Stochastic Systems
arXiv: 1106.4288

CONTINUUM LIMITS OF MARKOV CHAINS WITH APPLICATION TO WIRELESS NETWORK MODELING[*,†]

By Yang Zhang[‡], Edwin K. P. Chong[‡], Jan Hannig[§] and Donald Estep[‡]

Colorado State University[‡] *and The University of North Carolina at Chapel Hill*[§]

We investigate the continuum limits of a class of Markov chains. The investigation of such limits is motivated by the desire to model networks with a very large number of nodes. We show that a sequence of such Markov chains indexed by N, the number of components in the system that they model, converges in a certain sense to its continuum limit, which is the solution of a partial differential equation (PDE), as N goes to infinity. We provide sufficient conditions for the convergence and characterize the rate of convergence. As an application we approximate Markov chains modeling large wireless networks by PDEs. While traditional Monte Carlo simulation for very large networks is practically infeasible, PDEs can be solved with reasonable computation overhead using well-established mathematical tools.

1. Introduction.

In this paper we analyze the convergence of a class of Markov chains to their continuum limits, which are the solutions of certain partial differential equations (PDEs). As an application of the results of such analysis to network modeling, we use PDEs to approximate large wireless networks modeled by such Markov chains.

[*]Yang Zhang and Edwin K. P. Chong's research was supported in part by NSF under ECCS-0700559 and ONR under grant N00014-08-1-110. Jan Hannig's research was supported in part by NSF under Grant No. 1007543 and 1016441. Donald Estep's work is supported in part by the Defense Threat Reduction Agency (HDTRA1-09-1-0036), Department of Energy (DE-FG02-04ER25620, DE-FG02-05ER25699, DE-FC02-07ER54909, DE-SC0001724, DE-SC0005304, INL00120133), Lawrence Livermore National Laboratory (B573139, B584647, B590495), the National Aeronautics and Space Administration (NNG04GH63G), the National Institutes of Health (5R01GM096192-02), the National Science Foundation (DMS-0107832, DMS-0715135, DGE-0221595003, MSPA-CSE-0434354, ECCS-0700559, DMS-1016268, DMS-FRG-1065046), and Idaho National Laboratory (00069249, 00115474).

[†]A preliminary version of parts of the work of this paper was presented at the 49th IEEE Conference on Decision and Control (CDC 2010).

AMS 2000 subject classifications: Primary 60J10, 60J20, 90B15, 90B10, 35Q80; secondary 40A05, 90B18

Keywords and phrases: Continuum limit, Markov chain, partial differential equation, large network modeling, wireless sensor network

Modeling is an important tool in the analysis and design of communication networks, which tend to present very complex behavior that is difficult to describe and comprehend, and experimentally expensive to determine. For example, in many cases, networks have an inherently stochastic nature, e.g., arising from data traffic and communication protocols themselves. In these cases, analysis tends to focus on computing statistical information about the network behavior, e.g., expected throughput of a particular design, and on comparing network designs based on these statistical properties.

One traditional approach to computing network statistics is direct Monte Carlo computer simulation [24]. Such simulation is expensive in both time and hardware for large and complex networks. Simulating the behavior of even one realization may take weeks of computer time, and computing sufficient realizations to produce accurate statistics is prohibitive. This is one reason that there is widespread interest in statistical analysis of stochastic networks that does *not* depend on raw simulation.

Some approaches use a direct asymptotic analysis; see for example [26, 23, 30]. Other approaches are based on devising a continuum model that approximates particular behaviors of the network; see for example [27, 17, 46, 35, 4, 56]. Some results presented in our recent papers [16, 58, 10], which fall into the second category, use PDEs to approximate large sensor or cellular networks modeled by a certain class of Markov chains. The convergence analysis in this paper is motivated by the network modeling strategies in those papers, and by the need for a rigorous description of the heuristic limiting process underlying the construction of their PDE models. We analyze the convergence of a general Markov chain model in an abstract setting instead of that of any particular network model. We do this for two reasons: first, our network modeling results involve a class of Markov chains modeling a variety of communication networks; second, similar Markov chain models akin to ours arise in several other contexts. For example, a very recent paper [11] on human crowd modeling derives a limiting PDE in a fashion similar to our approach.

In the convergence analysis, for a class of Markov chains, we show that a sequence of such Markov chains indexed by N, the number of components in the system that they model, converges in a certain sense to its continuum limit, which is the solution of a partial differential equation (PDE), as N goes to ∞. The PDE solution describes the global spatio-temporal behavior of the model in the limit of large system size. We apply this abstract result to the modeling of a large wireless sensor network by approximating a particular global aspect of the network states (queue length) by a nonlinear convection-diffusion-reaction PDE. This network model includes the network example discussed in [16] as a special case.

There are well-established mathematical tools to solve PDEs, such as the finite element method [43] and the finite difference method [48], incorporated into computer software packages such as Matlab and Comsol. We can use these tools to greatly reduce computation time, which makes it possible to carry out the analysis, design, and optimization for very large networks.

1.1. *Markov chain model.* We first describe our model in full generality. Consider N points $V_N = \{v_N(1), \dots, v_N(N)\}$ placed over a compact, convex Euclidean domain $\mathcal{D}$ representing a spatial region. We assume that these points form a *uniform* grid, though the model generalizes to nonuniform spacing of points under certain conditions (see Sec. 4 for discussion). We refer to these N points in $\mathcal{D}$ as *grid* points.

We consider a discrete-time Markov chain

$$X_{N,M}(k) = [X_{N,M}(k,1), \dots, X_{N,M}(k,N)]^T \in \mathbb{R}^N$$

(the superscript T represents transpose) whose evolution is described by the stochastic difference equation

$$X_{N,M}(k+1) = X_{N,M}(k) + F_N(X_{N,M}(k)/M, U(k)). \tag{1.1}$$

Here, $X_{N,M}(k,n)$ is the real-valued state associated with the grid point $v_N(n)$ at time k, where $n = 1, \dots, N$ is a *spatial* index and $k = 0, 1, \dots$ is a *temporal* index; $U(k)$ are i.i.d. random vectors that do not depend on the state $X_{N,M}(k)$; M is an "averaging" parameter (explained later); and F_N is a given function.

Treating N and M as indices that grow, the equation (1.1) defines a doubly indexed family $X_{N,M}$ of Markov chains indexed by both N and M. (We will later take M to be a function of N, and treat this family as a sequence X_N of the *single* index N.) Below we give a concrete example of a system described by (1.1).

1.2. *A stochastic network model.* In this subsection we demonstrate the various objects in the abstract Markov chain model analyzed in this paper on a prototypical example. We begin by describing a stochastic model of a wireless sensor network.

Consider a network of wireless sensor nodes uniformly placed over a domain. In a random fashion, the sensor nodes generate data messages that need to be communicated to the destination nodes located on the boundary of the domain, which represent specialized devices that collect the sensor data. The sensor nodes also serve as relays in the routing of the messages to

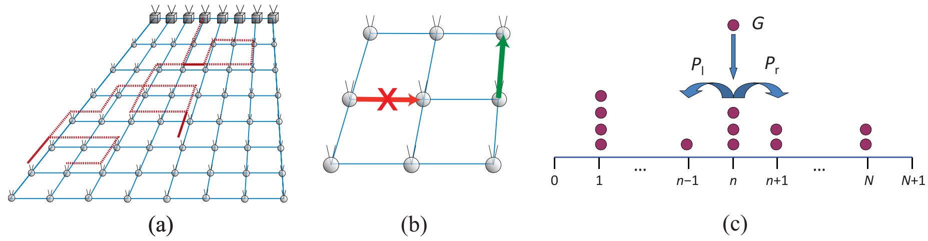


FIG 1. *(a) An illustration of a wireless sensor network over a two-dimensional domain. Destination nodes are located at the far edge. We show the possible path of a message originating from a node located in the left-front region. (b) An illustration of the collision protocol: reception at a node fails when one of its other neighbors transmits (regardless of the intended receiver). (c) An illustration of the time evolution of the queues in the one-dimensional network model.*

the destination nodes. Each sensor node has the capacity to store messages in a *queue*, and is capable of either receiving or transmitting messages from or to its immediate neighbors. (Generalization to further ranges of transmission can be found in our paper [60].) At each time instant $k = 0, 1, \dots,$ each sensor node probabilistically decides to be a transmitter or receiver, but not both. This simplified rule of transmission allows for a relatively simple representation. We illustrate such a network over a two-dimensional domain in Fig. 1. In this network, communication between nodes is interference-limited because all nodes share the same wireless channel. We assume a simple collision protocol: a transmission from a transmitter to a neighboring receiver is successful if and only if none of the other neighbors of the receiver is a transmitter, as illustrated in Fig. 1.

Let us, for the sake of explanation, simplify the problem even further and consider a one-dimensional domain (a two-dimensional example will be given in Sec. 2.4.3). Here, N sensor nodes are equidistributed in an interval $\mathcal{D} \subset \mathbb{R}$ and labeled by $n = 1, \dots, N$. The destination nodes are located on the boundary of $\mathcal{D}$, labeled by $n = 0$ and $n = N + 1$.

Let $G(k, n)$ be the number of messages generated at node n at time k. We model $G(k, n)$ by independent Poisson random variables with mean $g(n)$. We assume that the probability that a node decides to be a transmitter is a function of its normalized queue length (normalized by an "averaging" parameter M). That is, at time k, node n decides to be a transmitter with probability $W(n, X_{N,M}(k, n)/M)$, where $X_{N,M}(k, n)$ is the queue length of node n at time k. We assume that if node n is a transmitter, it randomly chooses to transmit *one* message to the right or the left immediate neighbor with probability $P_r(n)$ and $P_l(n)$, respectively, where $P_r(n) + P_l(n) \leq 1$

(inequality here allows for a more general transmission rule). The special destination nodes at the boundaries of the domain do not have queues; they simply receive any message transmitted to them and never themselves transmit anything. We illustrate the time evolution of the queues in the network in Fig. 1.

The queue lengths $X_{N,M}(k) = [X_{N,M}(k,1), \ldots, X_{N,M}(k,N)]^T \in \mathbb{R}^N$ form a Markov chain network model given by (1.1), where

$$U(k) = [G(k,1), \ldots, G(k,N), Q(k,1), \ldots, Q(k,N), T(k,1), \ldots, T(k,N)]^T$$

is a random vector comprising independent random variables: $G(k,n)$ are as described above; $Q(k,n)$ are uniform random variables on $[0,1]$ used to determine if a transmitter is on or off; and $T(k,n)$ are ternary random variables used to determine the direction a message is passed, which take values R, L, and S (representing transmitting to the right, the left, and giving up the transmission, respectively) with probabilities $P_r(n)$, $P_l(n)$ and $1-(P_r(n)+P_l(n))$, respectively. For a generic $x = [x_1, \ldots, x_N]^T \in \mathbb{R}^N$, the nth component of $F_N(x, U(k))$, where $n = 1, \ldots, N$, is

(1.2)
$$\begin{cases} 1 + G(k,n) & \text{if} \quad Q(k,x_{n-1}) < W(n-1,x_{n-1}),\ T(k,n-1) = R, \\ & \qquad Q(k,x_n) > W(n,x_n),\ Q(k,x_{n+1}) > W(n+1,x_{n+1}), \\ & \text{or} \quad Q(k,x_{n+1}) < W(n+1,x_{n+1}),\ T(k,n+1) = L, \\ & \qquad Q(k,x_n) > W(n,x_n),\ Q(k,x_{n-1}) > W(n-1,x_{n-1}); \\ -1 + G(k,n) & \text{if} \quad Q(k,x_n) < W(n,x_n),\ T(k,n) = L, \\ & \qquad Q(k,x_{n-1}) > W(n-1,x_{n-1}), \\ & \qquad Q(k,x_{n-2}) > W(n-2,x_{n-2}); \\ & \text{or} \quad Q(k,x_n) < W(n,x_n),\ T(k,n) = R, \\ & \qquad Q(k,x_{n+1}) > W(n+1,x_{n+1}), \\ & \qquad Q(k,x_{n+2}) > W(n+2,x_{n+2}); \\ G(k,n) & \qquad \text{otherwise,} \end{cases}$$

where x_n with $n \leq 0$ or $n \geq N+1$ are defined to be zero. Here, the three possible values of the function correspond to the three events that at time k, node n successfully receives one message, successfully transmits one message, and does neither of the above, respectively. The inequalities and equations on the right describe conditions under which these three events occur: for example, $Q(k,x_{n-1}) < W(n-1,x_{n-1})$ corresponds to the choice of node $n-1$ to be a transmitter at time k, $T(k,n-1) = R$ corresponds to its choice to transmit to the right, and so on.

We simplify the situation further by assuming that $W(n,y) = \min(1,y)$. With the collision protocol described earlier, this provides the analog of a network with backpressure routing [49].

After presenting the main results of the paper, we will revisit this network model in Sec. 2.4 and derive a PDE that approximates its global behavior as an application of the main results.

1.3. *Overview of results in this paper.* In this subsection, we provide a brief description of the main results in Sec. 2.

The Markov chain model (1.1) is related to a deterministic difference equation. We set

$$f_N(x) = EF_N(x, U(k)), \quad x \in \mathbb{R}^N, \tag{1.3}$$

and define $x_{N,M}(k) = [x_{N,M}(k,1), \dots, x_{N,M}(k,N)]^T \in \mathbb{R}^N$ by

$$x_{N,M}(k+1) = x_{N,M}(k) + \frac{1}{M} f_N(x_{N,M}(k)), \quad x_{N,M}(0) = \frac{X_{N,M}(0)}{M} \text{ a.s.} \tag{1.4}$$

("a.s." is short for "almost surely").

Example. *For the Markov chain network model introduced in Sec. 1.2, it follows from (1.2) (with the particular choice of $W(n,y) = \min(1,y)$) that for $x = [x_1, \dots, x_N]^T \in [0,1]^N$, the nth component of $f_N(x)$ in its corresponding deterministic difference equation (1.4), where $n = 1, \dots, N$, is (after some tedious algebra, as described in [16])*

$$\begin{aligned} &(1 - x_n)[P_r(n-1)x_{n-1}(1 - x_{n+1}) + P_l(n+1)x_{n+1}(1 - x_{n-1})] \\ &- x_n[P_r(n)(1 - x_{n+1})(1 - x_{n+2}) + P_l(n)(1 - x_{n-1})(1 - x_{n-2})] \\ &+ g(n), \end{aligned} \tag{1.5}$$

where x_n with $n \leq 0$ or $n \geq N+1$ are defined to be zero.

We analyze the convergence of the Markov chain to the solution of a PDE using a two-step procedure. The first step depends heavily on the relation between $X_{N,M}$ and $x_{N,M}$. We show that as $M \to \infty$ and N remains fixed, the difference between $X_{N,M}/M$ and $x_{N,M}$ vanishes, by proving that they both converge in a certain sense to the solution of the same ordinary differential equation (ODE). The basic idea of this convergence is that as the "fluctuation size" of the system decreases and the "fluctuation rate" of the system increases, the stochastic system converges to a deterministic "small-fast-fluctuation" limit, which can be characterized as the solution of a particular ODE. In our case, the smallness of the fluctuation size and largeness of the fluctuation rate is quantified by the "averaging" parameter M. We use a weak convergence theorem of Kushner [39] to prove this convergence.

In the second step, we treat M as a function of N, written M_N (therefore treating X_{N,M_N} and x_{N,M_N} as sequences of the *single* index N, written X_N and x_N, respectively), and show that for *any sequence* $\{M_N\}$ of N, as $N \to \infty$, x_N converges to the solution of a certain PDE (and we show how to construct the PDE). This is essentially a convergence analysis on the truncation error between x_N and the PDE solution. We stress that this is different from the numerical analysis on classical finite difference schemes (see, e.g., [48, 34, 33]), because our difference equation (1.4), which originates from particular system models, differs from those designed specifically for the purpose of numerically *solving* differential equations. The difficulty in our convergence analysis arises from both the different form of (1.4) and the fact that it is in general nonlinear. We provide not only sufficient conditions for the convergence, but also a practical criterion for verifying such conditions otherwise difficult to check.

Finally, based on these two steps, we show that as N and M_N go to ∞ in a *dependent* way, the normalized Markov chain X_N/M_N converges to the PDE solution. We also characterize the rate of convergence. We note that special caution is needed for specifying the details of this dependence between the two indices N and M of the doubly indexed family $X_{N,M}$ of Markov chains in the limiting process.

1.4. *Related literature.* The modeling and analysis of stochastic systems such as networks is a large field of research, and much of the previous contributions share goals with the work in this paper.

Kushner's ODE method, which forms the basis for the first step of our analysis, is closely related to the line of research called stochastic approximation. This line of research, started by Robbins and Monro [53] and Kiefer and Wolfowitz [36] in the early 1950s, studies stochastic processes similar to those addressed by Kushner's ODE method, and has been widely used in many areas (see, e.g., [3, 40], for surveys). These convergence results differ from our results in the sense that they essentially study only the single-step "small-fast-fluctuation" limit as the "averaging" parameter (in our case, M) goes to ∞, but do not have the second-step convergence to the "large-system" PDE limit (as $N \to \infty$). In other words, while Kushner's method and related work deal with a fixed state space with fixed N, we treat a sequence of state spaces $\{\mathbb{R}^N\}$ indexed by increasing N. There are systems in which the "averaging" parameter represents some "size" of the system (e.g., population in epidemic models [19, 45]). However, it is still the case that the convergence requires a fixed dimension of the state space of the Markov chain, like the case of Kushner's ODE convergence, and does not apply to the "large-system" limit in our second step.

Markov chains modeling various systems have been shown to converge to differential equations [38, 18], abstract Cauchy problems [45], or other stochastic processes [39, 20]. These results use methods different from Kushner's, but share with it the principle idea of "averaging out" of the randomness of the Markov chain. Their deeper connection lies in weak convergence theory [39, 20, 6] and methods to prove such convergence that they have in common: the operator semigroup convergence theorem, the martingale characterization method, and identification of the limit as the solution to a stochastic differential equation. The reader is referred to [39, 20] and references therein for additional information on these methods.

There are a variety of other analysis methods for large network systems taking completely different approaches. For example, the well-cited work of Gupta and Kumar [26], followed by many others (e.g., [23, 30]), derives scaling laws of network performance parameters (e.g., throughput); and many efforts based on mean field theory [21, 4, 12] or on the theory of large deviations [56, 51, 52] study the limit of the so-called empirical (or occupancy) measure or distribution, which essentially represents the proportion of components in certain states. These approaches differ from our work because they do not study the spatio-temporal characteristics of the system. Note that we can directly compute many such limiting deterministic characteristics of the network once we have computed the solution of the limiting PDE.

Of course, there do exist numerous continuum models in a wide spectrum of areas that formulate spatio-temporal phenomena (e.g., [7, 31, 59, 11]), many of which use PDEs. All these works differ from the work presented here both by the properties of the system being studied and the analytic approaches. In addition, most of them study distributions of limiting processes that are random, while our limiting functions themselves are deterministic. We especially emphasize the difference between our results and those of the mathematical physics of hydrodynamics [25, 37, 2], because the latter have a similar style by deducing macroscopic behavior from microscopic interactions of individual particles, and in some special cases result in similar PDEs. However, they use an entirely different approach, which usually requires different assumptions on the systems such as translation invariant transition probabilities, conservation of the number of particles, and particular distributions of the initial state; and their limiting PDE is not the direct approximation of system state, but the density of some associated probability measure.

There is a vast literature on the convergence of a large variety of network models different from ours, to essentially two kinds of limits: the fluid limit [13, 17, 46, 35, 42, 22] and the diffusion limit [32, 27, 28, 57, 29, 9, 8], with the latter limit mostly studied in networks in heavy traffic. (Some papers study both limits [14, 50, 15].) Unlike our work, this field of research focuses primarily on networks with a fixed number of nodes.

Our work is to be distinguished from approaches where the model is constructed to be a continuum representation from the start. For example, many papers treat nodes as a continuum by considering only the average density of nodes [54, 55, 1]; and others model network traffic as a continuum by capturing certain average characteristics of the data packet traffic [47, 44, 41].

1.5. *Outline of the paper.* The remainder of the paper is organized as follows. In Sec. 2, we present the main theoretical results and apply the results to the wireless sensor network introduced in Sec. 1.2, and present some numerical experiments. In Sec. 3, we present the proofs of the main results. Finally, we conclude the paper and discuss future work in Sec. 4.

2. Main results and applications.

2.1. *Construction of the limiting PDE.* We begin with the construction of the PDE whose solution describes the limiting behavior of the abstract Markov chain model.

For each N and the grid points $V_N = \{v_N(1), \ldots, v_N(N)\} \subset \mathcal{D}$ as introduced in Sec. 1.1, we denote the distance between any two neighboring grid points by ds_N. For any continuous function $w : \mathcal{D} \to \mathbb{R}$, let y_N be the vector in $\mathbb{R}^N$ composed of the values of w at the grid points $v_N(n)$, i.e., $y_N = [w(v_N(1)), \ldots, w(v_N(N))]^T$. Given a point $s \in \mathcal{D}$, we let $\{s_N\} \subset \mathcal{D}$ be any sequence of *grid points* $s_N \in V_N$ such that as $N \to \infty$, $s_N \to s$. Let $f_N(y_N, s_N)$ be the component of the vector $f_N(y_N)$ corresponding to the location s_N, i.e., if $s_N = v_N(n) \in V_N$, then $f_N(y_N, s_N)$ is the nth component of $f_N(y_N)$.

In order to obtain a limiting PDE, we have to make certain technical assumptions on the asymptotic behavior of the sequence of functions $\{f_N\}$ that insure that $f_N(y_N, s_N)$ is asymptotically close to an expression that looks like the right-hand side of a time-dependent PDE. Such conditions are familiar in the context of PDE limits of Brownian motion. Checking these conditions often amounts to a simple algebraic exercise. We provide a concrete example in Sec. 2.4.

We assume that there exist sequences $\{\delta_N\}$, $\{\beta_N\}$, $\{\gamma_N\}$, and $\{\rho_N\}$, functions f and h, and a constant $c < \infty$, such that as $N \to \infty, \delta_N \to 0, \delta_N/\beta_N \to 0, \gamma_N \to 0, \rho_N \to 0$, and:

- Given s is in the interior of $\mathcal{D}$, there exists a sequence of functions $\{\phi_N\} : \mathcal{D} \to \mathbb{R}$ such that

  $$f_N(y_N, s_N)/\delta_N = f(s_N, w(s_N), \nabla w(s_N), \nabla^2 w(s_N)) + \phi_N(s_N), \tag{2.1}$$

  for any sequence of grid points $s_N \to s$, and for N sufficiently large,

  $$|\phi_N(s_N)| \leq c\gamma_N; \tag{2.2}$$

  and
- Given s on the boundary of $\mathcal{D}$, there exists a sequence of functions $\{\varphi_N\} : \mathcal{D} \to \mathbb{R}$ such that

  $$f_N(y_N, s_N)/\beta_N = h(s_N, w(s_N), \nabla w(s_N), \nabla^2 w(s_N)) + \varphi_N(s_N), \tag{2.3}$$

  for any sequence of grid points $s_N \to s$, and for N sufficiently large, $|\varphi_N(s_N)| \leq c\rho_N$.

Here, $\nabla^i w$ represents all the ith order derivatives of w, where $i = 1, 2$.

Fix $T > 0$ for the rest of this section. Assume that there exists a unique function $z : [0, T] \times \mathcal{D} \to \mathbb{R}$ that solves the limiting PDE

$$\dot{z}(t, s) = f(s, z(t, s), \nabla z(t, s), \nabla^2 z(t, s)), \tag{2.4}$$

with boundary condition

$$h(s, z(t, s), \nabla z(t, s)\nabla^2 z(t, s)) = 0 \tag{2.5}$$

and initial condition $z(0, s) = z_0(s)$. Define

$$dt_{N,M} = \frac{\delta_N}{M},\ t_{N,M}(k) = k\, dt_{N,M},\ K_{N,M} = \left\lfloor \frac{T}{dt_{N,M}} \right\rfloor,\ \tilde{T}_N = \frac{T}{\delta_N}. \tag{2.6}$$

Define

(2.7)

$$z_{N,M}(k, n) = z(t_{N,M}(k), v_N(n)),\ z_{N,M}(k) = [z_{N,M}(k, 1), \ldots, z_{N,M}(k, N)]^T.$$

2.2. *Main results for continuum limits of the abstract Markov chain model.* In this subsection, we present the main theorem, which states that under some conditions, the Markov chain converges uniformly to the PDE solution, as N and M go to ∞ in a *dependent* way. By this we mean that we set M to be a function of N, written M_N, such that $M_N \to \infty$ as $N \to \infty$. Then we can treat $X_{N,M_N}(k)$, $x_{N,M_N}(k)$, $z_{N,M_N}(k)$, dt_{N,M_N}, t_{N,M_N}, and K_{N,M_N} all as sequences of the *single* index N, written $X_N(k)$, $x_N(k)$, $z_N(k)$, dt_N,

t_N, and K_N respectively. We apply such changes of notation throughout the rest of the paper *whenever* M is treated as a function of N.

Let $X_N = [X_N(1)/M_N, \dots, X_N(K_N)/M_N]$, $x_N = [x_N(1), \dots, x_N(K_N)]$, and $z_N = [z_N(1), \dots, z_N(K_N)]$ denote vectors in $\mathbb{R}^{K_N \times N}$. Note that X_N is normalized by M_N. Define

$$\varepsilon_N(k,n) = x_N(k,n) - z_N(k,n), \quad k = 0, \dots, K_N,\ n = 1, \dots, N, \tag{2.8}$$

$\varepsilon_N(k) = [\varepsilon_N(k,1), \dots, \varepsilon_N(k,N)]^T \in \mathbb{R}^N$, and $\varepsilon_N = [\varepsilon_N(1), \dots, \varepsilon_N(K_N)] \in \mathbb{R}^{K_N \times N}$.

We denote the ∞-norms on $\mathbb{R}^N$ and $\mathbb{R}^{K_N \times N}$ both by $\| \cdot \|_\infty^{(N)}$. That is, for $x = [x_1, \dots, x_N]^T \in \mathbb{R}^N$,

$$\|x\|_\infty^{(N)} = \max_{1 \le n \le N} |x_n|;$$

and for $x = [x(1), \dots, x(K_N)] \in \mathbb{R}^{K_N \times N}$, where $x(k) = [x(k,1), \dots, x(k,N)]^T \in \mathbb{R}^N$,

$$\|x\|_\infty^{(N)} = \max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} |x(k,n)|.$$

Define a sequence $u_N(k) \in \mathbb{R}^N$ such that for $k = 0, \dots, K_N - 1$,

$$z_N(k+1) - z_N(k) = \frac{1}{M_N} f_N(z_N(k)) - dt_N u_N(k). \tag{2.9}$$

Define $u_N = [u_N(0), \dots, u_N(K_N - 1)] \in \mathbb{R}^{K_N \times N}$. By (1.4), (2.8), and (2.9),

$$\varepsilon_N(k+1) = \varepsilon_N(k) + \frac{1}{M_N}(f_N(x_N(k)) - f_N(z_N(k))) + dt_N u_N(k). \tag{2.10}$$

Assume that

$$\|\varepsilon_N(0)\|_\infty^{(N)} = 0. \tag{2.11}$$

Then by (1.4) and (2.9), for fixed z_N, x_N is a function of u_N; hence ε_N is a function of u_N. Then by (2.10), there exists a function $H_N : \mathbb{R}^{K_N \times N} \to \mathbb{R}^{K_N \times N}$ such that

$$\varepsilon_N = H_N(u_N). \tag{2.12}$$

Define

$$\mu_N = \lim_{\alpha \to 0} \sup_{\|u\|_\infty^{(N)} \le \alpha} \frac{\|H_N(u)\|_\infty^{(N)}}{\|u\|_\infty^{(N)}}. \tag{2.13}$$

THEOREM 2.1. *Assume that:*

1. *for each N, there exists an identically distributed sequence $\{\lambda(k)\}$ of integrable random variables such that for each k and x, $|F_N(x, U(k))| \leq \lambda(k)$ a.s.;*
2. *for each N, the function $F_N(x, U(k))$ is continuous in x a.s.;*
3. *for each N, the ODE $\dot{y} = f_N(y)$ has a unique solution on $[0, \tilde{T}_N]$ for any initial condition $y(0)$, where $\tilde{T}_N$ is as defined by (2.6);*
4. *z is continuously differentiable in t;*
5. *for each N, (2.11) holds; and*
6. *the sequence $\{\mu_N\}$ is bounded.*

Then a.s., there exist $c < \infty$, N_0, and $\hat{M}_1 < \hat{M}_2 < \hat{M}_3, \ldots$ such that for each $N \geq N_0$ and for each $M_N \geq \hat{M}_N$, $\|X_N - z_N\|_\infty^{(N)} \leq c\gamma_N$, where γ_N is as defined by (2.2).

This theorem states that as N and M_N go to ∞ in a dependent way, X_N converges uniformly to z_N a.s., and with the rate $O(\gamma_N)$. We prove this in Sec. 3.3.

Next, we provide a result treating the convergence of the continuous-time-space extension of the Markov chain $X_{N,M}$ to the limiting PDE solution z. Let $\tilde{T}_N$ be as defined by (2.6). We construct the continuous-time extension $X_{N,M}^{(o)}(\tilde{t})$ of $X_{N,M}(k)$, as the piecewise-constant time interpolant with interval length $1/M$ and normalized by M:

$$X_{N,M}^{(o)}(\tilde{t}) = X_{N,M}(\lfloor M\tilde{t} \rfloor)/M, \quad \tilde{t} \in [0, \tilde{T}_N]. \tag{2.14}$$

Similarly, define the continuous-time extension $x_{N,M}^{(o)}(\tilde{t})$ of $x_{N,M}(k)$ by

$$x_{N,M}^{(o)}(\tilde{t}) = x_{N,M}(\lfloor M\tilde{t} \rfloor), \quad \tilde{t} \in [0, \tilde{T}_N]. \tag{2.15}$$

Respectively, let $X_{N,M}^{(p)}(t, s)$ and $x_{N,M}^{(p)}(t, s)$, where $(t, s) \in [0, T] \times \mathcal{D}$, be the continuous-space extensions of $X_{N,M}^{(o)}(\tilde{t})$ and $x_{N,M}^{(o)}(\tilde{t})$ (with $\tilde{t} \in [0, \tilde{T}_N]$) by piecewise-constant space extensions on $\mathcal{D}$ and with time scaled by δ_N so that the time-interval length is $\delta_N/M := dt_{N,M}$. By *piecewise-constant space extension* of $X_{N,M}^{(o)}$, we mean the piecewise-constant function on $\mathcal{D}$ such that the value of this function at each point in $\mathcal{D}$ is the value of the component of the vector $X_{N,M}^{(o)}$ corresponding to the grid point that is "closest to the left" (taken one component at a time). Then for each t, $X_{N,M}^{(p)}(t, \cdot)$ and $x_{N,M}^{(p)}(t, \cdot)$ are real-valued functions defined on $\mathcal{D}$. We illustrate in Fig. 2.

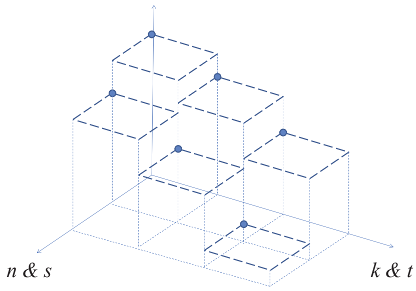


FIG 2. *An illustration of $x_{N,M}$ and $x^{(p)}_{N,M}$ in one dimension, represented by solid dots and dashed-line rectangles, respectively.*

Both $X^{(p)}_{N,M}(t,s)$ and $x^{(p)}_{N,M}(t,s)$ with $(t,s) \in [0,T] \times \mathcal{D}$ are in the space $D^{\mathcal{D}}[0,T]$ of functions from $[0,T] \times \mathcal{D}$ to $\mathbb{R}$ that are Càdlàg with respect to the time component, i.e., right-continuous at each $t \in [0,T)$, and have left-hand limits at each $t \in (0,T]$. Define the ∞-norm $\|\cdot\|^{(p)}_\infty$ on $D^{\mathcal{D}}[0,T]$, i.e., for $x \in D^{\mathcal{D}}[0,T]$,

$$\|x\|^{(p)}_\infty = \sup_{\substack{t\in[0,T]\\ s\in\mathcal{D}}} |x(t,s)|.$$

We again treat M as a function of N, written M_N, therefore treating $X^{(p)}_{N,M_N}$ and $x^{(p)}_{N,M_N}$ as sequences of the *single* index N, written $X^{(p)}_N$ and $x^{(p)}_N$, respectively. The following theorem states that as N and M_N go to ∞ in a dependent way, the continuous-time-space extension of the Markov chain converges uniformly to the PDE solution a.s., and with the rate $O(\max\{\gamma_N, ds_N\})$.

THEOREM 2.2. *Suppose that the assumptions of Theorem 2.1 hold. Then a.s., there exist $c < \infty$, N_0, and $\hat{M}_1 < \hat{M}_2 < \hat{M}_3, \ldots$ such that for each $N \geq N_0$ and for each $M_N \geq \hat{M}_N$, $\|X^{(p)}_N - z\|^{(p)}_\infty \leq c\max\{\gamma_N, ds_N\}$ on $[0,T] \times \mathcal{D}$.*

We prove this in Sec. 3.4.

2.3. *Sufficient conditions on f_N for the boundedness of $\{\mu_N\}$.* The key assumption of Theorems 2.1 and 2.2 is that the sequence $\{\mu_N\}$ is bounded. We present in the following theorem a result that gives specific sufficient conditions on f_N that guarantee that $\{\mu_N\}$ is bounded. This provide a practical criterion to verify this key assumption otherwise difficult to check. Again we treat M as a function of N, written M_N. In Sec. 3.6, we will show that these sufficient conditions hold for the network model described in Sec. 1.2, and use this theorem to prove the convergence of its underlying Markov chain to a PDE.

Consider fixed z_N for each N. We assume that $f_N \in \mathcal{C}^1$ and denote the jacobian matrix of f_N at x by $Df_N(x)$. Define for each N and for $k = 0, \dots, K_N - 1$,

$$A_N(k) = I_N + \frac{1}{M_N} Df_N(z_N(k)), \tag{2.16}$$

where I_N is the identity matrix in $\mathbb{R}^{N\times N}$. We denote the induced ∞-norm on $\mathbb{R}^{N\times N}$ again by $\|\cdot\|_\infty^{(N)}$.

We then have

THEOREM 2.3. *Assume that:*

1. *z is continuously differentiable in t;*
2. *for each N, (2.11) holds;*
3. *for each N, $f_N \in \mathcal{C}^1$; and*
4. *there exists $c < \infty$ such that for N sufficiently large and for $k = 1, \dots, K_N - 1$, $\|A_N(k)\|_\infty^{(N)} \le 1 + c\, dt_N$.*

Then $\{\mu_N\}$ is bounded.

We prove this in Sec. 3.5.

2.4. *Application to network models.* In this subsection, we apply the main results to show how the Markov chain modeling the network introduced in Sec. 1.2 can be approximated by the solution of a PDE. This approximation was heuristically developed in [16].

We first deal with the one-dimensional network model. Its corresponding stochastic and deterministic difference equations (1.1) and (1.4) were specified by (1.2) and (1.5), respectively.

For this model we set δ_N (introduced in Sec. 2.1) to be ds_N^2. Then

$$dt_{N,M} := \delta_N/M = ds_N^2/M. \tag{2.17}$$

Assume that

$$P_l(n) = p_l(v_N(n)) \text{ and } P_r(n) = p_r(v_N(n)), \tag{2.18}$$

where $p_l(s)$ and $p_r(s)$ are real-valued functions defined on $\mathcal{D}$ such that

$$p_l(s) = b(s) + c_l(s) ds_N \text{ and } p_r(s) = b(s) + c_r(s) ds_N. \tag{2.19}$$

Let $c = c_l - c_r$. The values $b(s)$ and $c(s)$ correspond to diffusion and convection quantities in the limiting PDE. Because $p_l(s)+p_r(s) \leq 1$, it is necessary that $b(s) \leq 1/2$. In order to guarantee that the number of messages entering the system from outside over finite time intervals remains finite throughout the limiting process, we set $g(n) = M g_p(v_N(n))dt_N$, where $g_p : \mathcal{D} \to \mathbb{R}$ is called the message generation rate. Assume that b, c_l, c_r, and g_p are in $\mathcal{C}^1$. Further assume that $x_{N,M}(k) \in [0,1]^N$ for each k. Then f_N is in $\mathcal{C}^1$.

We have assumed above that the probabilities P_l and P_r of the direction of transmission are the values of the continuous functions p_l and p_r at the grid points, respectively. This may correspond to stochastic routing schemes where nodes in close vicinity behave similarly based on some local information that they share; or to those with an underlying network-wide directional configuration that are continuous in space, designed to relay messages to destination nodes at known locations. On the other hand, the results can be extended to situations with certain levels of discontinuity, as discussed in Sec. 4.

By these assumptions and definitions, it follow from (1.5) that the function f in (2.4) for this network model is:

$$
\begin{aligned}
f(s, z(t,s), \nabla z(t,s), \nabla^2 z(t,s)) = b(s)\frac{\partial}{\partial s}&\left((1-z(t,s))(1+3z(t,s))z_s(t,s)\right) \\
&+ 2(1-z(t,s))z_s(t,s)b_s(s) \\
&+ z(t,s)(1-z(t,s))^2 b_{ss}(s) \\
&+ \frac{\partial}{\partial s}(c(s)z(t,s)(1-z(t,s))^2) + g_p(s).
\end{aligned}
\tag{2.20}
$$

Here, a single subscript s represents first derivative and a double subscript ss represents second derivative.

Note that the computations needed to obtain (2.20) (and later, (2.21) and (3.22)) require tedious but elementary algebraic manipulations. In practice, we use the symbolic tools in Matlab.

Based on the behavior of nodes $n = 1$ and $n = N$ next to the destination nodes, we derive the boundary condition (2.5) of the PDE of this network. For example, the node $n = 1$ receives messages only from the right and encounters no interference when transmitting to the left. Replacing x_n with $n \leq 0$ or $n \geq N+1$ by 0, it follows that the 1st component of $f_N(x)$ is

$$(1-x_n)P_l(n+1)x_{n+1} - x_n[P_l(n) + P_r(n)(1-x_{n+1})(1-x_{n+2})] + g(n).$$

Similarly, the Nth component of $f_N(x)$ is

$$(1-x_n)P_r(n-1)x_{n-1} - x_n[P_r(n) + P_l(n)(1-x_{n-1})(1-x_{n-2})] + g(n).$$

Set β_N, defined in Sec. 2.1, to be 1. Then from each of the above two functions we get the function h in (2.5) for the one-dimensional network:

$$h(s, z(t,s), \nabla z(t,s), \nabla^2 z(t,s)) = -b(s)z(s)^3 + b(s)z(s)^2 - b(s)z(s). \tag{2.21}$$

Note that the function h is the limit of $f_N(y_N, s_N)/\beta_N$, not $f_N(y_N, s_N)/\delta_N$ (whose limit is f). Solving $h = 0$ for real z, we have the boundary condition $z(t, s) = 0$.

Let z be the solution of the PDE (2.4) with f specified by (2.20) and with boundary condition $z(t, s) = 0$ and initial condition $z(0, s) = z_0(s)$. Assume that (2.11) holds. As in Sec. 2.2, we treat M as a sequence of N, written M_N. Then X_N and z_N are as in the general case. In the following theorem we show the convergence of the Markov chain modeling the one-dimensional network to the PDE solution.

THEOREM 2.4. *For the one-dimensional network model, assume that the function* $\max\{|z|, |z_s|, |z_{ss}|, |b_s|, |b_{ss}|, |c|, |c_s|\}$ *is bounded on* $[0, T] \times \mathcal{D}$. *Then a.s., there exist* $c < \infty$, N_0, *and* $\hat{M}_1 < \hat{M}_2 < \hat{M}_3, \ldots$ *such that for each* $N \geq N_0$ *and for each* $M_N \geq \hat{M}_N$, $\|X_N - z_N\|_\infty^{(N)} \leq c\, ds_N$.

We prove this in Sec. 3.6.

There is an analogous result for the continuous-time-space extension $X_N^{(p)}$.

2.4.1. *Interpretation of limiting PDE.* Now we make some remarks on how to interpret a given limiting PDE. First, for fixed N and M, the normalized queue length of node n at time k, is approximated by the value of the PDE solution z at the corresponding point in $[0, T] \times \mathcal{D}$, i.e., $\frac{X_{N,M}(k,n)}{M} \approx z(t_{N,M}(k), v_N(n))$.

Second, we discuss how to interpret $C(t_o) := \int_{\mathcal{D}} z(t_o, s)ds$, the area below the curve $z(t_o, s)$ for fixed $t_o \in [0, T]$. Let $k_o = \lfloor t_o/dt_{N,M} \rfloor$. Then we have that $z(t_o, v_N(n))ds_N \approx \frac{X_{N,M}(k_o,n)}{M} ds_N$, the area of the nth rectangle in Fig. 3. Therefore

$$C(t_o) \approx \sum_{n=1}^{N} z(t_o, v_N(n))ds_N \approx \sum_{n=1}^{N} \frac{X_{N,M}(k_o, n)}{M} ds_N,$$

the sum of all rectangles. If we assume that all messages in the queue have roughly the same bits, and think of ds_N as the "coverage" of each node, then the area under any segment of the curve measures a kind of "data-coverage product" of the nodes covered by the segment, in the unit of "bit·meter." As $N \to \infty$, the total normalized queue length $\sum_{n=1}^{N} X_{N,M}(k_o, n)/M$ of the network does go to ∞; however, the coverage ds_N of each node goes to 0. Hence the sum of the "data-coverage product" can be approximated by the finite area $C(t_o)$.

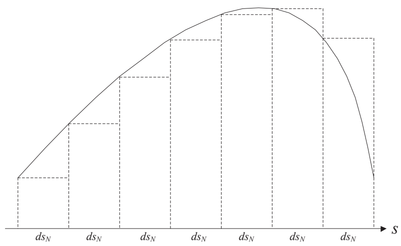


FIG 3. *The PDE solution at a fixed time that approximates the normalized queue lengths of the network.*

2.4.2. *Comparison between the PDE solution and Monte Carlo simulations of the network.* We compare the limiting PDE solution with Monte Carlo simulations for a network over the domain $\mathcal{D} = [-1, 1]$. We use the initial condition $z_0(s) = l_1 e^{-s^2}$, where $l_1 > 0$ is a constant, so that initially the nodes in the middle have messages to transmit, while those near the boundaries have very few. We set the message generation rate $g_p(s) = l_2 e^{-s^2}$, where $l_2 > 0$ is a parameter determining the total load of the system.

We use three sets of values of $N = 20, 50, 80$ and $M = N^3$, and show the PDE solution and the Monte Carlo simulation results with different N and M at $t = 1s$. The networks have diffusion $b = 1/2$ and convection $c = 0$ in Fig. 4 and $c = 1$ in Fig. 5, respectively, where the x-axis denotes the node location and y-axis denotes the normalized queue length.

For the three sets of the values of $N = 20, 50, 80$ and $M = N^3$, with $c = 0$, the maximum absolute errors of the PDE approximation are 5.6×10^{-3}, 1.3×10^{-3}, and 1.1×10^{-3}, respectively; and with $c = 1$, the errors are 4.4×10^{-3}, 1.5×10^{-3}, and 1.1×10^{-3}, respectively. As we can see, as N and M increase, the resemblance between the Monte Carlo simulations and the PDE solution becomes stronger. In the case of very large N and M, it is difficult to distinguish the results.

We stress that the PDEs only took fractions of a second to solve on a computer, while the Monte Carlo simulations took on the order of tens of hours.

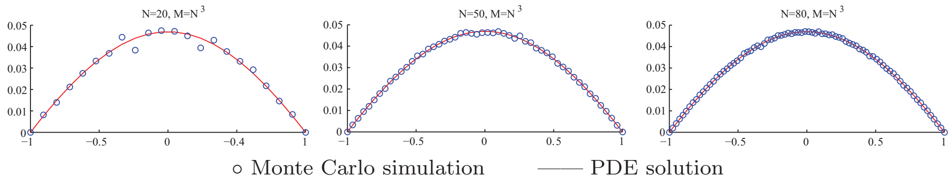


Fig 4. *The Monte Carlo simulations (with different N and M) and the PDE solution of a one-dimensional network, with $b = 1/2$ and $c = 0$, at $t = 1s$.*

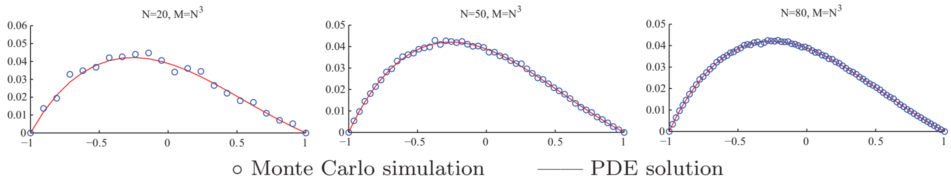


Fig 5. *The Monte Carlo simulations (with different N and M) and the PDE solution of a one-dimensional network, with $b = 1/2$ and $c = 1$, at $t = 1s$.*

2.4.3. *A two dimensional network.* The generalization of the continuum model to higher dimensions is straightforward, except for more arduous algebraic manipulation. Likewise, the convergence analysis is similar to the one dimensional case.

We consider the two-dimensional network of $N = N_1 \times N_2$ sensor nodes. The nodes are uniformly placed over a domain $\mathcal{D} \subset \mathbb{R}^2$ and labeled by (n, m), where $n = 1, \ldots, N_1$ and $m = 1, \ldots, N_2$. Denote the grid point in $\mathcal{D}$ corresponding to node (n, m) by $v_N(n, m)$. Again let the distance between any two neighboring nodes be ds_N. Assume that the node labeled by (n, m) randomly chooses to transmit to the east, west, north, or south immediate neighbor with probabilities $P_e(n, m) = b_1(v_N(n, m)) + c_e(v_N(n, m))ds_N$, $P_w(n, m) = b_1(v_N(n, m)) + c_w(v_N(n, m))ds_N$, $P_n(n, m) = b_2(v_N(n, m)) + c_n(v_N(n, m))ds_N$, and $P_s(n, m) = b_2(v_N(n, m)) + c_s(v_N(n, m))ds_N$, respectively, where $P_e(n, m) + P_w(n, m) + P_n(n, m) + P_s(n, m) \leq 1$. Therefore it is necessary that $b_1(s) + b_2(s) \leq 1/2$. Define $c_1 = c_w - c_e$ and $c_2 = c_s - c_n$.

The derivation of the limiting PDE is similar to those of the one-dimensional case, except that we now have to consider transmission to and interference from four directions instead of two. We present the limiting PDE here without the detailed derivation:

$$\dot{z} = \sum_{j=1}^{2} b_j \frac{\partial}{\partial s_j} \left((1 + 5z)(1 - z)^3 \frac{\partial z}{\partial s_j} \right) + 2(1 - z)^3 \frac{\partial z}{\partial s_j} \frac{d b_j}{d s_j}$$

$$+ z(1-z)^4 \frac{d^2 b_j}{ds_j^2} + \frac{\partial}{\partial s_j}\left(c_j z(1-z)^4\right) + g_p,$$

with boundary condition $z(t,s) = 0$ and initial condition $z(0,s) = z_0(s)$, where $t \in [0,T]$ and $s = (s_1, s_2) \in \mathcal{D}$.

We now compare the PDE approximation and the Monte Carlo simulations of a network over the domain $\mathcal{D} = [-1,1] \times [-1,1]$. We use the initial condition $z_0(s) = l_1 e^{-(s_1^2+s_2^2)}$, where $l_1 > 0$ is a constant. We set the message generation rate $g_p(s) = l_2 e^{-(s_1^2+s_2^2)}$, where $l_2 > 0$ is a constant.

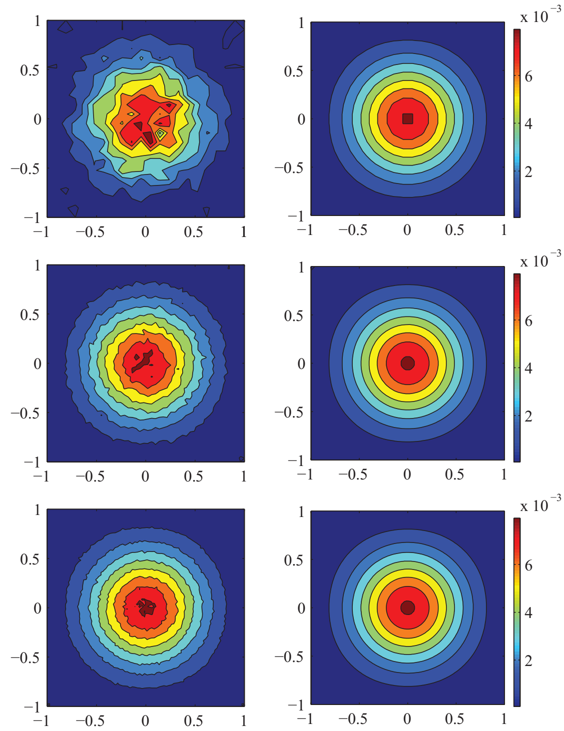
We use three different sets of the values of $N_1 \times N_2$ and M, where $N_1 = N_2 = 20, 50, 80$ and $M = N_1^3$. We show the contours of the normalized queue length from the PDE solution and the Monte Carlo simulation results with different sets of values of N_1, N_2, and M, at $t = 0.1s$. The networks have diffusion $b_1 = b_2 = 1/4$ and convection $c_1 = c_2 = 0$ in Fig. 6 and $c_1 = -2, c_2 = -4$ in Fig. 7, respectively.

For the three sets of values of $N_1 = N_2 = 20, 50, 80$ and $M = N_1^3$ and with $c_1 = c_2 = 0$, the maximum absolute errors are 3.2×10^{-3}, 1.1×10^{-3}, and 6.8×10^{-4}, respectively; and with $c_1 = -2, c_2 = -4$, the errors are 4.1×10^{-3}, 1.0×10^{-3}, and 6.6×10^{-4}, respectively. Again the accuracy of the continuum model increases with N_1, N_2, and M.

It took 3 days to do the Monte Carlo simulation of the network at $t = 0.1s$ with 80×80 nodes and the maximum queue length $M = 80^3$, while the PDE solved on the same machine took less than a second. We could not do Monte Carlo simulations of any larger networks or greater values of t because of prohibitively long computation time.

3. Proofs of the main results. This section is devoted solely to the proofs of the results in Sec. 2. As such, the material here is highly technical and might be tedious to follow in detail, though we have tried our best to make it as readable as possible. The reader can safely skip this section without doing violence to the main ideas of the paper, though much of our hard work is reflected here.

We first prove Theorems 2.1 and 2.2 by analyzing the convergence of the Markov chains $X_{N,M}$ to the solution of the limiting PDE in a two-step procedure. In the first step, for fixed N, we show in Sec. 3.1 that as $M \to \infty$, $X_{N,M}/M$ converges to $x_{N,M}$. In the second step, we treat M as a function of N, written M_N, and for any sequence $\{M_N\}$, we show in Sec. 3.2 that as $N \to \infty$, x_N converges to the PDE solution. Based on the two steps, we show in Sec. 3.3 that as N and M_N go to ∞ in a dependent way, X_N/M_N converges to the PDE solution, proving Theorem 2.1; and we prove Theorem 2.2 in Sec. 3.4.



Monte Carlo simulations PDE solution

FIG 6. *The Monte Carlo simulations (from top to bottom, with $N_1 = N_2 = 20, 50, 80$, respectively, and $M = N_1^3$) and the PDE solution of a two-dimensional network, with $b_1 = b_2 = 1/4$ and $c_1 = c_2 = 0$, at $t = 0.1s$.*

We then prove Theorem 2.3 in Sec. 3.5, and use it to prove Theorem 2.4 in Sec. 3.6.

3.1. *Convergence of $X_{N,M}$ and $x_{N,M}$ to the solution of the same ODE.* In this subsection, we show that for fixed N, $X_{N,M}/M$ and $x_{N,M}$ are close in a certain sense for large M under certain conditions, by proving that both their continuous-time extensions converge to the solution of the same ODE.

For fixed T and N, by (2.6), $\tilde{T}_N$ is fixed. As defined by (2.14) and (2.15) respectively, both $X_{N,M}^{(o)}(\tilde{t})$ and $x_{N,M}^{(o)}(\tilde{t})$ with $\tilde{t} \in [0, \tilde{T}_N]$ are in the space $D^N[0, \tilde{T}_N]$ of $\mathbb{R}^N$-valued Càdlàg functions on $[0, \tilde{T}_N]$. Since they both depend on M, each one of them forms a sequence of functions in $D^N[0, \tilde{T}_N]$ indexed by $M = 1, 2, \ldots$. Define the ∞-norm $\|\cdot\|_\infty^{(o)}$ on $D^N[0, \tilde{T}_N]$, i.e., for $x \in$

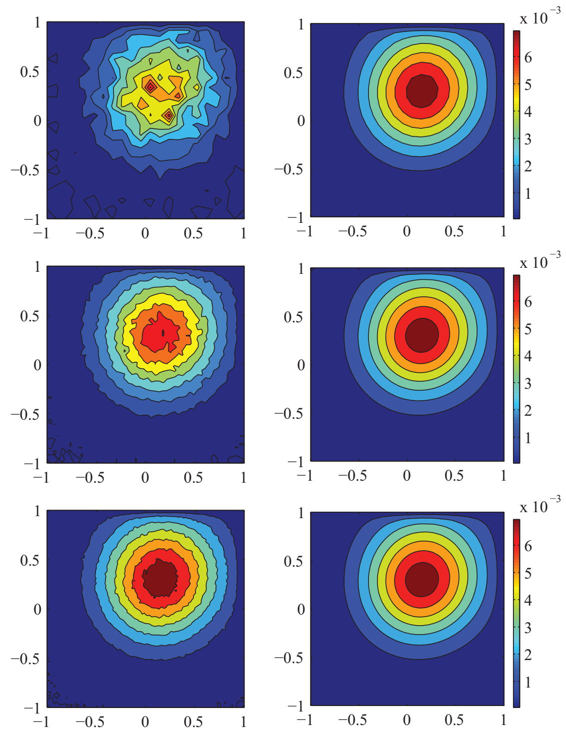


FIG 7. *The Monte Carlo simulations (from top to bottom, with $N_1 = N_2 = 20, 50, 80$, respectively, and $M = N_1^3$) and the PDE solution of a two-dimensional network, with $b_1 = b_2 = 1/4$ and $c_1 = -2, c_2 = -4$, at $t = 0.1s$.*

$D^N[0, \tilde{T}_N]$,

$$\|x\|_\infty^{(o)} = \max_{n=1,\ldots,N} \sup_{t\in[0,\tilde{T}_N]} |x_n(t)|,$$

where x_n is the nth components of x.

Now we present a lemma stating that under some conditions, as $M \to \infty$, $X_{N,M}^{(o)}$ converges uniformly to the solution of the ODE $\dot{y} = f_N(y)$, and $x_{N,M}^{(o)}$ converges uniformly to the same solution, both on $[0, \tilde{T}_N]$.

LEMMA 1. *Assume that:*

1. *there exists an identically distributed sequence $\{\lambda(k)\}$ of integrable random variables such that for each k and x, $|F_N(x, U(k))| \leq \lambda(k)$ a.s.;*
2. *the function $F_N(x, U(k))$ is continuous in x a.s.; and*
3. *the ODE $\dot{y} = f_N(y)$ has a unique solution on $[0, \tilde{T}_N]$ for any initial condition $y(0)$.*

Suppose that as $M \to \infty$, $X_{N,M}^{(o)}(0) \xrightarrow{P} y(0)$ *and* $x_{N,M}^{(o)}(0) \to y(0)$, *where* "$\xrightarrow{P}$" *represents convergence in probability. Then, as* $M \to \infty$, $\|X_{N,M}^{(o)} - y\|_\infty^{(o)} \xrightarrow{P} 0$ *and* $\|x_{N,M}^{(o)} - y\|_\infty^{(o)} \to 0$ *on* $[0, \tilde{T}_N]$, *where* y *is the unique solution of* $\dot{y} = f_N(y)$ *with initial condition* $y(0)$.

To prove Lemma 1, we first present a lemma due to Kushner [39].

LEMMA 2. *Assume that:*

$\hat{1}$. *the set* $\{|F_N(x, U(k))| : k \geq 0\}$ *is uniformly integrable;*
$\hat{2}$. *for each* k *and each bounded random variable* X,

$$\lim_{\delta \to 0} E \sup_{|Y| \leq \delta} |F_N(X, U(k)) - F_N(X + Y, U(k))| = 0;$$

and
$\hat{3}$. *there is a function* $\hat{f}_N(\cdot)$ [*continuous by* $\hat{2}$] *such that as* $n \to \infty$,

$$\frac{1}{n} \sum_{k=0}^{n} F_N(x, U(k)) \xrightarrow{P} \hat{f}_N(x).$$

Suppose that $\dot{y} = \hat{f}_N(y)$ *has a unique solution on* $[0, \tilde{T}_N]$ *for any initial condition, and that* $X_{N,M}^{(o)}(0) \Rightarrow y(0)$, *where* "$\Rightarrow$" *represents weak convergence. Then as* $M \to \infty$, $\|X_{N,M}^{(o)} - y\|_\infty^{(o)} \Rightarrow 0$ *on* $[0, \tilde{T}_N]$.

We note that in Kushner's original theorem, the convergence of $X_{N,M}^{(o)}$ to y is stated in terms of Skorokhod norm [39], but it is equivalent to the ∞-norm in our case where the time interval $[0, \tilde{T}_N]$ is finite and the limit y is continuous [5].

We now prove Lemma 1 by showing that the assumptions $\hat{1}$–$\hat{3}$ of Lemma 2 hold under the assumptions 1–3 of Lemma 1.

Proof of Lemma 1: Since $\lambda(k)$ is integrable, as $a \to \infty$, $E|\lambda(k)|1_{\{|\lambda(k)|>a\}} \to 0$, where 1_A is the indicator function of set A. By Assumption 1, for each k, x, and $a > 0$,

$$\begin{aligned} E|F_N(x, U(k))|1_{\{|F_N(x,U(k))|>a\}} &\leq E|\lambda(k)|1_{\{|F_N(x,U(k))|>a\}} \\ &\leq E|\lambda(k)|1_{\{|\lambda(k)|>a\}}. \end{aligned}$$

Therefore as $a \to \infty$,

$$\sup_{k \geq 0} E|F_N(x, U(k))|1_{\{|F_N(x,U(k))|>a\}} \to 0,$$

i.e., the family $\{|F_N(x, U(k))| : k \geq 0\}$ is uniformly integrable, and hence Assumption $\hat{1}$ holds.

By Assumption 2, for each k and each bounded X, a.s.,

$$\lim_{\delta\to 0} \sup_{|Y|\leq\delta} |F_N(X, U(k)) - F_N(X+Y, U(k))| = 0.$$

By Assumption 1, for each k and each bounded X and Y, a.s.,

$$\begin{aligned}&|F_N(X, U(k)) - F_N(X+Y, U(k))|\\ &\leq |F_N(X, U(k))| + |F_N(X+Y, U(k))| \leq 2\lambda(k).\end{aligned}$$

Therefore for each k, each bounded X, and each δ, a.s.,

$$\left| \sup_{|Y|\leq\delta} |F_N(X, U(k)) - F_N(X+Y, U(k))| \right| \leq 2\lambda(k),$$

an integrable random variable. By the dominant convergence theorem,

$$\begin{aligned}&\lim_{\delta\to 0} E \sup_{|Y|\leq\delta} |F_N(X, U(k)) - F_N(X+Y, U(k))|\\ &= E \lim_{\delta\to 0} \sup_{|Y|\leq\delta} |F_N(X, U(k)) - F_N(X+Y, U(k))| = 0.\end{aligned}$$

Hence Assumption $\hat{2}$ holds.

Since $U(k)$ are i.i.d., by the weak law of large numbers and the definition of f_N in (1.3), as $n \to \infty$,

$$\frac{1}{n} \sum_{k=0}^{n} F_N(x, U(k)) \xrightarrow{P} f_N(x).$$

Hence Assumption $\hat{3}$ holds.

Therefore, by Lemma 2, as $M \to \infty$, $\|X_{N,M}^{(o)} - y\|_\infty^{(o)} \Rightarrow 0$ on $[0, \tilde{T}_N]$. For any sequence of random processes $\{X_n\}$, if A is a constant, $X_n \Rightarrow A$ if and only if $X_n \xrightarrow{P} A$. Therefore, as $M \to \infty$, $\|X_{N,M}^{(o)} - y\|_\infty^{(o)} \xrightarrow{P} 0$ on $[0, \tilde{T}_N]$. The same argument implies the deterministic convergence of $x_{N,M}^{(o)}$: as $M \to \infty$, $\|x_{N,M}^{(o)} - y\|_\infty^{(o)} \to 0$ on $[0, \tilde{T}_N]$. ∎

Based on Lemma 1, we get the following lemma, which states that $X_{N,M}^{(o)}$ and $x_{N,M}^{(o)}$ are close with high probability for large M.

Lemma 3. *Let the assumptions of Lemma 1 hold. Then for any sequence* $\{\zeta_N\}$*, for each* N *and for* M *sufficiently large,*

$$P\{\|X_{N,M}^{(o)} - x_{N,M}^{(o)}\|_\infty^{(o)} > \zeta_N\} \le 1/N^2 \text{ on } [0, \tilde{T}_N].$$

Proof. By the triangle inequality,

$$\|X_{N,M}^{(o)} - x_{N,M}^{(o)}\|_\infty^{(o)} \le \|X_{N,M}^{(o)} - y\|_\infty^{(o)} + \|x_{N,M}^{(o)} - y\|_\infty^{(o)}.$$

By Lemma 1, for each N, as $M \to \infty$, $\|X_{N,M}^{(o)} - x_{N,M}^{(o)}\|_\infty^{(o)} \xrightarrow{P} 0$ on $[0, \tilde{T}_N]$. This finishes the proof. □

Since $X_{N,M}^{(o)}$ and $x_{N,M}^{(o)}$ are the continuous-time extensions of $X_{N,M}$ and $x_{N,M}$ by piecewise-constant extensions, respectively, we have the following corollary stating that for each N, as $M \to \infty$, $X_{N,M}/M$ converges uniformly to $x_{N,M}$.

Corollary 1. *Let the assumptions of Lemma 1 hold. Then for any sequence* $\{\zeta_N\}$*, for each* N *and for* M *sufficiently large, we have that*

$$P\left\{\max_{\substack{k=0,\ldots,K_{N,M}\\ n=1,\ldots,N}} \left|\frac{X_{N,M}(k,n)}{M} - x_{N,M}(k,n)\right| > \zeta_N\right\} \le \frac{1}{N^2}.$$

3.2. *Convergence of* x_N *to the limiting PDE.* For the remainder of this section, we treat M as a function of N, written M_N. We state conditions under which x_N is asymptotically close to z_N for *any sequence* $\{M_N\}$ as $N \to \infty$. The basic idea is this. Recall that $x_N(k)$ is defined by (1.4). Suppose that we associate the discrete time k with points on the real line spaced apart by a distance proportional to δ_N. Then, the technical assumptions (2.1) and (2.3) imply that $x_N(k)$ is, in a certain sense, close to the solution of the limiting PDE (2.4) with boundary condition (2.5). The remainder of this subsection is devoted to developing this argument rigorously.

Lemma 4. *Assume that:*

1. z *is continuously differentiable in* t*;*
2. *for each* N*, (2.11) holds; and*
3. *the sequence* $\{\mu_N\}$ *is bounded.*

Then there exists $c < \infty$ *such that for any sequence* $\{M_N\}$ *and for* N *sufficiently large,* $\|\varepsilon_N\|_\infty^{(N)} \le c \max\{\gamma_N, dt_N\}$.

PROOF. Since z is continuously differentiable in t, by (2.6), there exists $c_1 < \infty$ such that for each N for $k = 0, \dots, K_N - 1$ and $n = 1, \dots, N$, there exists a function $r_N : [0, T] \times \mathcal{D} \to \mathbb{R}$ such that

$$\begin{aligned} \frac{z_N(k+1,n) - z_N(k,n)}{dt_N} &= \frac{z(t_N(k+1), v_N(n)) - z(t_N(k), v_N(n))}{dt_N} \\ &= \dot{z}(t_N(k), v_N(n)) + r_N(t_N(k), v_N(n)), \end{aligned} \tag{3.1}$$

and for N sufficiently large, $|r_N(t_N(k), v_N(n))| \le c_1 dt_N$.

By (2.1), (2.4), and (2.6), there exists $c_2 < \infty$ such that there exists a function $\phi_N : [0, T] \times \mathcal{D} \to \mathbb{R}$ such that

$$\begin{aligned} &\dot{z}(t_N(k), v_N(n)) \\ &= f(v_N(n), z(t_N(k), v_N(n)), \nabla z(t_N(k), v_N(n)), \nabla^2 z(t_N(k), v_N(n))) \\ &= f_N(z_N(k), v_N(n))/\delta_N + \phi_N(t_N(k), v_N(n)), \end{aligned} \tag{3.2}$$

and for N sufficiently large, $|\phi_N(t_N(k), v_N(n))| \le c_2 \gamma_N$, where γ_N is as defined by (2.2). By (3.1) and (3.2),

$$\begin{aligned} \frac{z_N(k+1,n) - z_N(k,n)}{dt_N} = \frac{f_N(z_N(k), v_N(n))}{\delta_N} &+ \phi_N(t_N(k), v_N(n)) \\ &+ r_N(t_N(k), v_N(n)). \end{aligned}$$

Notice that $f_N(z_N(k), v_N(n))$ is the nth component of $f_N(z_N(k))$. Let

$$u_N(k,n) = -\phi_N(t_N(k), v_N(n)) - r_N(t_N(k), v_N(n)).$$

Then by (2.6) and (2.9), $u_N(k) = [u_N(k,1), \dots, u_N(k,N)]^T$, and there exists $c_3 < \infty$ such that N sufficiently large,

$$\|u_N\|_\infty^{(N)} \le c_3 \max\{\gamma_N, dt_N\}. \tag{3.3}$$

By the definition of μ_N (2.13), for each N, there exists $\delta > 0$ such that for $\alpha < \delta$,

$$\sup_{\|u\|_\infty^{(N)} \le \alpha} \frac{\|H_N(u)\|_\infty^{(N)}}{\|u\|_\infty^{(N)}} \le \mu_N + 1.$$

By (3.3), there exists α_1 such that for N sufficiently large, $\|u_N\|_\infty^{(N)} \le \alpha_1 < \delta$, and hence

$$\frac{\|H_N(u_N)\|_\infty^{(N)}}{\|u_N\|_\infty^{(N)}} \le \sup_{\|u\|_\infty^{(N)} \le \alpha_1} \frac{\|H_N(u)\|_\infty^{(N)}}{\|u\|_\infty^{(N)}} \le \mu_N + 1.$$

Therefore, since the sequence $\{\mu_N\}$ is bounded, there exists $c < \infty$ such that for N sufficiently large,

$$\|\varepsilon_N\|_\infty^{(N)} = \|H_N(u_N)\|_\infty^{(N)} \leq (\mu_N + 1)\|u_N\|_\infty^{(N)} \leq c \max\{\gamma_N, dt_N\}.$$

Because the derivation above does not depend on the choice of the sequence $\{M_N\}$, the proof is finished. □

This lemma states that for any sequence $\{M_N\}$, as $N \to \infty$, x_N converges uniformly to z_N with the rate $O(\max\{\gamma_N, dt_N\})$.

3.3. *Proof of Theorem 2.1.* We now prove Theorem 2.1.

Proof of Theorem 2.1: By (2.6), there exists a sequence $\{\bar{M}_N\}$ such that if for each N, $M_N \geq \bar{M}_N$, then for N sufficiently large, $\gamma_N \geq dt_N := \delta_N / M_N$, and hence by Lemma 4, there exists $c_1 < \infty$ such that $\|\varepsilon_N\|_\infty^{(N)} \leq c_1 \gamma_N$.

By Corollary 1, there exists a sequence $\{\tilde{M}_N\}$ such that if for each N, $M_N \geq \tilde{M}_N$, then

$$\sum_{N=1}^{\infty} P\{\|X_N - x_N\|_\infty^{(N)} > \gamma_N\} \leq \sum_{N=1}^{\infty} 1/N^2 < \infty.$$

It follows from the first Borel-Cantelli Lemma that

$$P\left\{\limsup_{N\to\infty} \{\|X_N - x_N\|_\infty^{(N)} > \gamma_N\}\right\} = 0,$$

which implies that a.s., for N sufficiently large and for $M_N \geq \tilde{M}_N$, $\|X_N - x_N\|_\infty^{(N)} \leq \gamma_N$.

By the triangle inequality,

$$\|X_N - z_N\|_\infty^{(N)} \leq \|X_N - x_N\|_\infty^{(N)} + \|\varepsilon_N\|_\infty^{(N)}.$$

Setting $\hat{M}_N = \max\{\bar{M}_N, \tilde{M}_N\}$ finishes the proof. ■

3.4. *Proof of Theorem 2.2.* We now prove Theorem 2.2.

Proof of Theorem 2.2: For each N, by the definition of $x_N^{(p)}$, we have that $x_N^{(p)}(t_N(k), v_N(n)) = x_N(k, n)$. Let $\Omega_N(k, n)$ be the subset of $[0, T] \times \mathcal{D}$ containing $(t_N(k), v_N(n))$ where $x_N^{(p)}$ is piecewise constant, i.e., $(t_N(k), v_N(n)) \in \Omega_N(k, n)$ and for all $(t, s) \in \Omega_N(k, n)$, $x_N^{(p)}(t, s) = x_N^{(p)}(t_N(k), v_N(n))$. Then

$$\|x_N^{(p)} - z\|_\infty^{(p)} \leq \|\varepsilon_N\|_\infty^{(N)} + \max_{\substack{k=0,\ldots,K_N \\ n=1,\ldots,N}} \sup_{(t,s)\in\Omega_N(k,n)} |z(t_N(k), v_N(n)) - z(t, s)|.$$

Since $z(t,s)$ is continuously differentiable in t on a compact domain, it is Lipschitz continuous in t. Similarly, it is Lipschitz continuous in s. Then there exist $c_1, c_2 \leq \infty$ such that for some norm $\|\cdot\|$ on $[0,T]\times\mathcal{D}$,

$$\begin{aligned}&\max_{\substack{k=0,\dots,K_N\\ n=1,\dots,N}} \sup_{(t,s)\in\Omega_N(k,n)} |z(t_N(k),v_N(n)) - z(t,s)|\\ &\leq c_1 \max_{\substack{k=0,\dots,K_N\\ n=1,\dots,N}} \sup_{(t,s)\in\Omega_N(k,n)} \|(t_N(k),v_N(n)) - (t,s)\| \leq c_2 \max\{ds_N, dt_N\}.\end{aligned}$$

Therefore, by Lemma 4, there exists $c_3 < \infty$ such that for N sufficiently large, $\|x_N^{(p)} - z\|_\infty^{(p)} \leq c_3 \max\{\gamma_N, dt_N, ds_N\}$.

By (2.6), there exists a sequence $\{\bar{M}_N\}$ such that if for each N, $M_N \geq \bar{M}_N$, then for N sufficiently large, $\max\{\gamma_N, ds_N\} \geq dt_N$, and hence there exists $c_4 < \infty$ such that $\|x_N^{(p)} - z\|_\infty^{(p)} \leq c_4 \max\{\gamma_N, ds_N\}$.

Since M_N is a function of N, $X_{N,M_N}^{(o)}$ and $x_{N,M_N}^{(o)}$ can be treated as sequences of the single index N, written $X_N^{(o)}$ and $x_N^{(o)}$, respectively. By Lemma 3, there exists a sequence $\{\tilde{M}_N\}$ such that if for each N, $M_N \geq \tilde{M}_N$, then

$$\sum_{N=1}^{\infty} P\{\|X_N^{(o)} - x_N^{(o)}\|_\infty^{(o)} > \max\{\gamma_N, ds_N\}\} \leq \sum_{N=1}^{\infty} 1/N^2 < \infty.$$

Using arguments analogous to those in the last proof, one can show that a.s., for N sufficiently large and for $M_N \geq \tilde{M}_N$, $\|X_N^{(o)} - x_N^{(o)}\|_\infty^{(o)} \leq \max\{\gamma_N, ds_N\}$ on $[0,\tilde{T}_N]$. Since $X_N^{(p)}$ and $x_N^{(p)}$ are the continuous-space extensions of $X_N^{(o)}$ and $x_N^{(o)}$ by piecewise-constant extensions, respectively, it follows that a.s., for N sufficiently large and for $M_N \geq \tilde{M}_N$, $\|X_N^{(p)} - x_N^{(p)}\|_\infty^{(p)} \leq \max\{\gamma_N, ds_N\}$ on $[0,T]\times\mathcal{D}$.

By the triangle inequality,

$$\|X_N^{(p)} - z\|_\infty^{(p)} \leq \|X_N^{(p)} - x_N^{(p)}\|_\infty^{(p)} + \|x_N^{(p)} - z\|_\infty^{(p)}.$$

Setting $\hat{M}_N = \max\{\bar{M}_N, \tilde{M}_N\}$ finishes the proof. ∎

3.5. *Proof of Theorem 2.3.* To prove Theorem 2.3, we first prove Lemma 5 and 6 presented below. First we provide in Lemma 5 a sequence bounding $\{\mu_N\}$ from above. By (2.12), for each N, for $k = 1,\dots,K_N$ and $n = 1,\dots,N$, we can write $\varepsilon_N(k,n) = H_N^{(k,n)}(u_N)$, where $H_N^{(k,n)}$ is from $\mathbb{R}^{K_N\times N}$ to $\mathbb{R}$. Suppose that H_N is differentiable at 0. Define

$$DH_N = \max_{\substack{k=1,\dots,K_N\\ n=1,\dots,N}} \sum_{\substack{i=1,\dots,K_N\\ j=1,\dots,N}} \left|\frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0)\right|, \tag{3.4}$$

where 0 is in $\mathbb{R}^{K_N \times N}$. Notice that DH_N is essentially the induced ∞-norm of the linearized version of the operator H_N.

LEMMA 5. *Assume that:*

1. *z is continuously differentiable in t;*
2. *for each N, (2.11) holds; and*
3. *for each N, $H_N \in \mathcal{C}^1$ locally at 0.*

Then we have that for each N, $\mu_N \leq DH_N$.

PROOF. For all $u \neq 0$, we have by the definition of DH_N that

$$DH_N \geq \frac{\max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left| \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) u(i,j) \right|}{\|u\|_\infty^{(N)}}. \tag{3.5}$$

Define

$$\nu_N(i,j) = \operatorname{sgn} \frac{\partial H_N^{(k_0,n_0)}}{\partial u(i,j)}(0),$$

where

$$(k_0, n_0) \in \operatorname*{arg\,max}_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \left| \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) \right|.$$

Let $\nu_N(k) = [\nu_N(k,1), \dots, \nu_N(k,N)]^T$ and $\nu_N = [\nu_N(1), \dots, \nu_N(K_N)]$. Then

$$DH_N = \frac{\max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left| \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) \nu_N(i,j) \right|}{\|\nu_N\|_\infty^{(N)}}.$$

By this and (3.5), we have that

$$DH_N = \sup_{u \neq 0} \frac{\max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left| \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) u(i,j) \right|}{\|u\|_\infty^{(N)}}. \tag{3.6}$$

Note that if $u_N = 0$, then by (1.4), (2.9), and (2.11), x_N and z_N are identical, i.e., $\varepsilon_N = 0$. Therefore

$$H_N(0) = 0. \tag{3.7}$$

Hence $H_N^{(k,n)}(0) = 0$. Since $H_N \in \mathcal{C}^1$ locally at 0, by Taylor's theorem, there exists a function $\tilde{H}_N^{(k,n)}$ such that for sufficiently small u,

$$H_N^{(k,n)}(u) = \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0)u(i,j) + \tilde{H}_N^{(k,n)}(u), \tag{3.8}$$

and $\lim_{u\to 0} \frac{\tilde{H}_N^{(k,n)}(u)}{\|u\|_\infty^{(N)}} = 0$. Hence for each $\varepsilon > 0$, there exists δ such that for $\|u\|_\infty^{(N)} < \delta$, $\frac{\left|\tilde{H}_N^{(k,n)}(u)\right|}{\|u\|_\infty^{(N)}} < \varepsilon$. Then for $\|u\|_\infty^{(N)} \le \alpha \le \delta$, $\sup_{\|u\|_\infty^{(N)} \le \alpha} \frac{\left|\tilde{H}_N^{(k,n)}(u)\right|}{\|u\|_\infty^{(N)}} < \varepsilon$. Therefore,

$$\lim_{\alpha\to 0} \sup_{\|u\|_\infty^{(N)} \le \alpha} \frac{\left|\tilde{H}_N^{(k,n)}(u)\right|}{\|u\|_\infty^{(N)}} = 0. \tag{3.9}$$

By (3.8),

$$\|H_N(u)\|_\infty^{(N)} \le \max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left|\tilde{H}_N^{(k,n)}(u)\right| + \max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left| \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0)u(i,j) \right|.$$

Therefore by (2.13),

$$\mu_N \le \lim_{\alpha\to 0} \sup_{\|u\|_\infty^{(N)} \le \alpha} \left(\frac{\max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left|\tilde{H}_N^{(k,n)}(u)\right|}{\|u\|_\infty^{(N)}} + \frac{\max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left|\sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0)u(i,j)\right|}{\|u\|_\infty^{(N)}} \right).$$

Hence by (3.6) and (3.9), we finish the proof. □

Next we present in Lemma 6 a relationship between f_N and DH_N. Define for each N and for $k, l = 1, \dots, K_N$,

$$B_N^{(k,l)} = \begin{cases} A_N(k-1)A_N(k-2)\dots A_N(l), & 1 \le l < k; \\ I_N, & l = k; \\ 0, & l > k, \end{cases} \tag{3.10}$$

where $A_N(l)$ is as defined by (2.16).

LEMMA 6. *Assume that:*

1. *z is continuously differentiable in t;*
2. *for each N, (2.11) holds; and*
3. *for each N, $f_N \in \mathcal{C}^1$.*

Then we have that for each N, for $k, i = 1, \dots, K_N$ and $n, j = 1, \dots, N$,

$$\frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) = B_N^{(k,i)}(n,j) dt_N.$$

PROOF. It follows from (2.11) that $x_N(0) = z_N(0)$. Then by (1.4) and (2.9), $x_N(1) = z_N(1) + dt_N u_N(0)$, and for $k = 2, 3, \dots$,

$$x_N(k) = z_N(k) + \frac{1}{M_N} \sum_{l=1}^{k-1} (f_N(x_N(l)) - f_N(z_N(l))) + dt_N \sum_{l=0}^{k-1} u_N(l).$$

Therefore, by Assumption 3 and by induction, for fixed z_N, x_N is a $\mathcal{C}^1$ function of u_N, because the composition of functions in $\mathcal{C}^1$ is still in $\mathcal{C}^1$. Similarly, by (2.10), for fixed z_N, ε_N is a $\mathcal{C}^1$ function of u_N. Hence Assumption 3 here implies Assumption 3 of Lemma 5.

By Taylor's theorem, for fixed z_N, there exists a function $\tilde{f}_N$ such that

$$f_N(x_N(k)) - f_N(z_N(k)) = Df_N(z_N(k)) \varepsilon_N(k) + \tilde{f}_N(z_N(k) + \varepsilon_N(k), z_N(k)),$$

and for each z,

$$\tilde{f}_N(z, z) = 0, \tag{3.11}$$

and

$$\lim_{\|\varepsilon\|_\infty^{(N)} \to 0} \frac{\left\| \tilde{f}_N(z + \varepsilon, z) \right\|_\infty^{(N)}}{\|\varepsilon\|_\infty^{(N)}} = 0. \tag{3.12}$$

Then we have from (2.10) that for $k = 0, \dots, K_N - 1$,

$$\begin{aligned} \varepsilon_N(k+1) &= \varepsilon_N(k) + dt_N u_N(k) \\ &\quad + \frac{1}{M_N} \left(Df_N(z_N(k)) \varepsilon_N(k) + \tilde{f}_N(z_N(k) + \varepsilon_N(k), z_N(k)) \right). \end{aligned}$$

Therefore

$$\varepsilon_N(k+1) = A_N(k) \varepsilon_N(k) + dt_N u_N(k) + \frac{\tilde{f}_N(z_N(k) + \varepsilon_N(k), z_N(k))}{M_N}.$$

For $k = 0, \dots, K_N - 1$, define

$$\eta_N(k) = dt_N u_N(k) + \frac{\tilde{f}_N(z_N(k) + \varepsilon_N(k), z_N(k))}{M_N}. \tag{3.13}$$

Then $\varepsilon_N(k+1) = A_N(k)\varepsilon_N(k) + \eta_N(k)$. Therefore for $k = 1, \dots, K_N$,

$$\begin{aligned}\varepsilon_N(k) = {} & A_N(k-1)\dots A_N(1)\eta_N(0) + A_N(k-1)\dots A_N(2)\eta_N(1) \\ & + \dots + A_N(k-1)A_N(k-2)\eta_N(k-3) + A_N(k-1)\eta_N(k-2) \\ & + \eta_N(k-1).\end{aligned}$$

Then it follows from (3.10) that for $k = 1, \dots, K_N$,

$$\varepsilon_N(k) = \sum_{l=1}^{k} B_N^{(k,l)} \eta_N(l-1). \tag{3.14}$$

Write $\varepsilon_N(k) = H_N^{(k)}(u_N)$. By (3.13),

$$\eta_N(k) = dt_N u_N(k) + \frac{\tilde{f}_N\left(z_N(k) + H_N^{(k)}(u_N), z_N(k)\right)}{M_N}.$$

Hence by (3.14), for $k = 1, \dots, K_N$,

$$\begin{aligned}& \varepsilon_N(k) \\ & = \sum_{l=1}^{k} B_N^{(k,l)} \left(dt_N u_N(l-1) + \frac{\tilde{f}_N\left(z_N(l-1) + H_N^{(l-1)}(u_N), z_N(l-1)\right)}{M_N} \right).\end{aligned}$$

Denote by $g_N^{(k,l,n)}(\cdot) : \mathbb{R}^{K_N \times N} \to \mathbb{R}^N$ the nth component of

$$B_N^{(k,l)} \tilde{f}_N\left(z_N(l-1) + H_N^{(l-1)}(\cdot), z_N(l-1)\right).$$

By (3.11) and (3.7), $g_N^{(k,l,n)}(0) = 0$.

Let $\{e(i,j) : i = 1, \dots, K_N, j = 1, \dots, N\}$ be the standard basis for $\mathbb{R}^{K_N \times N}$, i.e., each $e(i,j)$ is the element of $\mathbb{R}^{K_N \times N}$ with the (i,j)th entry being 1 and all other entries being 0. Then

$$\frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) = B_N^{(k,i)}(n,j) dt_N + \frac{1}{M_N} \sum_{l=1}^{k} \left(\lim_{h \to 0} \frac{g_N^{(k,l,n)}(h\, e(i,j))}{h} \right).$$

It remains to show that

$$\lim_{h\to 0}\frac{g_N^{(k,l,n)}(h\,e(i,j))}{h}=0.$$

Denote by $\theta_N^{(l,d)}(\cdot):\mathbb{R}^{K_N\times N}\to\mathbb{R}$ the dth component of $\tilde{f}_N(z_N(l)+H_N^{(l)}(\cdot),z_N(l))$. Then

$$g_N^{(k,l,n)}(u)=\sum_{d=1}^{N}B_N^{(k,l)}(n,d)\theta_N^{(l-1,d)}(u).$$

Denote by $\tilde{f}_N^{(l,d)}(\cdot):\mathbb{R}^N\to\mathbb{R}$ the dth component of $\tilde{f}_N(z_N(l)+(\cdot),z_N(l))$. Then

$$\theta_N^{(l,d)}(u)=\tilde{f}_N^{(l,d)}(H_N^{(l)}(u)). \tag{3.15}$$

Then it remains to show that

$$\lim_{\|u\|_\infty^{(N)}\to 0}\frac{\theta_N^{(l,d)}(u)}{\|u\|_\infty^{(N)}}=0. \tag{3.16}$$

Since $H_N\in\mathcal{C}^1$ locally at 0, by (3.7), there exists c such that $|c|<\infty$, and for each $\varepsilon_1>0$, there exists $\delta_1(\varepsilon_1)$ such that for $\|u\|_\infty^{(N)}<\delta_1(\varepsilon_1)$, $\left|\frac{\left\|H_N^{(l)}(u)\right\|_\infty^{(N)}}{\|u\|_\infty^{(N)}}-c\right|<\varepsilon_1$. Hence for $\|u\|_\infty^{(N)}<\delta_1(\varepsilon_1)$,

$$\left\|H_N^{(l)}(u)\right\|_\infty^{(N)}<(|c|+\varepsilon_1)\|u\|_\infty^{(N)}. \tag{3.17}$$

By (3.12), $\lim_{\|x\|_\infty^{(N)}\to 0}\frac{\tilde{f}_N^{(l,d)}(x)}{\|x\|_\infty^{(N)}}=0$. Hence for each $\varepsilon_2>0$, there exists $\delta_2(\varepsilon_2)$ such that for $\|x\|_\infty^{(N)}<\delta_2(\varepsilon_2)$, $\frac{\left|\tilde{f}_N^{(l,d)}(x)\right|}{\|x\|_\infty^{(N)}}<\frac{\varepsilon_2}{|c|+1}$. Hence for $0<\|x\|_\infty^{(N)}<\delta_2(\varepsilon_2)$,

$$\left|\tilde{f}_N^{(l,d)}(x)\right|<\frac{\varepsilon_2}{|c|+1}\|x\|_\infty^{(N)}. \tag{3.18}$$

For each ε, let $\hat{\varepsilon}(\varepsilon)$ be sufficiently small such that

$$(|c|+\hat{\varepsilon}(\varepsilon))\delta_1(\hat{\varepsilon}(\varepsilon))<\delta_2(\varepsilon), \tag{3.19}$$

and

$$\hat{\varepsilon}(\varepsilon)<1. \tag{3.20}$$

Then by (3.17) and (3.19), for $\|u\|_\infty^{(N)} < \delta_1(\hat{\varepsilon}(\varepsilon))$, $\left\|H_N^{(l)}(u)\right\|_\infty^{(N)} < \delta_2(\varepsilon)$. Therefore, in the case that $\left\|H_N^{(l)}(u)\right\|_\infty^{(N)} > 0$, by (3.15) and (3.18),

$$\left|\theta_N^{(l,d)}(u)\right| = \left|\tilde{f}_N^{(l,d)}\left(H_N^{(l)}(u)\right)\right| < \frac{\varepsilon}{|c|+1}\left\|H_N^{(l)}(u)\right\|_\infty^{(N)}.$$

By (3.17) and (3.20),

$$\left\|H_N^{(l)}(u)\right\|_\infty^{(N)} < (|c| + \hat{\varepsilon}(\varepsilon))\|u\|_\infty^{(N)} < (|c|+1)\|u\|_\infty^{(N)}.$$

By the above two inequalities,

$$\frac{\left|\theta_N^{(l,d)}(u)\right|}{\|u\|_\infty^{(N)}} < \varepsilon. \tag{3.21}$$

By (3.11), $\tilde{f}_N^{(l,d)}(0) = 0$. Therefore, in the case that $\left\|H_N^{(l)}(u)\right\|_\infty^{(N)} = 0$, $\theta_N^{(l,d)}(u) = 0$, and thus (3.21) still holds. Therefore, (3.16) holds. □

Now we prove Theorem 2.3 using the two preceding lemmas.

Proof of Theorem 2.3:
By (3.4), Lemma 5, and Lemma 6,

$$\mu_N \leq \max_{\substack{k=1,\dots,K_N\\ n=1,\dots,N}} \sum_{\substack{i=1,\dots,K_N\\ j=1,\dots,N}} \left|B_N^{(k,i)}(n,j)\right| dt_N.$$

Therefore, by (3.10) and the assumption on $\|A_N(k)\|_\infty^{(N)}$, and by the sub-multiplicative property of the induced ∞-norm $\|\cdot\|_\infty^{(N)}$ on $\mathbb{R}^{N\times N}$, there exists $c < \infty$ such that for N sufficiently large,

$$\mu_N \leq K_N\left(\max_{k=1,\dots,K_N-1}\|A_N(k)\|_\infty^{(N)}\right)^{K_N} dt_N \leq K_N dt_N(1 + c\, dt_N)^{K_N}.$$

Since $T < \infty$, by (2.6), $K_N dt_N$ is bounded. As $N \to \infty$, $K_N \to \infty$, and

$$(1 + c\, dt_N)^{K_N} = \left(1 + \frac{cT}{K_N}\right)^{K_N} \to e^{cT}.$$

Therefore $\{\mu_N\}$ is bounded. ∎

3.6. *Proof of Theorem 2.4.* We now prove Theorem 2.4. using Theorem 2.1 and 2.3.

Proof of Theorem 2.4: By (1.5), for fixed N, for $x = [x_1, \dots, x_N]^T \in [0,1]^N$, the (n,m)th component of $Df_N(x)$, where $n, m = 1, \dots, N$, is

$$\begin{cases} P_l(n)x_n(1-x_{n-1}), & m=n-2; \\ (1-x_n)[P_r(n-1)(1-x_{n+1}) - P_l(n+1)x_{n+1}] & \\ \quad +P_l(n)x_n(1-x_{n-2}), & m=n-1; \\ -P_r(n-1)x_{n-1}(1-x_{n+1}) - P_l(n+1)x_{n+1}(1-x_{n-1}) & \\ -P_r(n)(1-x_{n+1})(1-x_{n+2}) - P_l(n)(1-x_{n-1})(1-x_{n-2}), & m=n; \\ (1-x_n)[P_l(n+1)(1-x_{n-1}) - P_r(n-1)x_{n-1}] & \\ \quad +P_r(n)x_n(1-x_{n+2}), & m=n+1; \\ P_r(n)x_n(1-x_{n+1}), & m=n+2; \\ 0 & \text{other wise,} \end{cases}$$

where x_n with $n \le 0$ or $n \ge N+1$ are defined to be zero. Then

$$\begin{aligned} \|A_N(k)\|_\infty^{(N)} &= \max_{n=1,\dots,N} \frac{1}{M_N}(|P_l(n)z_N(k,n)(1-z_N(k,n-1))| \\ &+ |(1-z_N(k,n))[P_r(n-1)(1-z_N(k,n+1)) - P_l(n+1)z_N(k,n+1)] \\ &+ P_l(n)z_N(k,n)(1-z_N(k,n-2))| \\ &+ |M_N - P_r(n-1)z_N(k,n-1)(1-z_N(k,n+1)) \\ &- P_l(n+1)z_N(k,n+1)(1-z_N(k,n-1)) \\ &- P_r(n)(1-z_N(k,n+1))(1-z_N(k,n+2)) \\ &- P_l(n)(1-z_N(k,n-1))(1-z_N(k,n-2))| \\ &+ |(1-z_N(k,n))[P_l(n+1)(1-z_N(k,n-1)) - P_r(n-1)z_N(k,n-1)] \\ &+ P_r(n)z_N(k,n)(1-z_N(k,n+2))| + |P_r(n)z_N(k,n)(1-z_N(k,n+1))|). \end{aligned}$$

Put (2.7), (2.18), (2.19), and the Taylor's expansions

$$z(t, s \pm ds_N) = z(t,s) \pm z_s(t,s)ds_N + z_{ss}(t,s)\frac{ds_N^2}{2} + o(ds_N^2),$$

$$b(s \pm ds_N) = b(s) \pm b_s(s)ds_N + b_{ss}(s)\frac{ds_N^2}{2} + o(ds_N^2),$$

and

$$c(s \pm ds_N) = c(s) \pm c_s(s)ds_N + o(ds_N)$$

into the above equation and rearrange. Then it follows that there exists $c_1 < \infty$ such that for N sufficiently large,

$$\|A_N(k)\|_\infty^{(N)} \le 1 + \max_{n=1,\dots,N} |-c_s(v_N(n)) - b_{ss}(v_N(n))$$

$$
\begin{aligned}
&- 2b(v_N(n))z_{ss}(t_N(k), v_N(n)) + 4b_{ss}(v_N(n))z(t_N(k), v_N(n)) \\
&+ 2b_s(v_N(n))z_s(t_N(k), v_N(n)) + 4c_s(v_N(n))z(t_N(k), v_N(n)) \\
&+ 4c(v_N(n))z_s(t_N(k), v_N(n)) + 6b(v_N(n))z_s(t_N(k), v_N(n))^2 \\
&- 3b_{ss}(v_N(n))z(t_N(k), v_N(n))^2 - 3c_s(v_N(n))z(t_N(k), v_N(n))^2 \\
&+ 6b(v_N(n))z(t_N(k), v_N(n))z_{ss}(t_N(k), v_N(n)) \\
&- 6c(v_N(n))z(t_N(k), v_N(n))z_s(t_N(k), v_N(n))| \frac{ds_N^2}{M_N} + c_1 \frac{ds_N^3}{M_N} \\
:= {} & 1 + \max_{n=1,\dots,N} |q(t_N(k), v_N(n))| \frac{ds_N^2}{M_N} + c_1 \frac{ds_N^3}{M_N}.
\end{aligned}
$$

Then by (2.17), by the assumption of the theorem, and by the fact that b is bounded by $1/2$, there exists $c_3 < \infty$ such that for N sufficiently large, $\|A_N(k)\|_\infty^{(N)} \leq 1 + c_3\, dt_N$. Hence the last assumption of Theorem 2.3 holds. Then one can verify that the assumptions of Theorem 2.1 hold.

One can show from (1.5) and (2.20) (after some tedious algebra) that there exists $c < \infty$ such that for N sufficiently large,

$$\gamma_N \leq c\, ds_N. \tag{3.22}$$

By Theorem 2.1, we finish the proof. ∎

4. Conclusion. In this paper we analyze the convergence of a sequence of Markov chains to its continuum limit, the solution of a PDE, in a two-step procedure. We provide precise sufficient conditions for the convergence and the explicit rate of convergence. Based on such convergence we approximate the Markov chain modeling a large wireless sensor network by a nonlinear diffusion-convection PDE.

With the well-developed mathematical tools available for PDEs, this approach provides a framework to model and simulate networks with a very large number of components, which is practically infeasible for Monte Carlo simulation. Such a tool enables us to tackle problems such as performance analysis and prototyping, resource provisioning, network design, network parametric optimization, network control, network tomography, and inverse problems, for very large networks. For example, we can now use the PDE model to optimize certain performance metrics (e.g., throughput) of a large network by adjusting the placement of destination nodes or the routing parameters (e.g., coefficients in convection terms), with relatively negligible computation overhead compared with that of the same task done by Monte Carlo simulations.

For simplicity, we have treated sequences of grid points that are uniformly located. As with finite difference methods for differential equations, the convergence results can be extended to models that have nonuniform points spacing under assumptions that insure the points in the sequence should become dense in the underlying domain uniformly in the limit. For example, we could consider a double sequence of minimum point spacing $\{h_i\}$ and maximum point spacing $\{H_i\}$ with $H_i/h_i =$ constant, and for each i, we can consider a model with nonhomogeneous point spacing between h_i and H_i. We can also introduce a spatial change of variables that maps a nonuniform model to a uniform model. This changes the coefficients in the resulting PDE, by substitution and the chain rule. In this way we can extend our approach to nonuniform, even mobile, networks. We can further consider the control of nodes such that global characteristics of the network are invariant under node locations and mobility. (See our paper [60] for details.)

The assumption made in (2.18) that the probabilities of transmission behave continuously insures that there is a limiting behavior in the limit of large numbers of nodes and relates the behavior of networks with different numbers of nodes. The convergence results can be extended to the situation in which the probabilities change discontinuously at a finite number of lower dimensional linear manifolds (e.g., points in one dimension, lines in two dimensions, planes in three dimensions) in space provided that all of the discrete networks under consideration have nodes on the manifolds of discontinuity.

There are other considerations regarding the network that can significantly affect the derivation of the continuum model. For example, transmissions could happen beyond immediate nodes, and the interference between nodes could behave differently in the presence of power control; we can consider more boundary conditions other than sinks, including walls, semi-permeating walls, and their composition; and we can seek to establish continuum models for other domains such as the Internet, cellular networks, traffic networks, and human crowds.

References.

[1] Altman, E., Silva, A., Bernhard, P. and Debbah, M. (2007). Continuum equilibria for routing in dense ad-hoc networks. In *Proc. of Forty-Fifth Annual Allerton Conference on Communication, Control, and Computing*.

[2] Bahadoran, C., Guiol, H., Ravishankar, K. and Saada, E. (2006). Euler hydrodynamics of one-dimensional attractive particle systems. *Ann. Probab.* **34** 1339-1369.

[3] Benaïm, M. (1999). Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités XXXIII. Lecture Notes in Mathematics* **1709** 1-68. Springer Berlin/Heidelberg.

[4] Benaïm, M. and Boudec, J.-Y. L. (2008). A class of mean field interaction models for computer and communication systems. *Perform. Eval* 11-12.

[5] Billingsley, P. (1995). *Probability and Measure*. New York, NY: Wiley.

[6] Billingsley, P. (1999). *Convergence of Probability Measures*. New York, NY: Wiley.

[7] Blount, D. (1992). Law of large numbers in the supremum norm for a chemical reaction with diffusion. *Ann. Appl. Probab.* **2** 131-141.

[8] Bramson, M., Dai, J. G. and Harrison, J. M. (2010). Positive recurrence of reflecting Brownian motion in three dimensions. *Ann. Appl. Probab.* **20**.

[9] Budhiraja, A. and Ghosh, A. P. (2006). Diffusion approximations for controlled stochastic networks: An asymptotic bound for the value function. *Ann. Appl. Probab.* **16** 1962-2006.

[10] Burch, N., Chong, E., Estep, D. and Hannig, J. (2011). Analysis of Routing Protocols and Interference-limited Communication in Large Wireless Networks. *Journal of Engineering Mathematics*. in revision.

[11] Burger, M., Markowich, P. and Pietschmann, J. (2011). Continuous limit of a crowd motion and herding model: Analysis and numerical simulations. *Kinetic and Related Methods* **4** 1025-1047.

[12] Caines, P. E. (2009). Bode lecture: Mean Field stochastic control. In *Proceedings of the 48th IEEE Conference on Decision and Control, Shanghai, China*.

[13] Chen, H. (1995). Fluid Approximations and Stability of Multiclass Queueing Networks: Work-Conserving Disciplines. *Ann. Appl. Probab.* **5** 637-665.

[14] Chen, H. and Shanthikumar, J. (1994). Fluid limits and diffusion approximations for networks of multi-server queues in heavy traffic. *Discrete Event Dynamic Systems* **4** 269-291.

[15] Chen, W., Huang, D., Kulkarni, A. A., Unnikrishnan, J., Zhu, Q., Mehta, P., Meyn, S. and Wierman, A. (2009). Approximate dynamic programming using fluid and diffusion approximations with applications to power management. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on* 3575-3580.

[16] Chong, E. K. P., Estep, D. and Hannig, J. (2008). Continuum modeling of large networks. *Int. J. Numer. Model.* **21** 169-186.

[17] Dai, J. G. (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Ann. Appl. Probab.* **5** 49-77.

[18] Darling, R. W. R. (2002). Fluid limits of pure jump Markov processes: a practical guide. *arXiv:math/0210109v3*.

[19] Darling, R. W. R. and Norris, J. R. (2008). Differential equation approximations for Markov chains. *Probability Surveys* **5** 37.

[20] Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons Inc.

[21] Graham, C. (2000). Chaoticity on path space for a queueing network with selection of the shortest queue among several. *Ann. Appl. Probab.* **37** 198-211.

[22] Gromoll, H. C. and Williams, R. J. (2009). Fluid limits for networks with bandwidth sharing and general document size distributions. *Ann. Appl. Probab.* **19** 243-280.

[23] Grossglauser, M. and Tse, D. N. C. (2002). Mobility increases the capacity of ad hoc wireless networks. *Networking, IEEE/ACM Transactions on* **10** 477-486.

[24] Guizani, M., Rayes, A., Khan, B. and Al-Fuqaha, A. (2010). *Network Modeling and Simulation: A Practical Perspective*. John Wiley & Sons.

[25] Guo, M. Z., Papanicolaou, G. C. and Varadhan, S. R. S. (1988). Nonlinear

diffusion limit for a system with nearest neighbor interactions. *Communications in Mathematical Physics* **118** 31-59.

[26] Gupta, P. and Kumar, P. R. (2000). The capacity of wireless networks. *Information Theory, IEEE Transactions on* **46** 388-404.

[27] Harrison, J. M. (1973). A Limit Theorem for Priority Queues in Heavy Traffic. *Journal of Applied Probability* **10** 907-912.

[28] Harrison, J. M. and Reiman, M. I. (1981). Reflected Brownian Motion on an Orthant. *Ann. Probab.* **9** 302-308.

[29] Harrison, J. M. and Williams, R. J. (1992). Brownian models of feedforward queueing networks: quasireversibility and product form solutions. *Ann. Appl. Probab.* **2** 263-293.

[30] Herdtner, J. D. and Chong, E. K. P. (2005). Throughput-storage tradeoff in ad hoc networks. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE* **4** 2536-2542.

[31] Huang, C. F. and Pages, H. (1992). Optimal consumption and portfolio policies with an infinite horizon: existence and convergence. *Ann. Appl. Probab.* **2** 36-64.

[32] Iglehart, D. L. (1971). Multiple channel queues in heavy traffic. *Probability Theory and Related Fields* **17** 168-180.

[33] Isaacson, E. and Keller, H. B. (1966). *Analysis of Numerical Methods*. Wiley.

[34] Jain, M. K. (1984). *Numerical Solution of Differential Equations*. Wiley Eastern Halsted Press.

[35] Kelly, F. P. and Williams, R. J. (2004). Fluid model for a network operating under a fair bandwidth-sharing policy. *Ann. Appl. Probab.* **14** 1055-1083.

[36] Kiefer, J. and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Ann. Math. Statis* **23** 462-466.

[37] Koukkous, A. (1999). Hydrodynamic behavior of symmetric zero-range processes with random rates. *Stochastic Processes and their Applications* **84** 297-312.

[38] Kurtz, T. G. (1970). Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes. *Journal of Applied Probability* **7** 49-58.

[39] Kushner, H. J. (1984). *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*. Cambridge, MA: MIT Press.

[40] Lai, T. L. (2003). Stochastic approximation: invited paper. *Ann. Statist.* **31** 391-406.

[41] Lee, J., Bohacek, S., Hespanha, J. P. and Obraczka, K. (2007). Modeling Communication Networks With Hybrid Systems. *Networking, IEEE/ACM Transactions on* **15** 630-643.

[42] Lelarge, M. (2005). Fluid limit of generalized Jackson queueing networks with stationary and ergodic arrivals and service times. *J. Appl. Probab.* **42** 491-512.

[43] Liu, G. R. and Quek, S. S. (2003). *The Finite Element Method: A Practical Course*. Butterworth-Heinemann.

[44] Liu, L., Parag, P., Tang, J., Chen, W. Y. and Chamberland, J. F. (2007). Resource allocation and quality of service evaluation for wireless communication systems using fluid models. *Information Theory, IEEE Transactions on* **53** 1767-1777.

[45] McVinish, R. and Pollett, P. K. The deterministic limit of heterogeneous density dependent Markov chains. submitted to *Ann. Appl. Probab.*

[46] Meyn, S. P. (1995). Transience of Multiclass Queueing Networks Via Fluid Limit Models. *Ann. Appl. Probab.* **5** 946-957.

[47] Misra, V., Gong, W. B. and Towsley, D. (2000). Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Com-*

puter Communication. SIGCOMM '00 151-160. ACM, New York, NY, USA.

[48] Mitchell, A. R. and Griffiths, D. F. (1980). *The Finite Difference Method in Partial Differential Equations*. Wiley.

[49] Neely, M. J., Modiano, E. and Rohrs, C. E. (2005). Dynamic power allocation and routing for time-varying wireless networks. *Selected Areas in Communications, IEEE Journal on* **23** 89-103.

[50] Nguyen, V. (1995). Fluid and Diffusion Approximations of a Two-Station Mixed Queueing Network. *Mathematics of Operations Research* **20** 321-354.

[51] Paschalidis, I. C. and Chen, Y. (2008). Anomaly detection in sensor networks based on large deviations of Markov chain models. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on* 2338-2343.

[52] Paschalidis, I. C. and Smaragdakis, G. (2009). Spatio-Temporal Network Anomaly Detection by Assessing Deviations of Empirical Measures. *Networking, IEEE/ACM Transactions on* **17** 685-697.

[53] Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Statis* **22** 400-407.

[54] Sirkeci-Mergen, B., Scaglione, A. and Mergen, G. (2006). Asymptotic analysis of multistage cooperative broadcast in wireless networks. *Information Theory, IEEE Transactions on* **52** 2531-2550.

[55] Toumpis, S. and Tassiulas, L. (2006). Optimal deployment of large wireless sensor networks. *Information Theory, IEEE Transactions on* **52** 2935-2953.

[56] Weiss, A. (1995). An introduction to large deviations for communication networks. *Selected Areas in Communications, IEEE Journal on* **13** 938-952.

[57] Whitt, W. (1982). On the Heavy-Traffic Limit Theorem for $GI/G/\infty$ Queues. *Advances in Applied Probability* **14** 171-190.

[58] Wong, K. D., Gin, C. W., Rajeswari, R., Woon, W. L. and Chong, E. K. P. (2007). Feasability of a PDE-based teletraffic model for cellular networks. In *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on* 383 -388.

[59] Xu, G. L. and Shreve, S. E. (1992). A duality method for optimal consumption and investment under short-selling prohibition. II. constant market coefficients. *Ann. Appl. Probab.* **2** 314-328.

[60] Zhang, Y., Chong, E. K. P., Hannig, J. and Estep, D. (2011). Continuum modeling and control of large nonuniform networks. In *the 49th annual Allerton conference on Communication, control, and computing* 1670-1677.

Yang Zhang
Edwin K. P. Chong
Department of Electrical
and Computer Engineering
Colorado State University
Fort Collins, CO 80523-1373
E-mail: yang.zhang@colostate.edu
edwin.chong@colostate.edu

Jan Hannig
Department of Statistics
and Operation Research
The University of North Carolina
at Chapel Hill
Chapel Hill, NC 27599-3260
E-mail: jan.hannig@unc.edu

Donald Estep
Department of Statistics
Colorado State University
Fort Collins, CO 80523
E-mail: estep@math.colostate.edu