Asymptotics of the invariant distribution in a mean-field model with jumps

Rajesh Sundaresan

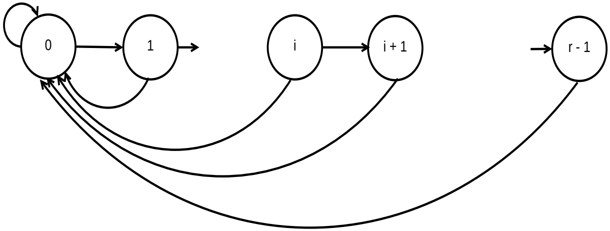
Indian Institute of Science

28 February 2013

Based on joint work with Vivek S. Borkar

Wireless Local Area Network (WLAN) interactions

- N nodes or particles accessing the medium in a wireless LAN
- State space for each particle: $\mathcal{Z} = \{0, 1, \cdots, r-1\}$
- Transitions: From state i to either $i+1$ or 0



- r: Maximum number of transmission attempts before discard
- Coupled dynamics: Transition rate for success or failure depends on empirical distribution of nodes across states

Mean field dynamics

- Empirical measure $\mu_N(t)$: Fraction of nodes in each state
  - $X_n^{(N)}(t)$: state of nth particle at time t

$$\mu_N(t) = \frac{1}{N}\sum_{n=1}^{N} \delta_{\{X_n^{(N)}(t)\}} \in \mathcal{M}_1(\mathcal{Z}), \text{ space of probability measures on } \mathcal{Z}$$

- A node transits from state i to state j at time t with rate $\lambda_{i,j}(\mu_N(t))$
- In general, allowed transitions are specified by the directed edges $\mathcal{E}$ on the vertex set $\mathcal{Z}$

Example rate functions

This example comes from a discrete-time (slotted-ALOHA) model

- Slot size $1/N$, access probability in each slot is c_i/N when node is in state i, with $c \in \mathbb{R}_+^r$
- Assume three states $r = 3$.

With $\xi \in \mathcal{M}_1(\mathcal{Z})$, the rate matrix is

$$\Lambda(\xi) = \begin{bmatrix} -(\cdot) & c_1(1 - e^{-c^T\xi}) & 0 \\ c_2 e^{-c^T\xi} & -(\cdot) & c_2(1 - e^{-c^T\xi}) \\ c_3 e^{-c^T\xi} & 0 & -(\cdot) \end{bmatrix}.$$

Interpretation: $c^T\xi$ is like a load factor when $\mu_N(t) = \xi$. Success probability is $e^{-c^T\xi}$ if an attempt is made.

The Markov processes, big and small

- $(X_n^{(N)}(\cdot),\ 1 \leq n \leq N)$ is Markov
- State space grows exponentially with N: size r^N
- Study $\mu_N(\cdot)$ instead, also a Markov process
  Its state space size is at most $(N+1)^r$, and is a subset of $\mathcal{M}_1(\mathcal{Z})$
- We will focus mostly on $\mu_N(\cdot)$

The smaller Markov process

- $\mu_N(\cdot)$ is a Markov process
- The transition from ξ to $\xi + \frac{1}{N}e_j - \frac{1}{N}e_i$ occurs with rate $N\xi(i)\lambda_{i,j}(\xi)$
- For large N, changes are small, $O(1/N)$, at higher rates, $O(N)$.
- Familiar setting of Kurtz's theorem where μ_N converges to a deterministic limit given by an ODE.

The conditional expected drift in μ_N

- For each k,

$$\dot{\xi}_k = \sum_{i:i\neq k} \xi_i \lambda_{i,k}(\xi) - \xi_k \sum_{i:i\neq k} \lambda_{k,i}$$

- With $\Lambda(\xi) = [\ \lambda_{i,j}(\xi)\ ]$, we get

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

- First order approximation: ignore the randomness in $\mu_N(\cdot)$, and set it to its mean evolution given by

$$\dot{\mu}(t) = \Lambda(\mu(t))^* \ \mu(t), \quad t > 0 \qquad \text{[McKean-Vlasov equation]}$$

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

- State space is a more familiar compact set, but evolution is nonlinear

The conditional expected drift in μ_N

- For each k,

$$\dot{\xi}_k = \sum_{i:i \neq k} \xi_i \lambda_{i,k}(\xi) - \xi_k \sum_{i:i \neq k} \lambda_{k,i}$$

- With $\Lambda(\xi) = [\ \lambda_{i,j}(\xi)\ ]$, we get

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

- First order approximation: ignore the randomness in $\mu_N(\cdot)$, and set it to its mean evolution given by

$$\dot{\mu}(t) = \Lambda(\mu(t))^* \ \mu(t), \quad t > 0 \qquad \text{[McKean-Vlasov equation]}$$

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

- State space is a more familiar compact set, but evolution is nonlinear

Assumptions

- The graph with vertex set $\mathcal{Z}$ and edge set $\mathcal{E}$ is irreducible.
  Holds in our WLAN setting

- There exist positive constants $c > 0$ and $C < +\infty$ such that, for every $(i, j) \in \mathcal{E}$, we have

$$c \leq \lambda_{i,j}(\cdot) \leq C$$

- The mapping $\mu \mapsto \lambda_{i,j}(\mu)$ is Lipschitz continuous over $\mathcal{M}_1(\mathcal{Z})$

Kurtz's theorem

- Consider $D([0,T],\mathcal{M}_1(\mathcal{Z}))$, cadlag, measure-valued paths, and equip it with the metric

$$\rho_T(\xi,\xi') = \sup_{t\in[0,T]} ||\xi(t) - \xi'(t)||_1$$

where $||\cdot||_1$ is the L^1 metric

- Finer than Skorohod topology; not separable

Theorem

Let $\mu_N(0) \to \mu(0)$ weakly. Let $T > 0$ be arbitrary, but finite. Then, for every $\varepsilon > 0$, we have

$$\lim_{N\to+\infty} \Pr\{\rho_T(\mu_N,\mu) > \varepsilon\} = 0.$$

Approximation over *finite time durations*

Formally ...

- For any $\Phi : \mathcal{M}_1(\mathcal{Z}) \to \mathbb{R}$ that is bounded and continuous, the conditional expected drift starting from ξ

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

if Φ has bounded second order derivatives.

Back to the individual particles

- Let $\mu(\cdot)$ be the solution to the McKean-Vlasov dynamics
- Tag a particle.
  - It is likely to be in state i with probability $\mu(t)(i)$.
  - Its evolution is described asymptotically by a Markov process with time-dependent transition rates $\lambda_{i,j}(\mu(t))$
- Tag k particles.
  - If their states are independent at time 0, then they evolve (in the asymptotics of large N) independently of each other under the mean field $\mu(\cdot)$

Large deviation principle?

- From simulations, exponentially fast concentration
- A large deviation principle holds for some class of Markov processes. Shwartz and Weiss, Freidlin and Wentzell, Leonard, and others.
- The case under consideration does not satisfy the conditions assumed in these works.

Large deviation principle (LDP)

- *Definition*: The sequence $(p^{(N)}, N \geq 1)$ of probability measures on the metric space $D([0,T], \mathcal{M}_1(\mathcal{Z}))$ satisfies the LDP with speed N and good rate function $S_{[0,T]}(\mu)$ if
  - For every open set G and closed set F of the metric space $D([0,T], \mathcal{M}_1(\mathcal{Z}))$, we have

$$\begin{aligned} \liminf_{N \to +\infty} \frac{\log p^{(N)}(G)}{N} &\geq -\inf_{\mu \in G} S_{[0,T]}(\mu) \\ \limsup_{N \to +\infty} \frac{\log p^{(N)}(F)}{N} &\leq -\inf_{\mu \in F} S_{[0,T]}(\mu) \end{aligned}$$

  - For each $a \in [0, +\infty)$, the level sets $\{\mu : S_{[0,T]}(\mu) \leq a\}$ are compact

Relative entropy between two inhomogenous Poisson point processes

Let us understand a simpler case first ...

- P: Poisson point process on $[0, T]$ with intensity $\eta(t)$
  Q: Poisson point process on $[0, T]$ with intensity $\zeta(t)$

- Sanov's theorem: Sample N iid paths from Q. The probability that the empirical measure is in a small neighbourhood near P is $\approx e^{-NI(P||Q)}$ where

$$
\begin{aligned}
I(P||Q) &= \int_{[0,T]} \left[\eta(t) \log \frac{\eta(t)}{\zeta(t)} - \eta(t) + \zeta(t) \right] dt \\
&= \int_{[0,T]} \left[\zeta(t)\ \tau^{*} \left(\frac{\eta(t)}{\zeta(t)} - 1 \right) \right] dt
\end{aligned}
$$

  where $\tau^{*}(u) = (u+1)\log(u+1) - u,\ u \geq -1$

Heuristics

- Find probability of being near a deviant path μ, a solution to

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

- Normal intensity for an (i,j) jump at time t is $(\mu(t)(i)) \times \lambda_{i,j}(\mu(t))$

- Empirical distribution should be near that of iid sampling from $(\mu(t)(i)) \times \lambda_{i,j}(\mu(t))$

- But the path μ appears to have intensity for (i,j) jump at time t given by $(\mu(t)(i)) \times l_{i,j}(t)$

- Add up the relative entropies for each jump process indexed by $(i,j) \in \mathcal{E}$

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

Finite duration LDP

Theorem

Suppose that the initial conditions $\nu_N \to \nu$ *weakly.*

Then the sequence $(p^{(N)}_{\nu_N}, N \geq 1)$ *satisfies the LDP on* $D([0,T], \mathcal{M}_1(\mathcal{Z}))$ *(with metric* ρ_T*) with speed* N *and a good rate function* $S_{[0,T]}(\mu|\nu)$*.*

Proof steps

- Apply Sanov's theorem to noninteracting system on path space
- Use the Laplace-Varadhan principle to extract a path space LDP
- Then use the contraction principle (from an LDP for the empirical measure in path space to an LDP for the law of $\mu_N(\cdot)$).

Corollary:
$p_{\nu_N}^{(N)} \to \delta_{\mu(\cdot)}$ weakly, where $\mu(\cdot)$ is the McKean-Vlasov solution

Proof steps

- Apply Sanov's theorem to noninteracting system on path space
- Use the Laplace-Varadhan principle to extract a path space LDP
- Then use the contraction principle (from an LDP for the empirical measure in path space to an LDP for the law of $\mu_N(\cdot)$).

Corollary:
$p^{(N)}_{\nu_N} \to \delta_{\mu(\cdot)}$ weakly, where $\mu(\cdot)$ is the McKean-Vlasov solution

Assumptions again

- The graph with vertex set $\mathcal{Z}$ and edge set $\mathcal{E}$ is irreducible
  Holds in the our WLAN case

- There exist positive constants $c > 0$ and $C < +\infty$ such that, for every $(i, j) \in \mathcal{E}$, we have

$$c \leq \lambda_{i,j}(\cdot) < C$$

- The mapping $\mu \mapsto \lambda_{i,j}(\mu)$ is Lipschitz continuous over $\mathcal{M}_1(\mathcal{Z})$

- Take $\lim_{t\to+\infty} \lim_{N\to+\infty}(\cdots)$
  Assume that the McKean-Vlasov equation $\dot{\mu}(t) = \Lambda(\mu(t))^* \mu(t)$
  - Has a unique equilibrium ξ_0 (i.e., $\Lambda(\xi_0)^*\xi_0 = 0$)
  - The equilibrium ξ_0 is globally asymptotically stable

  Then $\lim_{t\to+\infty} \mu(t) = \xi_0$ for any initial condition

Assumptions again

- The graph with vertex set $\mathcal{Z}$ and edge set $\mathcal{E}$ is irreducible
  Holds in the our WLAN case

- There exist positive constants $c > 0$ and $C < +\infty$ such that, for every $(i, j) \in \mathcal{E}$, we have

$$c \leq \lambda_{i,j}(\cdot) < C$$

- The mapping $\mu \mapsto \lambda_{i,j}(\mu)$ is Lipschitz continuous over $\mathcal{M}_1(\mathcal{Z})$

- Take $\lim_{t\to+\infty} \lim_{N\to+\infty}(\cdots)$
  Assume that the McKean-Vlasov equation $\dot{\mu}(t) = \Lambda(\mu(t))^* \mu(t)$
  - Has a unique equilibrium ξ_0 (i.e., $\Lambda(\xi_0)^* \xi_0 = 0$)
  - The equilibrium ξ_0 is globally asymptotically stable

  Then $\lim_{t\to+\infty} \mu(t) = \xi_0$ for any initial condition

When $t \to +\infty$ first ... large time behaviour

- Let the directed graph $G(\mathcal{Z}, \mathcal{E})$ be irreducible. Then, for a fixed N, the Markov chain μ_N is irreducible with a finite state space. It therefore has a unique stationary distribution: $\wp^{(N)} = \mathscr{L}_{st}(\mu_N(+\infty))$

- Does $\wp^{(N)} \to \delta_{\xi_0}$?
(Stolyar 1989, Anantharam 1991, Anantharam and Benchekroun 1993, Bordenave et al. 2005/2007, Benaim and Le Boudec 2008)

- Decoupling approximation

- Large deviations from this limit?

Large deviations for the invariant measure

- If $\mu_N(+\infty)$ is near ξ, then this is most likely due to an excursion that began at ξ_0, worked against the attractor ξ_0, and took the lowest cost path to ξ over all possible time durations

- Looking backwards in time, the dynamics must be

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

with $\hat{\mu}(0) = \xi$, $\lim_{t \to +\infty} \hat{\mu}(t) = \xi_0$, and $\hat{L}(t)$ is some family of rate matrices. Define

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

Theorem

Under the stated assumptions, the sequence $(\wp^{(N)}, N \geq 1)$ satisfies the LDP with speed N and good rate function $s(\cdot)$.

Large deviations for the invariant measure

- If $\mu_N(+\infty)$ is near ξ, then this is most likely due to an excursion that began at ξ_0, worked against the attractor ξ_0, and took the lowest cost path to ξ over all possible time durations

- Looking backwards in time, the dynamics must be

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

with $\hat{\mu}(0) = \xi$, $\lim_{t \to +\infty} \hat{\mu}(t) = \xi_0$, and $\hat{L}(t)$ is some family of rate matrices. Define

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

Theorem

Under the stated assumptions, the sequence $(\wp^{(N)}, N \geq 1)$ *satisfies the LDP with speed* N *and good rate function* $s(\cdot)$.

Summary

- Asymptotics of mean field limits in WLANs
- A finite duration LDP
- When there is a unique globally stable equilibrium ξ_0 for the McKean-Vlasov equation, the invariant measure satisfies the LDP. The rate function $s(\xi)$ is characterised by the cost of an optimal control that moves the system from ξ to ξ_0 in reversed time
- Extension to cases with multiple equilibria.
- arXiv:1107.4142

Proof steps

- Given $\nu_N \to \nu$, extract LDP for the laws for terminal state (finite T), via contraction principle, with rate function
$$S_T(\xi|\nu) = \inf \{S_{[0,T]}(\mu|\nu) \mid \mu(0) = \nu, \mu(T) = \xi\}$$
- If the laws for initial states satisfy the LDP with a good rate function $s(\nu)$, argue that joint laws for initial and terminal states satisfy the LDP with a good rate function $s(\nu) + S_T(\xi|\nu)$. Then apply contraction principle to get that the laws for the terminal states satisfy the LDP with good rate function
$$\inf_{\nu \in \mathcal{M}_1(\mathcal{Z})} \{s(\nu) + S_T(\xi|\nu)\}$$
- The invariant measures $(\wp^{(N)}, N \geq 1)$ live on a compact space. So, given any subsequence, there is a further subsequential LDP with appropriate speed, and with rate function $s(\xi)$ that satisfies
$$s(\xi) = \inf_{\nu \in \mathcal{M}_1(\mathcal{Z})} \{s(\nu) + S_T(\xi|\nu)\}$$

Proof steps continued

- By the assumption that ξ_0 is a unique equilibrium that is globally stable, we can show $s(\xi_0) = 0$.

- Extract a single infinite duration path $\hat{\mu}(\cdot)$ that is optimal, i.e., it attains the infimum for each duration $[0, mT]$, $\hat{\mu}(0) = \xi$, and satisfies

$$\begin{aligned} s(\xi) &= s(\hat{\mu}(mT)) + S_{mT}(\xi|\nu), \quad \forall m \geq 1 \\ &= s(\hat{\mu}(mT)) + \int_{[0,mT]} [\cdots]\, dt \end{aligned}$$

- The integrand in the second term is nonnegative; the second term increases with m, and so the first term $s(\hat{\mu}(mT))$ decreases with m. Since $s(\cdot)$ is bounded below by 0, $s(\hat{\mu}(mT))$ must converge to a constant as $m \to +\infty$

Proof steps continued even further

- So the increment $\int_{mT}^{mT+T}[\cdots]dt \to 0$ in the second term, and in the limit, integrand must be 0 a.e., which is a McKean-Vlasov path in reversed time.
  More precisely, $\hat{\mu}(\cdot)$ has an ω-limit set that is positively invariant to (McKean-Vlasov dynamics in reversed time)

$$\hat{\mu}(t) = -\Lambda(\hat{\mu}(t))^* \hat{\mu}(t), \quad t \geq 0$$

- This limit set is also invariant to McKean-Vlasov dynamics. It is further compact and bounded within $\mathcal{M}_1(\mathcal{Z})$. The only such set invariant set is $\{\xi_0\}$. So $\hat{\mu}(mT) \to \xi_0$.

- Taking limit as $m \to +\infty$,

$$s(\xi) = s(\xi_0) + \int_{[0,+\infty)} [\cdots]\ dt = 0 + \int_{[0,+\infty)} [\cdots]\ dt$$

  This expression is the same regardless of the initial subsequence

- Thus every subsequence has a further subsequence that satisfies the LDP with appropriate speed and the same rate function $s(\cdot)$.