, we demonstrate that three common flow network models can be posed as linear stochastic flow networks.

Conservative Linear Fluid Flow Model

Classical fluid flow models track continuous-valued quantities (i.e., amounts of material) at network components. In a conservative linear model, deterministic or stochastic fractions of the quantity at each component are viewed as flowing to multiple components over each time interval (e.g., Berman, Neumann, and Stern (1989)). The update equation of such a fluid flow model can be expressed in the stochastic flow network formalism, as follows: $\mathbf{r}[k+1] = P[k]\mathbf{r}[k]$, where the flow state matrix $P[k]$ may be deterministic or randomly selected from a finite sample space, but has no dependence on $\mathbf{r}[k]$ or any other previous state. We note here that the jth entry in $\mathbf{p}_i[k]$ is the exact fraction of $r_i[k]$ that flows to component j between times k and $k+1$. That is, only the independent selection of $P[k]$ may be stochastic, whereupon exact fractional flows of the quantities is enacted in the network. This simple fluid flow model has been used to represent such diverse phenomena as flow of goods and materials in transportation systems, fluids, and queueing dynamics in a high-traffic limit.

It is straightforward to ascertain that the model is a linear stochastic flow network. To this end, let us define $\bar{P}[k]$ as the expectation of $P[k]$ (or $E(P[k]) = \bar{P}[k]$). In this notation, we have $E(\mathbf{r}[k+1] \mid \mathbf{r}[k]) = E(P[k]\mathbf{r}[k] \mid \mathbf{r}[k]) = E(P[k] \mid \mathbf{r}[k])\mathbf{r}[k] = E(P[k])\mathbf{r}[k] = \bar{P}[k]\mathbf{r}[k]$. Thus, the model is linear, with $Q[k] = \bar{P}[k]$.

Markov Chain

A finite-state Markov chain is a classical stochastic model that captures a single discrete-valued state's Markov evolution, or equivalently the stochastic movement of a single object among a graph's vertices. Markov chains have found wide application in fields ranging from telecommunications to cell biology, and have been extensively analyzed (Kemeny and Snell (1976), Brémaud (1999), Krajci and Mrafko (1984)). Here, let us consider a Markov chain whose state can take on n possible values at each time), say $1, \ldots, n$.

The state is viewed as evolving from each value to one of the others with some probability, as captured in a transition matrix $A[k]$ (where, specifically, the entry $a_{ji}[k]$ is the transition probability from value i to value j at time k). We note that the initial value of the Markov chain's state is specified or described by a probability distribution.

Let us now argue that the Markov chain can be formulated as a linear stochastic flow network with n components. In the flow-network formulation, we will say that the quantity $r_i[k]$ of component i is 1 if the value of the Markov chain is i (i.e., the single object is at location i), and is 0 otherwise. From this definition, we see that the quantity vector $\mathbf{r}[k]$ is the $0-1$ **indicator vector** of i (typically denoted $\mathbf{e}_i$) when the Markov chain is in status i. The evolution of the Markov chain from time k to time $k+1$ can be captured as follows using the flow-network formalism: each column i of the matrix $P[k]$ (equivalently, the flow vector $\mathbf{p}_i[k]$) is chosen independently to be an indicator vector, specifically to equal the vector $\mathbf{e}_j$ with probability $a_{ji}[k]$. Then we notice that $\mathbf{r}[k+1] = P[k]\mathbf{r}[k]$ equals $\mathbf{e}_j$ with probability $a_{ji}[k]$ if $\mathbf{r}[k] = \mathbf{e}_i$ for all i and j, which matches exactly the Markov chain's update. In other words, this update captures that a unit quantity at component i will flow as a single unit to a component j, with appropriate probability. Thus, we have phrased the state update as as a stochastic flow network. To check linearity, we first note that $E(\mathbf{p}_i[k]) = \begin{bmatrix} a_{1i}[k] & \cdots & a_{ni}[k] \end{bmatrix}^T$, and hence that $E(P[k]) = A[k]$. Further noting the independence of $P[k]$ from the past quantity vector $\mathbf{r}[k]$, we have $E(\mathbf{r}[k+1] \,|\, \mathbf{r}[k]) = E(P[k]\mathbf{r}[k] \,|\, \mathbf{r}[k]) = E(P[k]|\mathbf{r}[k])\mathbf{r}[k] = E(P[k])\mathbf{r}[k] = A[k]\mathbf{r}[k]$. Thus, we have verified that this stochastic flow network is linear, with expectation flow matrix $Q[k]$ equal to $A[k]$.

A Probabilistic Routing Model

Let us describe a third common model, which we call a probabilistic routing model, that can be viewed as a stochastic flow network. In this model, component quantities comprise integral numbers of discrete units, each of which independently flow through the network in a Markovian fashion. Dynamics of this form are observed in infinite-server queueing network representations, and have found application in modeling data-packet transmission and air transportation networks among many other domains (e.g., Roy, Sridhar, and Verghese (2003), Ribeiro et al. (2007)).

Let us describe the dynamics of such a probabilistic routing model directly in the stochastic flow-network modeling framework. The quantity variables $r_i[k]$ that we track represent the number of discrete units at each component i at time k, and are constrained be integral. Let us begin by describing the flow-determination stage at each component i. At time k, each of the $r_i[k]$ units at component i are independently routed to other components (including itself) with certain probabilities, which are specified in a probability vector $\mathbf{d}_i[k]$ (i.e., the jth entry in $\mathbf{d}_i[k]$ is the probability with which each unit at component i will flow to component j). Therefore, the flow vector $\mathbf{f}_i[k]$ follows the **multinomial distribution** with parameters $r_i[k]$ and $\mathbf{d}_i[k]$ (see Mosimann (1962)), and has $\begin{pmatrix} n+r_i[k]-1 \\ n-1 \end{pmatrix}$ possible values (with probabilities specified by the multinomial distribution). The flow fraction vector $\mathbf{p}_i[k]$ is given by $\mathbf{p}_i[k] = \mathbf{f}_i[k]/r_i[k]$, and so is simply a scaled version of the multinomially-distributed vector $\mathbf{p}_i[k]$. For example, let us assume that $n=2$, $r_1[k]=3$, and $\mathbf{d}_1[k] = \begin{bmatrix} 0.2 & 0.8 \end{bmatrix}^T$. Then, $\mathbf{p}_1[k]$ has 4 possible values, which are $\begin{bmatrix} 0 & 1 \end{bmatrix}^T$, $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix}^T$, $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix}^T$, $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$, with probabilities $(0.8)^3$, $(0.2)(0.8)^2$, $(0.2)^2(0.8)$, $(0.2)^3$, respectively. Since we have described the generation rule for the flow fraction vector $\mathbf{p}_i[k]$, we have thus specified the instantiation matrix $P[k]$ of the probabilistic routing model.

Now, let us discuss the linearity of the model. We first note that the instantiation matrix $P[k]$ (which is $\begin{bmatrix} \mathbf{p}_1[k] & \ldots & \mathbf{p}_n[k] \end{bmatrix}$) actually depends on $\mathbf{r}[k]$, since the distribution of $\mathbf{p}_i[k]$ depends on $r_i[k]$. However, the model itself is still linear even with such a state dependence. Specifically, since the conditional first moment of the scaled multinomially-distributed vector $\mathbf{p}_i[k]$ is $E(\mathbf{p}_i[k] \,|\, \mathbf{r}[k]) = \mathbf{d}_i[k]$ (which is in fact independent on $\mathbf{r}[k]$), we have $E(\mathbf{r}[k+1] \,|\, \mathbf{r}[k]) = E(P[k]\mathbf{r}[k] \,|\, \mathbf{r}[k]) = E(P[k] \,|\, \mathbf{r}[k])\mathbf{r}[k] = \begin{bmatrix} \mathbf{d}_1[k] & \ldots & \mathbf{d}_n[k] \end{bmatrix} \mathbf{r}[k]$. Thus, the stochastic flow network is linear with $Q[k] = \begin{bmatrix} \mathbf{d}_1[k] & \ldots & \mathbf{d}_n[k] \end{bmatrix}$.

3 THE STOCHASTIC SYNCHRONIZATION NETWORK

Models for *synchronization* (or *consensus*) among networked autonomous agents have also found application in diverse fields, including distributed computing, analog-circuit design, cell-biology, and particle physics, among many others (e.g., Tsitsiklis (1984), Liggett (1999), Xue et al. (2010), Xue et al. (2010)). Fundamentally, synchronization models are often concerned with tracking either physical- or information-states of essentially autonomous agents, which through decentralized interaction or communication achieve identical state values. Although synchronization and flow dynamics are typically present in very different applications, we will see that the topological structures and analysis/design of flow and synchronization networks are strongly interconnected, and a common theory can be developed for both classes of models. Our purpose here is to introduce a general discrete-time stochastic linear modeling framework for synchronization (Section 3.1), and to show that the model captures many synchronization models available in the literature (Section 3.2).

3.1 Mathematical Formulation

In analogy with the flow networks, Let us first define stochastic synchronization models in some generality, and then focus on the the linear case. Formally, we consider a network of n agents, labelled $1,\dots,n$, and associate with each agent i a scalar variable $s_i[k]$ which we call the **opinion** of agent i. In contrast with the flow model, state variables may be positive or negative for the stochastic synchronization network. We find it convenient to assemble the opinions into a single **opinion vector** $\mathbf{s}[k] = \begin{bmatrix} s_1[k] & \cdots & s_n[k] \end{bmatrix}^T$.

Each agent is viewed as updating its opinion in discrete time, in such a way that its neighbors' current opinions are incorporated. Specifically, we model each agent's next opinion as a weighted non-negative *unitary* linear combination of multiple agents' current opinions. That is, agent i's opinion is updated as $s_i[k+1] = \mathbf{g}_i^T[k]\mathbf{s}[k]$, where the **influence vector** $\mathbf{g}_i[k]$ is entrywise nonnegative and sums to 1 ($\mathbf{g}_i^T\mathbf{1} = 1$). Let us stress here that each influence vector $\mathbf{g}_i[k]$ may be a deterministic or stochastic quantity, which may be correlated with $\mathbf{g}_j[k]$ ($j \neq i$) and may even depend on $\mathbf{s}[k]$ (but otherwise must be independent of the past history of the network). Assembling the opinion-update equations of the n agents as a single vector equation, we obtain $\mathbf{s}[k+1] = G^T[k]\mathbf{s}[k]$, where $G[k] = \begin{bmatrix} \mathbf{g}_1[k] & \dots & \mathbf{g}_n[k] \end{bmatrix}$ is a column-stochastic **influence matrix** that in general may be stochastic and concurrent-state dependent.

Let us define graphs that describe the interactions among agents in the stochastic synchronization model's opinion update. In particular, we define an n-vertex **influence instantiation graph** $\Lambda[k]$ at time k as $\Gamma(G^T[k])$, i.e. as the weighted and directed graph associated with the state matrix $G^T[k]$ at time k. The graph can be interpreted as follows: an arrow is drawn from vertex i to vertex j if component i's current opinion influences the next opinion of component j, with the weight equal to the strength of the influence. We notice that an influence instantiation graph may be stochastic, time-varying, and state-dependent, but always has the property that the sum of the weights on edges entering each vertex is 1.

Finally, let us identify a subset of the stochastic synchronization network models defined above, that are specially structured in that their expected dynamics are linear. Specifically, we call the synchronization model above a **linear stochastic synchronization network**, if the state dynamics satisfy $E(\mathbf{s}[k+1] \,|\, \mathbf{s}[k]) = H^T[k]\mathbf{s}[k]$, where the **influence expectation matrix** $H^T[k]$ is a state-independent matrix, for each k. In other words, we say that the synchronization network is linear, if the expectation of the next opinion vector given the current one is a linear function of the current opinion vector. If, further, the expected influence matrix $H^T[k]$ is time-independent (say equal to H^T), we shall call the model a **linear time invariant (LTI) stochastic synchronization network**. For linear and LTI stochastic synchronization networks, we find it convenient also to define graphs based on the influence expectation matrix. In particular, we define the **influence expectation graph** at time k for the linear model (or simply the influence expectation graph for the LTI model) as $\overline{\Lambda}[k] = \Gamma(H^T[k])$ (respectively, $\overline{\Lambda} = \Gamma(H^T)$).

Remark: Trivially, the transpose of an influence matrix is a flow matrix, and vice versa. Deeper consideration exposes a tighter duality, namely that a certain time-reversal of a synchronization process is a flow-network dynamics, and vice versa. Details are omitted in the interest of space

3.2 Examples

To highlight the score of the linear stochastic synchronization network, we introduce three widely-studied examples of the model.

Distributed Averaging Algorithm

Distributed averaging algorithms, in which network agents have continuous-valued opinions that evolve according to purely linear updates, have been extensively studied. Such models have been widely used to represent distributed decision-making processes in the computing sciences, and are classically the models considered in synchronization processes (Tsitsiklis (1984), Blondel et al. (2005), Roy, Saberi, and Herlugson (2007)). Let us give a precise formulation of such a model as a stochastic synchronization network. To this end, let us consider a network model with n agents, where agent i has associated with a scalar opinion $s_i[k]$ at time k. Then, agent i's next opinion, $s_i[k+1]$, is generated as a linear combination of all agents' current opinions, i.e. $s_i[k+1] = \mathbf{g}_i^T[k]\mathbf{s}[k]$, where the jth entry of $\mathbf{g}_i[k]$ is the weight of influence from agent j at time k. We note here that the influence vector $\mathbf{g}_i[k]$ can be either deterministic or independently selected from a sample space at each time.

Using a similar approach to the linearity analysis of the fluid flow model, we can show that the distributed averaging algorithm is linear, with $H^T[k] = E(G^T[k])$.

Voter Model

A voter model is a network model in which each agent independently stochastically chooses one other agent (maybe itself) according to a probability vector, and then copies the chosen agent's current opinion as its next opinion (Liggett (1999), Asavathiratham (2000), Xue et al. (2010), Xue et al. (2010)). Let us formulate the voter model as a stochastic synchronization network. We assume that each agent's opinion is initially an arbitrary real scalar. (Often, binary voter models, in which the opinions are constrained to be 0 or 1, are studied; however, our formulation permits propagation of arbitrary opinions.) At each time step, agent i's next opinion is determined as follows: the agent i selects a neighbor j with probability $c_{ji}[k]$ (where $c_{ji}[k] > 0$ and $\sum_j c_{ji}[k] = 1$), whereupon it copies the current status of the neighbor (i.e., $s_i[k+1] = s_j[k]$). Equivalently, we can write the updating dynamics as $s_i[k+1] = \mathbf{g}_i^T[k]\mathbf{s}[k]$, where $\mathbf{g}_i[k] = \mathbf{e}_j$ with probability $c_{ji}[k]$. Thus, we have posed the dynamics as that of a stochastic synchronization network.

Next, let us discuss the linearity of the voter model. According to the above description, we obtain that $E(\mathbf{g}_i[k]) = \mathbf{c}_i[k]$, where $\mathbf{c}_i[k] \stackrel{\triangle}{=} \begin{bmatrix} c_{1i}[k] & \cdots & c_{ni}[k] \end{bmatrix}^T$ (and $\mathbf{c}_i^T[k]\mathbf{1} = 1$).Since $\mathbf{g}_i[k]$ is independent of previous and current opinion, we then have $E(\mathbf{s}[k+1] \,|\, \mathbf{s}[k]) = E(G^T[k]\mathbf{s}[k] \,|\, \mathbf{s}[k]) = E(G^T[k] \,|\, \mathbf{s}[k])\mathbf{s}[k] = E(G^T[k])\mathbf{s}[k] = \begin{bmatrix} \mathbf{c}_1[k] & \cdots & \mathbf{c}_n[k] \end{bmatrix}^T \mathbf{s}[k]$, which indicates that the voter model is a linear synchronization network model, with $H^T[k] = \begin{bmatrix} \mathbf{c}_1[k] & \cdots & \mathbf{c}_n[k] \end{bmatrix}^T$.

One further note about the binary voter model is worthwhile. For this case, we note that the expected opinion vector contains the *probabilities* that each agent has status 1.

A Mixed Model

Finally, let us introduce an apparently-nonlinear stochastic synchronization network model that is a state-dependent mixture of a distributed-averaging algorithm and a voter model. Specifically, each agent's opinion at each time is assumed to be a real-valued scalar. We also associate with each agent i a threshold value, say b_i. At each time k, agent i updates its state as follows: the agent compares its current opinion $s_i[k]$

with its threshold b_i. If $s_i[k] \geq b_i$, agent i randomly picks one other agent j with probability g_{ij}, and then copies agent j's current opinion as its next opinion. For this case, the influence vector is $\mathbf{g}_i[k] = \mathbf{e}_j$ with probability g_{ij}. If $s_i[k] < b_i$, agent i computes a weighted average of the agents' current opinions based on weights $g_{i1}, \cdots, g_{in}$, and chooses this average to be its next opinion (i.e., $s_i[k+1] = \sum_{j=1}^{n} g_{ij} s_j[k]$). For this case, the influence vector $\mathbf{g}_i[k]$ is simply $\begin{bmatrix} g_{i1} & \cdots & g_{in} \end{bmatrix}^T$. We note that this stochastic synchronization network model is actually a mixture of a voting model and a distributed averaging model. We note that the instantiation matrix $G[k]$ (i.e., $G[k] = \begin{bmatrix} \mathbf{g}_1[k] & \cdots & \mathbf{g}_n[k] \end{bmatrix}$) depends on the current opinions of the agents.

Although the two synchronization mechanisms in the two operating regimes (i.e., above threshold and below threshold) are completely different, the model is linear. Let us use notation $\bar{G}[k]$ where the i, jth entry of $\bar{G}[k]$ has value g_{ji}. For the case that $s_i[k] \geq b_i$ (voting model phase), we can easily obtain that $E(\mathbf{s}[k+1] \,|\, \mathbf{s}[k]) = \bar{G}^T[k]\mathbf{s}[k]$. For the case that $s_i[k] < b_i$ for any i (distributed averaging model phase for agents i), we also obtain that $E(\mathbf{s}[k+1] \,|\, \mathbf{s}[k]) = \bar{G}^T[k]\mathbf{s}[k]$. Therefore, the model dynamics satisfy the linearity condition.

Remark: Let us highlight the strong analogy between the the first two examples of flow models (the fluid-flow and Markov chain models), and the corresponding examples of synchronization models (the distributed averaging and voter models, respectively). We also note the possibility for state-dependent update processes that are nevertheless linear in both settings.

4 GRAPH-THEORETIC ANALYSIS OF THE TWO MODELS

The infrastructure-network management applications that motivate our studies of flow/synchronization models require a comprehensive suite of analysis, parameterization, and estimation tools for network dynamics. In this first work, we provide only a few basic results regarding graph-theoretic analysis of the stochastic linear flow/synchronization models' dynamics. Although these analyses are very basic, they are important as a foundation for more intricate graph-theoretic analysis/estimation of network dynamics, and are important in and of themselves as enhancements of results for particular flow- or synchronization networks, or as tools for comparing networks.

Specifically, we begin with a few preliminaries on invariants of the model (Section 4.1). The main focus of the section is on graphical characterizations for the asymptotics of both the expected dynamics and the stochastic dynamics themselves are given (Section 4.2). Finally, we briefly pursue a comparison among flow models that fall within the broad class defined here (Section 4.3). Note: proofs have been taken out in the interest of space; please see the extended document (**?**) for the proofs.

4.1 Preliminary Observations: Invariants

As a preliminary step, let us formalize complementary invariances in the (general, possibly nonlinear) flow and synchronization models' dynamics. These invariances at their essence are simply formalizations of the principles underlying conservative flow and synchronization, respectively. Here are the results:

1. For the stochastic flow model, the sum of entries in the quantity vector remains unchanged with time. That is, for any trajectory of the stochastic flow model, $\mathbf{1}^T\mathbf{r}[k]$ is identical for $k = 0, 1, 2, \ldots$. For convenience, let us use r_s to represent the total quantities in the flow network ($r_s = \mathbf{1}^T\mathbf{r}[k]$).
2. For the stochastic synchronization network, an opinion vector whose entries are identical is an invariant of the dynamics. That is, if $\mathbf{s}[k_0] = c\mathbf{1}$ for some scalar c and some time-step k_0, then $\mathbf{s}[k] = \mathbf{s}[k_0] = c\mathbf{1}$ for all $k > k_0$.

4.2 Asymptotics

We develop graph-theoretic characterizations of the asymptotics of linear time-invariant stochastic flow- and synchronization- networks. Specifically, we first give graph-theoretic characterizations of the *expected* state dynamics, in terms of the expectation graph. We then build on the characterization of the expected state dynamics, to relate properties of the network model's stochastic dynamics (e.g., ergodicity) to the graph. In the interest of space, only a few preliminary results are given here.

Let us first study the asymptotics of the expected state. A graphical characterization for both models can be achieved, using classifications of the graph topology that are analogous to the ones used in analyzing Markov chains (Kemeny and Snell (1976), Br*é*maud (1999)). In developing this analysis, it is worth recognizing that an LTI flow-network's expectation graph is identical to a Markov chain's transition graph, and in fact the precise classification used for Markov chains will suffice for flow networks. Meanwhile, the state matrix for a synchronization network's expected state dynamics is the transpose of that for a flow network, and so the graph edges are reversed; thus, a slightly different classification is needed.

Let us begin by defining some terminologies for the LTI stochastic flow network's expectation graph: The expectation flow graph is said to have a **path** from vertex i to vertex j if it has a sequence of directed edges from vertex i to vertex j. If there is a path from i to j, and also from j to i, the two vertices i and j are said to **co-transport**. A set of vertices that co-transport with each other, and do not co-transport with any vertex outside the set, is called a **flow class**. A flow class is called **absorbing**, if there is no path from a vertex in the flow class to one outside. In other words, a flow class is absorbing if material/items cannot flow out of the corresponding network components. The vertices within an absorbing flow class are called **absorbing vertices**. A flow class that is not absorbing is called **transient**. The vertices within a transient flow class are called **transient vertices**. For an absorbing vertex i, a notion of periodicity needs to be defined. Specifically, the lowest common denominator among the path lengths from vertex i back to itself is termed the **period** of the vertex; the vertex is called **aperiodic** if the period is 1, and periodic otherwise. The vertices in an absorbing class can be shown to have the same period, so the period measure (and periodicity) can be associated with the whole flow class.

We also define complementary terminologies for the LTI stochastic synchronization network, in terms of its expectation graph. The expectation synchronization graph is said to have an **influence path** from vertex i to vertex j if the graph has a sequence of directed edges from vertex i to vertex j. If there is a path from vertex i to vertex j, and also from vertex j to vertex i, the two vertices i and j are said to **co-influence**. A set of vertices that co-influence with each other, and do not co-influence with any vertex outside the set, is called an **influence class**. An influence class is called **autonomous**, if there is no path from a vertex outside the influence class to one inside. In other words, an influence class is autonomous if the expected opinions of the corresponding network components are not dependent on previous opinions of other network components. The vertices within an autonomous influence class are called **autonomous vertices**. An influence class that is not autonomous is called **dependent**. The vertices within a dependent influence class are called **dependent vertices**. For an autonomous vertex i, a notion of periodicity needs to be defined. Specifically, the lowest common denominator among the path lengths from vertex i back to itself is termed the **period** of the vertex; the vertex is called **aperiodic** if the period is 1, and periodic otherwise. All the vertices in an autonomous influence class have the same period, so the period measure (and the periodicity concept) can be associated with the whole influence class.

We are now ready to present results regarding the asymptotics of the expected state dynamics. For both networks, we will characterize the asymptotics according to a classification of the graph topology, showing that qualitatively different asymptotics result depending on the topology. We begin with the flow-network result, which admits the well-known characterization of Markov chains as a special case.

Theorem 1 Consider an LTI stochastic flow network with expectation flow graph Γ. Then the asymptotics of the expected quantity vector $E(\mathbf{r}[k])$ are as follows:

1) The expected quantities for the network components associated with the transient vertices of Γ asymptotically approach 0, i.e. $\lim_{k\to\infty} E(r_i[k]) = 0$ if i is a transient vertex.

2) Consider vertices in Γ that are in absorbing aperiodic flow classes. The expected quantity for each corresponding network component reaches a limit asymptotically, i.e. $\lim_{k\to\infty} E(r_i[k])$ exists for all such i. If in fact Γ has only a single absorbing class which is aperiodic, then the asymptotic expectations at network components associated with this class are fixed positive fractions of the expected total quantity at the initial

time. That is, for each vertex i in the absorbing aperiodic flow class, we have $\lim_{k\to\infty} E(r_i[k]) = v_i r_s$, where $v_i > 0$ and the sum of v_i over the vertices i in the absorbing aperiodic flow class is 1.
3) In general, the expected quantities at network components corresponding to absorbing periodic vertices are not guaranteed to converge asymptotically. However, these expected quantities sampled at intervals equal to the period are convergent. That is, for an absorbing q-periodic vertex i in Γ, we have $\lim_{k\to\infty} E(r_i[qk+z])$ exists, for $z = 0,1,\ldots,q-1$.

Next, let us present complementary results for the LTI stochastic synchronization model.

Theorem 2 Consider an LTI stochastic synchronization network with expectation synchronization graph Λ. Then the asymptotics of the expected opinion vector $E(\mathbf{s}[k])$ are:
1) The expected opinions of network components associated with an autonomous aperiodic class each converge to a limit asymptotically, and further they synchronize (become equal) asymptotically. That is, for a vertex i in an autonomous class, $\lim_{k\to\infty} E(s_i[k])$ exists; also, for two vertices i and j in an autonomous class, we have $\lim_{k\to\infty} E(s_i[k]) - E(s_j[k]) = 0$. Further, the asymptotic value of the expected opinions is a positive unitary linear combination of the initial expected opinions of the agents. That is, $\lim_{k\to\infty} E(s_i[k]) = \sum_{j\in V_s} w_j E(s_j[0])$, where $w_j > 0$, $\sum_{j\in V_s} w_j = 1$, and V_s contains all the vertices in the autonomous class.
2) If all the autonomous classes in the graph are aperiodic, then the expected opinions of all network components converge to a limit asymptotically (i.e., $\lim_{k\to\infty} E(s_i[k])$ exists for all vertices i). The expected opinion of each network component associated with a dependent vertex converges to a unitary nonnegative linear combination of the limiting expected opinions of the autonomous classes. If in fact the graph has a single autonomous class that is aperiodic, then all network components synchronize asympotically, to the same limiting expected opinion as in the autonomous class.
3) The expected opinions at network components corresponding to autonomous periodic vertices are not guaranteed to converge asymptotically. However, these expected opinions sampled at intervals equal to the period are convergent. That is, for a autonomous q-periodic vertex i in Γ, we have that $\lim_{k\to\infty} E(s_i[qk+z])$ exists, for $z = 0,1,\ldots,q-1$. When one or more autonomous classes are periodic, then the expected opinions at the network components associated with dependent vertices also vary periodically, with period equal to the least common multiple of the periods of some or all of the autonomous classes' periods.

Remark 1: For synchronization networks, it is the autonomous classes' initial states that impact the asymptotic expected dynamics. The number of these autonomous classes and their periodicity structures modulate the asymptotic expected dynamics just as for flow networks, albeit with some subtle differences.

Remark 2: Tsitsiklis (1984) provides a comprehensive study of distributed averaging in the time-varying case; these results can permit generalization of the above results to the time-varying case.

The above theorems give necessary and sufficient graphical characterizations for the asymptotic expected dynamics. Such graphical analysis of Markov chains is already well-known, and a partial characterization of the steady-state dynamics was given for the binary voter model in Asavathiratham (2000). Otherwise, to the best of our knowledge, flow/synchronization dynamics have not been related to graph class structures as we have done here. Thus, the results provide immediate insight into the expected state's asymptotics, for several models including distributed-averaging and probabilistic routing ones.

Finally, let us also characterize asymptotic properties of the networks' stochastic state dynamics, not only the expected dynamics. We again give complementary results for synchronization and flow networks. Here are two results for LTI stochastic flow networks:

Theorem 3 Consider an LTI stochastic flow network. The quantities $r_i[k]$ at network components associated with transient vertices i converge to 0 in a mean square sense.

Theorem 4 Consider an LTI stochastic flow network that has a single absorbing class, which is aperiodic. Then, given the total quantity $\mathbf{1}^T\mathbf{r}[0]$, the evolution of quantity vector $\mathbf{r}[k]$ is mean-square ergodic.

The LTI stochastic synchronization network admits dual characterizations of its asymptotic state dynamics, in terms of the underlying expectation graph structure. Specifically, the asymptotic characterizations of the synchronization network are concerned with the dependence of an opinion at some time on initial or previous opinions of all the agents. In particular, we note that the opinion of each agent i at time k, or $s_i[k]$, can always be written as a unitary linear combination of the opinion vector at any previous time $\bar{k} < k$: $s_i[k] = \mathbf{w}_i^T[k,\bar{k}]\mathbf{s}[\bar{k}]$, where the **opinion-influence vector** for site i $\mathbf{w}_i[k,\bar{k}]$ is nonnegative and satisfies $\mathbf{w}_i^T[k,\bar{k}]\mathbf{1} = 1$. Here, let us present two results on the time-evolution of the opinion-influence vector:

Theorem 5 Consider an LTI stochastic synchronization network. For fixed initial time $\bar{k}$, any element of an opinion-influence vector corresponding to a dependent vertex in the expectation graph converges to 0 with respect to the time-index k, in a mean square sense.

Theorem 6 Consider an LTI stochastic flow network that has a single autonomous, aperiodic class. Then the opinion-influence vectors' dynamics for the network are ergodic, in the following sense: for fixed k_0 and any i, $\frac{1}{k-k_0+1}\sum_{z=k_0}^{k}\mathbf{w}_i[k,z]$ approaches $E(\mathbf{w}_i[k,k_0])$ in a mean-square sense, as k is made large.

We note that the ergodicity results are particularly informative, in that they suggest that certain unknown flow/synchronization processes can be partially characterized from time histories.

4.3 Comparisons among Flow and Synchronization Models

The broad framework introduced here also permits comparison among the dynamics of different flow models and synchronization models. That is, the formulation permits us to compare characteristics of stochastic flow or synchronization models with different update rules, and to develop general bounds for their dynamical properties. Here, let us present a first result of this sort, for linear stochastic flow networks. In particular, we bound the variability of a linear stochastic flow network's state with given expectation matrices $Q[k]$, and argue that a Markov-chain-type model achieves the maximum variability among all flow networks with these expectation matrices. Here is the result:

Theorem 7 Consider a linear stochastic flow network, with expected initial state $E(\mathbf{r}[0])$ and expectation flow matrices $Q[0], Q[1], \ldots$. Consider the variability of the quantity vector $\mathbf{r}[k]$. In particular, consider the measure $L = \sum_{i=1}^{n} var(r_i[k])$. Then $L \leq (\mathbf{1}^T E(\mathbf{r}[0]))^2 - \mathbf{z}^T\mathbf{z}$, where $\mathbf{z} = Q[k]...Q[0]E(\mathbf{r}[0])$. The maximum variability is achieved by a Markov chain-type dynamics. Specifically, a model in which the total quantity is placed at a single component at the initial time (with the component chosen probabilistically so that the expected initial state is as specified), and in which the flow state matrix at each time k is that of a Markov chain with transition matrix $Q[k]$, achieves the bound.

ACKNOWLEDGEMENTS

This project was partially supported by NSF grants ECS-0725589, ECS-0901137, CNS-1035369, and CNS-1058124.

REFERENCES

Asavathiratham, C. 2000. *The Influence Model: A Tractable Representation for the Dynamics of Networked Markov Chains*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Berman, A., M. Neumann, and R. J. Stern. 1989. *Nonnegative matrices in dynamical systems*. New York: Wiley Interscience.

Blondel, V. D., J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis. 2005. "Convergence in Multiagent Coordination, Consensus, and Flocking". In *Proceedings of the Joint 44th IEEE Conference on Decision and Control and European Control Conference*, 2996–3000. Seville, Spain.

Brémaud, P. 1999. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. New York: Springer.

Kemeny, J. G., and J. L. Snell. 1976. *Finite Markov chains*. New York: Springer-Verlag.

Krajci, M., and P. Mrafko. 1984. "Computer simulations of the structure of amorphous systems by the Markov chain method". *Journal of Physics F: Metal Physics* 14 (6): 1325–1332.

Li, X., and G. Chen. 2003. "Synchronization and desynchronization of complex dynamical networks: an engineering viewpoint". *IEEE Transactions on Circuits and Systems* 50:1381–1390.

Liggett, T. M. 1999. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Berlin: Springer.

Mosimann, J. E. 1962. "On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions". *Biometrika* 49:65–82.

Ribeiro, A., Z. Q. Luo, N. D. Sidiropoulos, and G. B. Giannakis. 2007. "Modelling and optimization of stochastic routing for wireless multihop networks". In *Proceedings of the 26th Annual Joint Conference of the IEEE Computer and Communications Societies*, 1748–1756. Anchorage, Alaska.

Roy, S., A. Saberi, and K. Herlugson. 2007. "A control-theoretic perspective on the design of distributed agreement protocols". *International Journal of Robust and Nonlinear Control* 17:1034–1066.

Roy, S., B. Sridhar, and G. C. Verghese. 2003. "An Aggregate Dynamic Stochastic Model for an Air Traffic System". In *Proceedings of the 5th Eurocontrol/Federal Aviation Agency Air Traffic Management Research and Development Seminar*. Budapest, Hungary. Available via http://atmseminar.eurocontrol.fr/past-seminars/5th-seminar-budapest-hungary-june-2003/papers/paper_003.

Strogatz, S. 2001. "Exploring complex networks". *Nature* 410:268–276.

Tsitsiklis, J. N. 1984. *Problems in decentralized decision making and computation*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA. Available via http://web.mit.edu/jnt/www/PhD-84-jnt.pdf [accessed March 26, 2011].

Wan, Y., S. Roy, and A. Saberi. 2008. "A New Focus in the Science of Networks: Towards Methods for Design". *Proceedings of the Royal Society A* 464:513–535.

Xue, M., and S. Roy. 2011. "Spectral and Graph-Theoretic Bounds on Steady-State-Probability Estimation Performance for An Ergodic Markov Chain". In *Proceedings of 2011 American Control Conference*, 2399–2404. San Francisco, CA.

Xue, M., S. Roy, A. Saberi, and B. Lesieutre. 2010. "On Generating Sets of Binary Random Variables with Specified First- and Second- Moments". In *Proceedings of the 2010 American Control Conference*, 1133–1138. Baltimore, MD.

Xue, M., S. Roy, Y. Wan, and A. Saberi. 2010. "Designing Asymptotics and Transients of Linear Stochastic Automaton Networks". In *Proceedings of the AIAA Guidance, Navigation and Control Conference*. Toronto, Canada.

Xue, M., E. Yeung, A. Rai, S. Roy, Y. Wan, and S. Warnick. 2011. "Initial-Condition Estimation in Network Synchronization Processes: Graphical Characterizations of Estimator Structure and Performance". In *Proceedings of the AIAA Guidance, Navigation and Control Conference*. Portland, OR.

Xue, M., S. M. Zobell, S. Roy, C. P. Taylor, Y. Wan, and C. Wanke. 2011. "Using Stochastic, Dynamic Weather-Impact Models in Strategic Traffic Flow Management". In *Proceedings of the 91st American Meteorological Society Annual Meeting*. Seattle, WA.