Forty-Ninth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 28 - 30, 2011

Asymptotics of the Invariant Measure in Mean Field Models with Jumps

Vivek Shripad Borkar
Indian Institute of Technology Bombay, India

Rajesh Sundaresan
Indian Institute of Science, Bangalore, India

***Abstract*—We consider the asymptotics of the invariant measure for the process of spatial distribution of N coupled Markov chains in the limit of a large number of chains. Each chain reflects the stochastic evolution of one particle. The chains are coupled through the dependence of transition rates on the spatial distribution of particles in the various states. Our model is a caricature for medium access interactions in wireless local area networks. Our model is also applicable in the study of spread of epidemics in a network. The limiting process satisfies a deterministic ordinary differential equation called the McKean-Vlasov equation. When this differential equation has a unique globally asymptotically stable equilibrium, the spatial distribution converges weakly to this equilibrium. Using a control-theoretic approach, we examine the question of a large deviation from this equilibrium.**

***Index Terms*—decoupling approximation, fluid limit, invariant measure, McKean-Vlasov equation, mean field limit, small noise limit, stationary measure, stochastic Liouville equation**

I. THE MODEL

This paper expands on the talk given at the 2011 Allerton Conference on Communications, Control and Computing. The presentation here will be somewhat expository and informal. Readers are referred to [9] for a more formal and detailed presentation of the results.

We begin with a description of the system under study.

System Description: There are N particles (nodes) in our caricature of a wireless local area network (WLAN). At each instant of time, a particle's state is a particular value taken from a finite state space, say $\mathcal{Z} = \{0, 1, \cdots, r-1\}$. This state represents the number of failed attempts at transmission of the head-of-the-line packet at that particle's queue. When a particle is in state i, a successful transmission gets the packet out of the system, and the particle moves to state 0 to service the next packet. A failed transmission moves the particle to state $i+1$ (mod r). In the case when i was initially $r-1$, i.e., $r-1$ unsuccessful transmission attempts were already made, another failed attempt results in the discarding of the packet. The particle then moves to state 0 with the next packet readied for transmission. We may interpret r as the maximum number of transmission attempts. The transition rate for a particle from state i to state j is governed by *mean field dynamics*, i.e., the transition rate is $\lambda_{i,j}(\mu_N(t))$ where $\mu_N(t)$ is the empirical distribution of the states of particles at time t. If $X_n^{(N)}(t)$ is the state of the nth particle at time t, then one may write $\mu_N(t)$ as

$$\mu_N(t) = \frac{1}{N}\sum_{n=1}^{N} \delta_{\{X_n^{(N)}(t)\}}.$$

The particles interact only through the dependence of their transition rates on the current empirical measure $\mu_N(t)$.

It must be noted that the above system model does not perfectly capture all aspects in a WLAN. In particular, interactions and changes of states occur in discrete-time units of slots, and multiple nodes may transit in one slot. Multiple transitions never occur, almost surely, in our continuous-time model. Nevertheless, the discrete-time model's transition rates and the slot sizes are appropriately scaled down as N grows so that the transition rates approach constants, and our continuous time model provides accurate predictions of behavior on the discrete-time model. See [6] for a similar continuous-time model.

The transitions allowed in the above model are i to either $i+1$ (mod r) or 0. Let us say that $\mathcal{E}$ denotes the set of pairs of allowed transitions. In the above model,

$$\mathcal{E} = \{(i, i+1), i = 0, 1, \cdots, r-1\} \cup \{(i, 0), i = 0, 1, \cdots, r-1\}$$

where the addition is taken modulo r.

The process $X^{(N)}(\cdot) = \{X_n^{(N)}(\cdot), 1 \le n \le N\}$ is clearly a Markov process. But one difficulty needs to be surmounted in analyzing this system: the size of the state space grows exponentially in the number of particles. As a step towards addressing this difficulty, we first consider a mean field limit as the number of particles grows to infinity.

II. THE MEAN FIELD LIMIT

Let us denote by $\mathcal{M}_1(\cdots)$, the space of probability measures on $\cdots$ with an associated σ-algebra that will usually be clear from the context. A moment's thought suggests that the empirical process $\mu_N(\cdot)$ is Markov, and that its infinitesimal generator is

$$\begin{aligned}
&\Omega^{(N)}\Phi(\xi) \\
&= \lim_{h\downarrow 0} \frac{1}{h}\mathbb{E}\left[\Phi(\mu_N(t+h)) - \Phi(\mu_N(t)) \mid \mu_N(t) = \xi\right] \\
&= \sum_{(i,j):j\neq i} N\xi(i)\lambda_{i,j}(\xi)\left[\Phi\left(\xi + \frac{1}{N}e_j - \frac{1}{N}e_i\right) - \Phi(\xi)\right]
\end{aligned}$$

where $\Phi : \mathcal{M}_1(\mathcal{Z}) \to \mathbb{R}$ is any bounded continuous function. Let us take $\Phi(\xi) = \xi(k)$, run k through the elements in $\mathcal{Z}$, and we will be able to verify the following ordinary differential equation: the expected drift in $\mu_N(t)$ satisfies

$$\lim_{h\downarrow 0} \frac{1}{h}\mathbb{E}\left[\mu_N(t+h) - \mu_N(t) \mid \mu_N(t) = \xi\right] = A_\xi^* \xi$$

978-1-4577-1818-2/11/$26.00 ©2011 IEEE

where A_ξ is the usual rate matrix describing the transitions for a tagged particle, i.e.,

$$(A_\xi)_{i,j} = \begin{cases} \lambda_{i,j}(\xi) & \text{if } i \neq j \\ -\sum_{j':j'\neq i} \lambda_{i,j'}(\xi) & \text{if } i = j. \end{cases}$$

The initial condition is taken to be $\mu_N(0)$. The lack of dependence N on the right-hand side above is due to the linear nature of the chosen Φ. If there were no dependence on N for all bounded continuous Φ, then the evolution would be deterministic. For general Φ, a Taylor's series expansion of Φ near ξ indicates that the dependence on N is from the second order onwards. This suggests a first order approximation: just ignore the randomness and set its evolution to be that of its mean, the deterministic $\mu(t)$, which is the solution to the ordinary differential equation (ODE)

$$\dot{\mu}(t) = A^*_{\mu(t)}\mu(t)$$

with an initial condition $\mu(0) = \mu_N(0)$. This ODE is called the *McKean-Vlasov equation*. One also anticipates that this approximation is good in the large number of particles limit, i.e., $\lim_{N\to+\infty} \mu_N(t) = \mu(t)$, where the limit is interpreted in an appropriate sense, with $\mu(0) = \lim_{N\to+\infty} \mu_N(0)$. See the third remark following Theorem IV.1.

The question of existence and uniqueness of the Markov process defined by the generator above is answered by assuming that the rates are bounded from above. The question of well-posedness of the ODE above, i.e., does the solution to the ODE exist for any initial condition and is it unique, is answered by assuming that the function $\lambda_{i,j}(\cdot)$ is Lipschitz. In addition, we make a uniform lower bound assumption on the rates of allowed transmissions. Let us summarize these assumptions as

(A) There exist positive constants $c > 0$ and $C < +\infty$ such that, for every $(i,j) \in \mathcal{E}$, the bounds $c \leq \lambda_{i,j}(\cdot) \leq C$ hold, and moreover, the mapping $\mu \mapsto \lambda_{i,j}(\mu)$ is Lipschitz continuous over $\mathcal{M}_1(\mathcal{Z})$.

III. Propagation of chaos and the decoupling approximation

The mean field approximation helps address the state-space explosion issue alluded to above. Let us focus attention on a finite number of tagged particles. If the initial conditions $\{X_n^{(N)}(0), n = 1, \cdots, N\}$ are independent and identically distributed across particles, and remain so for the tagged particles in the limit as $N \to +\infty$, a condition that we shall call *chaotic*, then the evolution of the k tagged particles (in the infinite particle limit, over any fixed and finite time duration) is independent and identically distributed. Further, the tagged particles' evolutions are characterized by the transition rates $\lambda_{i,j}(\mu(t))$, where $\mu(\cdot)$ is the solution to the ODE over the finite duration of interest, the McKean-Vlasov limit. The initial chaos thus propagates. The effect of all the other particles is to create the mean field. This idea was introduced Kac [20] as a simple model in kinetic theory and was later studied by McKean and others [22], with further developments in [29] and [19].

Under some conditions, even though the initial condition may not be chaotic, the system ends up being chaotic in the large time limit, so that the particles' evolutions are asymptotically independent, in a manner of speaking. The order of the limits $t \to +\infty$ and $N \to +\infty$ will be made precise in Section V. In the communication networks literature, this is termed as the *decoupling approximation* and was recently popularized by Bianchi [4]; some early precursors in this direction were [1], [2], [28]. See [8], [7], [3], and [23] for further justifications and reinterpretations of the decoupling approximation.

As remarked in [15], experimental evidence for CSMA protocols indicates that the model and the decoupling approximation are surprisingly accurate even for small populations. Indeed, this is one of the main reasons for the model's enormous popularity. One can explain this good fit, to some extent, if there is exponentially fast convergence to the mean field. Let us probe this notion a little further via the theory of large deviations.

IV. Large deviations from the McKean-Vlasov limit for finite time durations

Large deviation principles for interacting diffusions and interacting jump Markov processes have been well-studied by several authors, e.g., [11] for diffusions, [21], [14], [17], and [10] for processes with jumps. The initial condition is either a particular fixed state for each particle or is a chaotic measure (independent and identically distributed across particles) in [21], [17], and [10]. The initial conditions are more relaxed in [11] (for diffusions) and [14] (for a restricted class of jump processes that does not include our set up): they only need $\lim_{N\to+\infty} \mu_N(0) = \mu(0) = \nu$, say. The large deviation principle is quantified by a rate function $S_{[0,T]}(\mu|\nu)$ that measures the difficulty of passage of the empirical process $(\mu_N(t), t \in [0,T])$ for some $T \in (0, +\infty))$ in the neighborhood of a path $\mu : [0,T] \to \mathcal{M}_1(\mathcal{Z})$ having initial condition $\mu(0) = \nu$.

Some preliminaries on the large deviation principle are in order. Let us denote by ν_N the initial empirical measure, i.e., the initial value $\mu_N(0) = \nu_N$. Let $p_{\nu_N}^{(N)}$ denote the law of the Markov empirical process $\mu_N : [0,T] \to \mathcal{M}_1(\mathcal{Z})$ (whose generator we wrote earlier as $\Omega^{(N)}$) under the initial condition ν_N. Let us equip the space $D([0,T], \mathcal{M}_1(\mathcal{Z}))$, the space of cadlag functions on $[0,T]$ taking values in $\mathcal{M}_1(\mathcal{Z})$, with the metric

$$\rho_T(\xi, \xi') = \sup_{t\in[0,T]} \rho_0(\xi(t), \xi'(t))$$

where ρ_0 may be taken as the metric that metrizes the topology of weak convergence on $\mathcal{M}_1(\mathcal{Z})$. Two observations are that this topology is finer than the usual Skorohod topology (with the complete metric) and that the resulting topological space under the metric ρ_T is not separable.

Let $(p^{(N)}, N \geq 1)$ be a sequence of probability measures. We say that $(p^{(N)}, N \geq 1)$ satisfies the large deviation principle with speed N and a good rate function $S_{[0,T]}(\mu)$ if the following hold.

(i) For each open set G of the topological space $D([0,T],\mathcal{M}_1(\mathcal{Z}))$, we have

$$\liminf_{N\to+\infty} N^{-1}\log p^{(N)}(G) \geq -\inf_{\mu\in G} S_{[0,T]}(\mu).$$

(ii) For each closed set F of the topological space $D([0,T],\mathcal{M}_1(\mathcal{Z}))$, we have

$$\liminf_{N\to+\infty} N^{-1}\log p^{(N)}(F) \leq -\inf_{\mu\in F} S_{[0,T]}(\mu).$$

(iii) For each $a\in[0,+\infty)$, the level sets

$$\{\mu\in D([0,T],\mathcal{M}_1(\mathcal{Z}))\mid S_{[0,T]}(\mu)\leq a\}$$

are compact.

Our interest is in the large deviation principle for a sequence $(p^{(N)}_{\nu_N}, N\geq 1)$ whose initial conditions $\nu_N\to\nu$ weakly, and the rate functions will depend on the limiting initial condition ν. To highlight this dependence, we denote the rate function as $S_{[0,T]}(\mu|\nu)$. Note that μ is a measure-valued ($\mathcal{M}_1(\mathcal{Z})$-valued) path on $[0,T]$ while $\nu\in\mathcal{M}_1(\mathcal{Z})$ is simply a measure that denotes the initial condition for the path.

We shall now identify the rate function $S_{[0,T]}(\mu|\nu)$ for this sequence of empirical processes. Let $\xi\in\mathcal{M}_1(\mathcal{Z})$ and let $f:\mathcal{Z}\to\mathbb{R}$. The *Hamiltonian* $\mathscr{H}^{(N)}$ associated to the infinitesimal generator $\Omega(N)$ is defined as

$$\begin{aligned}\mathscr{H}^{(N)}(\xi,f) &= e^{-\langle\xi,f\rangle}\Omega^{(N)}e^{\langle\xi,f\rangle}\\ &= N\sum_{i,j:j\neq i}\xi(i)\lambda_{i,j}(\xi)\left[e^{(f(j)-f(i))/N}-1\right].\end{aligned}$$

where $\langle\xi,f\rangle:=\sum_i\xi(i)f(i)$. The *scaled Hamiltonian* is

$$\begin{aligned}\mathscr{H}(\xi,f) &= \lim_{N\to+\infty}\frac{1}{N}\mathscr{H}^{(N)}(\mu,Nf)\\ &= \sum_{i,j:j\neq i}\xi(i)\lambda_{i,j}(\xi)\left[e^{(f(j)-f(i))}-1\right].\end{aligned}$$

For a $\theta:\mathcal{Z}\to\mathbb{R}$, define the Legendre transform of the scaled Hamiltonian in the usual way as

$$\mathscr{L}(\xi,\theta)=\sup_{f:\mathcal{Z}\to\mathbb{R}}\{\langle\theta,f\rangle-\mathscr{H}(\xi,f)\}.$$

We will also need the subset $\mathcal{A}\subset D([0,T],\mathcal{M}_1(\mathcal{Z}))$ of *absolutely continuous* measure-valued functions on $[0,T]$. We refer the reader to [21, Defn. 3.1] for a precise definition of $\mathcal{A}$ and some consequences. We state two consequences that give an idea of the regularity of paths in $\mathcal{A}$. First, if $\mu\in\mathcal{A}$, then for any $f:\mathcal{Z}\to\mathbb{R}$, the function $\langle f,\mu(\cdot)\rangle$ is an absolutely continuous real-valued function. Second, if $\mu\in\mathcal{A}$ then the time derivative $\dot{\mu}$ exists for almost all $t\in[0,T]$.

Finally, let us define

$$\tau^*(u)=\begin{cases}(u+1)\log(u+1)-u & \text{if } u>-1\\ 1 & \text{if } u=-1\\ +\infty & \text{if } u<-1.\end{cases}$$

This is the Legendre dual of the function $\tau(u)=e^u-u-1$. With these elaborate preliminaries, the following result can be shown.

Theorem IV.1. *Let Assumption (***A***) hold. Suppose that the initial conditions $\nu_N\to\nu$ weakly. Then the sequence $(p^{(N)}_{\nu_N}, N\geq 1)$ satisfies the large deviation principle in the space $D([0,T],\mathcal{M}_1(\mathcal{Z}))$ (under the topology induced by the metric ρ_T) with speed N and good rate function $S_{[0,T]}(\mu|\nu)$ given by*

$$S_{[0,T]}(\mu|\nu)=\begin{cases}\int_{[0,T]}\mathscr{L}(\mu(t),\dot{\mu}(t))\,dt & \text{if } \mu\in\mathcal{A}\\ +\infty & \text{otherwise.}\end{cases}$$

If a path $\mu\in D([0,T],\mathcal{M}_1(\mathcal{Z}))$ has $S_{[0,T]}(\mu|\nu)<+\infty$, then $\mu\in\mathcal{A}$ and there exist rates $(l_{i,j}(t),t\in[0,T],\ (i,j)\in\mathcal{E})$ such that

- *$\dot{\mu}(t)=L(t)^*\mu(t)$ where $L(t)$ is the rate matrix associated with the time-varying rates $(l_{i,j}(t),(i,j)\in\mathcal{E})$ and $L(t)^*$ is its adjoint;*
- *the good rate function $S_{[0,T]}(\mu|\nu)$ is given by*

$$S_{[0,T]}(\mu|\nu)=\int_{[0,T]}\Big[\sum_{(i,j)\in\mathcal{E}}(\mu(t)(i))\lambda_{i,j}(\mu(t))\tau^*\left(\frac{l_{i,j}(t)}{\lambda_{i,j}(\mu(t))}-1\right)\Big]dt. \quad (1)$$

Remarks: 1. This result is a mild generalization of [21, Th. 3.1], and of the results in [10], [17], and [13] that assume chaotic initial conditions. We allow any arbitrary sequence of initial conditions ν_N so long as $\nu_N\to\nu$ weakly.

2. For any initial measure ν, it can be shown that the *cost* $S_{[0,T]}(\mu|\nu)$ of the limiting McKean-Vlasov path with the initial condition ν is 0. This is the path $\mu:[0,T]\to\mathcal{M}_1(\mathcal{Z})$ that satisfies the ODE

$$\dot{\mu}(t)=A^*_{\mu(t)}\mu(t)$$

with initial condition $\mu(0)=\nu$.

3. Since $\tau^*(u)=0$ if and only if $u=0$, we can conclude from (1) that if the rate function $S_{[0,T]}(\mu|\nu)=0$, then μ must be the unique solution to the McKean-Vlasov equation $\dot{\mu}(t)=A^*_{\mu(t)}\mu(t)$ with initial condition $\mu(0)=\nu$. (That the solution is unique follows from Lipschitz assumption in Assumption (**A**) which implies the well-posedness of the ODE). The limit law $p^{(N)}_{\nu_N}\to\delta_{\mu(\cdot)}$ in the topology of weak convergence automatically follows.

4. When a path μ is such that $S_{[0,T]}(\mu|\nu)<+\infty$, the first bullet in the second statement of Theorem IV.1 says that there is a control (tilt), given by the rate matrix $L(t)$, such that the normal limiting trajectory under this control is $\mu(\cdot)$. $S_{[0,T]}(\mu|\nu)$ is then the cost of this control.

V. Behavior for large time

The mean field approximation for finite durations is summarized as follows. Given an initial condition, and any finite time duration, the empirical measure follows the McKean-Vlasov dynamics. A finite number of tagged particles evolve approximately independently (in the large particle limit) in the mean field environment.

What happens as $t\to+\infty$?

Let us assume that the allowed transitions in $\mathcal{E}$ are such that the directed graph with vertices $\mathcal{Z}$ and edges $\mathcal{E}$ is irreducible. The corresponding continuous time Markov chain of empirical measures $\mu_N(\cdot)$ is then irreducible for each $N \in \mathbb{N}$. Its state space grows with N but has size at most $(N+1)^r$. It follows that for each finite N, there is a unique invariant measure, say $\wp^{(N)}$, to which the law of $\mu^{(N)}(t)$ converges as $t \to +\infty$. What can we say about $\lim_{N\to+\infty} \wp^{(N)}$?

Let us consider the good case first. Consider the McKean-Vlasov equation

$$\dot{\mu}(t) = A^*_{\mu(t)}\mu(t).$$

An *equilibrium* point is a $\xi \in \mathcal{M}_1(\mathcal{Z})$ that satisfies $A^*_\xi \xi = 0$. If $\mu(0) = \xi$, then $\mu(t) = \xi$ for all $t > 0$. Such a point is said to be *globally asymptotically stable* if the following hold:

- Regardless of the initial condition, we have $\lim_{t\to+\infty} \mu(t) = \xi_0$. The singleton set $\{\xi_0\}$ is therefore an *attractor*, and *all* trajectories converge to it.
- The singleton set $\{\xi_0\}$ is *Liapunov stable*, i.e., for every $\varepsilon > 0$, there is a $\delta > 0$ such that every trajectory initiated in the δ-neighborhood of ξ_0 remains in its ε-neighborhood.

A unique equilibrium that is globally asymptotically stable, say ξ_0, is a good thing to have. One anticipates in this case that $\wp^{(N)} \to \delta_{\xi_0}$ weakly.

In general, however, there can be many equilibria. Or there could even be a unique equilibrium, but it is not stable, and the system settles at a different limit set, say a limit cycle. Even worse, there could be multiple limiting sets, connected compact sets that are invariant to the McKean-Vlasov dynamics, with different basins of attraction. Such situations are not uncommon; see [3]. In such cases, we need to understand where $\lim_{N\to+\infty} \wp^{(N)}$ settles.

A large deviation principle, if we can find one, settles this question. Suppose that $(\wp^{(N)}, N \geq 1)$ satisfies the large deviation principle with a good rate function $s : \mathcal{M}_1(\mathcal{Z}) \to [0, +\infty]$, then the likely set to which $\wp^{(N)}$ will converge is

$$\arg \min_{\xi \in \mathcal{M}_1(\mathcal{Z})} s(\xi) = s^{-1}\{0\}.$$

It of course tells a lot more about likelihood of deviations from this set and difficulty of being in far off neighborhoods. We shall however focus only on the case when there is a unique equilibrium that is globally asympotically stable.

VI. A CONTROL THEORETIC VIEW

We now describe our approach to obtaining a large deviation principle for $(\wp^{(N)}, N \geq 1)$. Our work is naturally related to small noise asymptotics of invariant measures of diffusions and of certain jump processes, which were considered by Freidlin and Wentzell [18] and Shwartz and Weiss [27]. For jump processes, they provide results under rather strong assumptions on jump rates (for e.g., logarithm of the empirical process's jump rates is bounded). These are not satisfied by the Markov empirical process considered here. Since the state-space is constrained to lie within the probability simplex, there is a significant push inward at the boundary of the simplex, and the logarithm of the transition rates cannot be bounded. We therefore need to do a little more work.

Our approach for solving the large deviations of invariant measure exploits a control-theoretic view described in [5], which considers a class of diffusions and generalizes results of [26] and [12] on small noise asymptotics of invariant measures and exit probabilities in diffusions. See [25] for small noise asymptotics of exit probabilities in processes with jumps. The control-theoretic approach differs from those of Freidlin and Wentzell [18] and Shwartz and Weiss [27], which rely on a result of Khasminskii (see, e.g., [18, Ch. 6.4]) for the description of invariant measures of Markov chains. Our control-theoretic approach works when there is a unique globally asymptotically stable equilibrium for the McKean-Vlasov dynamics, but appears to require the Khasminskii result for resolution of certain boundary conditions in the case when there is more than one equilibrium. We do not address here the case of multiple equilibria or attractors.

1) *Nonchaotic initial conditions*: As in Freidlin-Wentzell [18], we will extract a large deviation principle for the invariant measure from a large deviation principle over finite time durations. Let us assume that the random variables $\{X_n^{(N)}(0), 1 \leq n \leq N\}$ that describe the state of the N particles are exchangeable, and that $\mu_N(0)$ is started at its invariant measure $\wp^{(N)}$, which may not be degenerate. The initial law of particles' states, with the invariant measure $\wp^{(N)}$ for the empirical measure, may not be independent across particles (i.e., not a product measure and so nonchaotic). To handle this, we first establish a subsequential large deviation result using Theorem IV.1, and then show that the rate function is unique. It is well known in the theory of weak convergence of probability measures that given a sequence of probability measures $(Q_N, N \geq 1)$, if every subsequence has a further subsequence that converges weakly to the same limit Q, then Q_N converges weakly to Q. An analogous statement holds for large deviations: if every subsequence $(Q_{N_k}, k \geq 1)$ has a further subsequence $(Q_{N_{k_l}}, l \geq 1)$ that satisfies the large deviation principle with speed N_{k_l} and the same rate function s, and the sequence is exponentially tight, then the original sequence satisfies the large deviation principle with speed N and rate s.

2) *Empirical measure at terminal times*: Fix a finite duration $[0, T]$. Let us use the notation $\mu_N(0) = \nu_N$ and $\mu(0) = \nu$. Suppose that the initial conditions satisfy $\nu_N \to \nu$ weakly. (No exchangeability assumption is made on the distribution of the particles, and so the initial conditions need not be chaotic). Recall that $p_{\nu_N}(N)$ is the law of $\mu_N : [0, T] \to \mathcal{M}_1(\mathcal{Z})$. Let $p_{\nu_N,T}(N)$ be the law of $\mu_N(T)$. Then, since the projection map $\mu \mapsto \mu(T)$ is continuous, Theorem IV.1 and the contraction principle imply that $(p_{\nu_N,T}(N), N \geq 1)$ satisfies the large deviation principle with speed N and rate

$$S_T(\xi|\nu) = \inf\{S_{[0,T]}(\mu|\nu) | \mu(0) = \nu, \mu(T) = \xi, \mu \in \mathcal{A}\}. \quad (2)$$

One can also extract a minimizing path $\overline{\mu} : [0, T] \to \mathcal{M}_1(\mathcal{Z})$ that moves from $\overline{\mu}(0) = \nu$ to $\overline{\mu}(T) = \xi$, and time varying rates

$(l_{i,j}, t \in [0,T])$ with an associated rate matrix $L(t)$, such that the minimizing path is now the normal path under control $L(t)$, i.e.,

$$\dot{\overline{\mu}}(t) = L(t)^* \overline{\mu}(t),$$

and $S_T(\xi|\nu) = S_{[0,T]}(\overline{\mu}|\nu)$.

3) *Empirical measure at initial and terminal times*: Suppose that the initial empirical measure $\mu_N(0)$ has law $\wp_0^{(N)}$. Let $\wp_{0,T}^{(0)}$ denote the joint law of the pair $(\mu_N(0), \mu_N(T))$. If the sequence of initial conditions $(\wp_0^{(N)}, N \geq 1)$ satisfies the large deviation principle with a good rate function s, then using a result for product distributions, see Feng and Kurtz [16, Prop. 3.25], and using some regularity properties of $S_T(\xi|\nu)$, one can establish that the sequence of joint laws $(\wp_{0,T}^{(N)}, N \geq 1)$ satisfies the large deviation principle with speed N and good rate function

$$S_{0,T}(\nu, \xi) = s(\nu) + S_T(\xi|\nu).$$

If one applies the contraction principle for the map $\mu \mapsto \mu(T)$, then the sequence of terminal measures $(\wp_T^{(N)}, N \geq 1)$ satisfies the large deviation principle with good rate function

$$\inf_{\nu \in \mathcal{M}_1(\mathcal{Z})} \{s(\nu) + S_T(\xi|\nu)\}. \quad (3)$$

4) *Invariant measure*: We saw that the Markov chain $\mu_N(\cdot)$ has a unique invariant measure $\wp^{(N)}$ for each N. Since $\mathcal{M}_1(\mathcal{Z})$ may be viewed as a compact subset of $\mathbb{R}^r$, there is a subsequential large deviation principle for the sequence of invariant measures. From this and the fact that invariance means $\wp_0^{(N)} = \wp_T^{(N)} = \wp^{(N)}$, one recognizes that (3) evaluates to $s(\xi)$, and so this rate function for the subsequential large deviation principle satisfies the Hamilton-Jacobi-Bellman equation

$$s(\xi) = \inf_{\nu \in \mathcal{M}_1(\mathcal{Z})} \{s(\nu) + S_T(\xi|\nu)\}. \quad (4)$$

Note that the above equation says that $s(\xi)$ has the interpretation that it is a minimum over all paths that last T units of time and end at ξ after taking into account the cost for the initial state; see (2). We then show that this equation holds for a countable set of time durations, say with T replaced by mT' with $m \geq 1$, and extract a single path of infinite duration ending at ξ, each of whose pieces of duration mT' is optimal and satisfies (4) with equality. If we view this path in *reverse time* and call it $\hat{\mu}(\cdot)$, it starts at $\hat{\mu}(0) = \xi$ and satisfies

$$\dot{\hat{\mu}}(t) = -\hat{L}(t)^* \hat{\mu}(t), \quad t > 0,$$

for some family of rate matrices $\hat{L}(\cdot)$. When there is a unique globally asymptotically stable equilibrium ξ_0, we show that $s(\xi_0) = 0$, and that the chosen infinite duration path must converge to ξ_0. Using (1) in reversed time, and letting $\hat{\mathcal{E}}$ denote the reversed edges of $\mathcal{E}$, one then gets

$$s(\xi) = \inf_{\hat{\mu}} \int_{[0,+\infty)} \Bigg[\sum_{(i,j) \in \hat{\mathcal{E}}} (\hat{\mu}(t)(j)) \hat{\lambda}_{i,j}(\hat{\mu}(t)) \, \tau^* \left(\frac{l_{i,j}(t)}{\hat{\lambda}_{i,j}(\hat{\mu}(t))} - 1 \right) \Bigg] \, dt \quad (5)$$

where the infimum is over all $\hat{\mu}$ that lie in $\mathcal{M}_1(\mathcal{Z})$ and are solutions to the dynamical system $\dot{\hat{\mu}}(t) = -\hat{L}(t)^* \hat{\mu}(t)$ for some family of rate matrices $\hat{L}(\cdot) = (l_{i,j}(\cdot))_{i,j \in \mathcal{Z}}$, with initial condition $\hat{\mu}(0) = \xi$ and terminal condition $\lim_{t \to +\infty} \hat{\mu}(t) = \xi_0$.

Since this does not depend on the subsequence, the subsequential large deviation principle is indeed a full large deviation principle. We now summarize the main theorem.

Theorem VI.1. *Let Assumption* (**A**) *hold. Let the graph with vertex set $\mathcal{Z}$ and edge set $\mathcal{E}$ be irreducible. Let the McKean-Vlasov equation $\dot{\mu}(t) = A^*_{\mu(t)} \mu(t)$ have a unique globally asymptotically stable equilibrium ξ_0. Then the sequence $(\wp^{(N)}, N \geq 1)$ satisfies the large deviation principle with speed N and good rate function s given by (5).*

ACKNOWLEDGMENTS

This work was supported by a J. C. Bose Fellowship and by IFCPAR (Indo-French Centre for the Promotion of Advanced Research), Project 4000-IT-1.

REFERENCES

[1] ANANTHARAM, V. (1991) "A mean field limit for a lattice caricature of dynamic routing in circuit switched networks", *Annals of Appl. Prob.* **1**, 481–503.

[2] ANANTHARAM, V. and BENCHEKROUN, M. (1993) "A technique for computing sojourn times in large networks of interacting queues", *Probability in the Engineering and Informational Sciences* **7**, 441–464.

[3] BENAIM, M. and LE BOUDEC J.-Y. (2008) "A class of mean field interaction models for computer and communication systems", *Perform. Eval.* **65**, 823–838.

[4] BIANCHI, G. (1998) "IEEE 802.11 - saturated throughput analysis", *IEEE Comm. Lett.* **12**, 318–320.

[5] BISWAS, A. and BORKAR, V. S. (2009) "Small noise asymptotics for invariant densities for a class of diffusions: a control theoretic view", *Journal of Mathematical Analysis and Applications* **360**, 476–484.

[6] BOORSTYN, R. R., KERSENBAUM, A., MAGLARIS, B. and SAHIN, V. (1987) "Throughput analysis in multihop CSMA packet radio networks", *IEEE Trans. on Commun.* **35**(3), 267-274.

[7] BORDENAVE, C., MCDONALD, D. and PROUTIÈRE, A. (2010) "A particle system in interaction with a rapidly varying environment: mean field limits and applications", *Networks and Heterogeneous Media*, **5**(1).

[8] BORDENAVE, C., MCDONALD, D. and PROUTIÈRE, A. (2005) "Random multi-access algorithms, a mean field analysis", *Proc. 43th Allerton conference*.

[9] BORKAR, V. and SUNDARESAN, R. (2011) "Asymptotics of the invariant measure in mean field models with jumps", http://arxiv.org/abs/1107.4142

[10] DAI PRA, P. and DEN HOLLANDER, F. (1996), "McKean-Vlasov limit for interacting random processes in random media", *Journal of Statistical Physics*, **84**(3/4), 735–772.

[11] DAWSON, D. A. and GÄRTNER, J. (1987) "Large deviations from the McKean-Vlasov limit for weakly interacting diffusions", *Stochastics*, **20**, 247–308.

[12] DAY, M. V. (1987) "Recent progress on the small parameter exit problem", *Stochastics*, **20**, 121–150.

[13] DEL MORAL, P. and ZAJIC, T. (2003) "Note on the Laplace-Varadhan Integral Lemma", *Bernoulli*, **9**(1), 49–65.

[14] DJEHICHE, B. and KAJ, I. (1995) "The rate function for some measure-valued jump processes", *The Annals of Probability*, **23**(3), 1414–1438.

[15] DUFFY, K. R. (2010) "Mean field Markov models of wireless local area networks", *Markov Processes and Related Fields* **16**(2), 295–328.

[16] FENG, J. and KURTZ, T. G. (2006) *Large Deviations for Stochastic Processes*, Mathematical Surveys and Monographs, **131**. American Math. Soc., Providence, RI.

[17] FENG, S. (1994) "Large deviations for empirical process of mean-field interacting particle system with unbounded jumps", *The Annals of Probability*, **22**(4), 2122–2151.

[18] FREIDLIN, M. I. and WENTZELL, A. D. (1998) *Random Perturbations of Dynamical Systems* (2nd ed.), Springer Verlag, Berlin–Heidelberg.

[19] Graham, C. (2000) "Chaoticity on path space for a queueing network with selection of the shortest queue among several", *J. Appl. Probab.* **37**(1), 198–211.

[20] KAC, M. (1956) "Foundations of kinetic theory", in '*Proc. of 3rd Berkeley Symp. on Math. Stat. and Prob.*' (J. Neyman, ed.), Uni. of California Press, 171–197.

[21] LÉONARD, C. (1995) "Large deviations for long range interacting particle systems with jumps", *Annales de l'I. H. P., section B*, **31**(2), 289–323.

[22] MCKEAN, H. P. (1966) "A class of Markov processes associated with nonlinear parabolic equations", *Proc. Nat. Acad. Sci.* **56**, 1907–1911.

[23] RAMAIYAN, V., KUMAR, A. and ALTMAN, E. (2008) "Fixed point analysis of single cell IEEE 802.11e WLANs: uniqueness and multistability", *IEEE/ACM Trans. Network.* **16**, 1080–1093.

[24] SHARMA, G., GANESH, A. J. and KEY, P. (2009) "Performance analysis of contention based medium access protocols", *IEEE Trans. Information Th.* **55**(4), 1665–1682. Also reported in *Proc. IEEE Infocom (2006)*.

[25] SHEU, S-J. (1985) "Stochastic control and exit probabilities of jump processes", *SIAM J. Control and Optimization* **23**(2), 306–328.

[26] SHEU, S-J. (1986) "Asymptotic behavior of the invariant density of a diffusion Markov process with small diffusion", *SIAM J. Math. Anal.* **17**(2), 451–460.

[27] SHWARTZ, A. and WEISS, A. (1995) *Large Deviations for Performance Analysis*, Chapman & Hall.

[28] STOLYAR, A. L. (1989) "Asymptotic behavior of the stationary distribution for a closed queueing system", *Problemy Peredachi Informatsii*, **25**(4), 80–92.

[29] SZNITMAN, A. S. (1991) "Topics in propogation of chaos", in ' *Ecole d'été de Probabilités de Saint Flour (1989)*', Springer Lecture Notes in Mathematics No. **1464**, 166–251