Large time behaviour of finite state mean-field interacting particle systems

Sarath A Y

Joint work with Rajesh Sundaresan

ECE Department, Indian Institute of Science

12 July 2020

System model

- N particles.
- State of the nth particle at time t is $X_n^N(t) \in \mathcal{Z}$.

System model

- N particles.
- State of the nth particle at time t is $X_n^N(t) \in \mathcal{Z}$.
- Certain allowed transitions: specified by a directed graph $(\mathcal{Z}, \mathcal{E})$.

System model

- N particles.
- State of the nth particle at time t is $X_n^N(t) \in \mathcal{Z}$.
- Certain allowed transitions: specified by a directed graph $(\mathcal{Z}, \mathcal{E})$.
- Empirical measure at time t

$$\mu_N(t) = \frac{1}{N} \sum_{n=1}^{N} \delta_{X_n^N(t)} \in M_1(\mathcal{Z}).$$

System model

- N particles.
- State of the nth particle at time t is $X_n^N(t) \in \mathcal{Z}$.
- Certain allowed transitions: specified by a directed graph $(\mathcal{Z}, \mathcal{E})$.
- Empirical measure at time t

$$\mu_N(t) = \frac{1}{N} \sum_{n=1}^{N} \delta_{X_n^N(t)} \in M_1(\mathcal{Z}).$$

- For each $(z, z') \in \mathcal{E}$, we have a function $\lambda_{z,z'} : M_1(\mathcal{Z}) \to [0, +\infty)$.
- Particle transitions: at time t, a $z \to z'$ transition occurs at rate $\lambda_{z,z'}(\mu_N(t))$. Mean-field interaction.

System model

- N particles.
- State of the nth particle at time t is $X_n^N(t) \in \mathcal{Z}$.
- Certain allowed transitions: specified by a directed graph $(\mathcal{Z}, \mathcal{E})$.
- Empirical measure at time t

$$\mu_N(t) = \frac{1}{N} \sum_{n=1}^{N} \delta_{X_n^N(t)} \in M_1(\mathcal{Z}).$$

- For each $(z, z') \in \mathcal{E}$, we have a function $\lambda_{z,z'} : M_1(\mathcal{Z}) \to [0, +\infty)$.
- Particle transitions: at time t, a $z \to z'$ transition occurs at rate $\lambda_{z,z'}(\mu_N(t))$. Mean-field interaction.
- $\{(X_n^N(t), 1 \le n \le N), t \ge 0\}$ is a Markov process on $\mathcal{Z}^N$.

The empirical measure process μ_N

- $\{\mu_N(t), t \geq 0\}$ is also a Markov process on $M_1(\mathcal{Z})$ with infinitesimal generator

$$L^N f(\xi) = \sum_{(z,z')\in\mathcal{E}} N\xi(z)\lambda_{z,z'}(\xi)\left[f\left(\xi + \frac{\delta_{z'}}{N} - \frac{\delta_z}{N}\right) - f(\xi)\right].$$

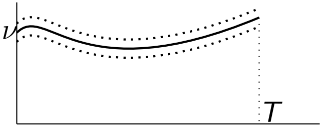
The empirical measure process μ_N

- $\{\mu_N(t), t \geq 0\}$ is also a Markov process on $M_1(\mathcal{Z})$ with infinitesimal generator

$$L^N f(\xi) = \sum_{(z,z')\in\mathcal{E}} N\xi(z)\lambda_{z,z'}(\xi)\left[f\left(\xi + \frac{\delta_{z'}}{N} - \frac{\delta_z}{N}\right) - f(\xi)\right].$$

- Typical behaviour of μ_N (mean-field limit):
Let $\mu_N(0) \to \nu$ weakly as $N \to \infty$. Then $\{\mu_N(t), t \geq 0\}$ is "close to" the solution to the McKean-Vlasov equation:

$$\dot{\mu}(t) = \Lambda^*(\mu(t))\mu(t), \; \mu(0) = \nu.$$



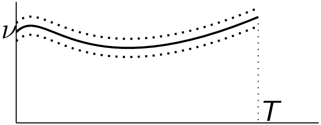
The empirical measure process μ_N

- $\{\mu_N(t), t \geq 0\}$ is also a Markov process on $M_1(\mathcal{Z})$ with infinitesimal generator

$$L^N f(\xi) = \sum_{(z,z')\in\mathcal{E}} N\xi(z)\lambda_{z,z'}(\xi)\left[f\left(\xi + \frac{\delta_{z'}}{N} - \frac{\delta_z}{N}\right) - f(\xi)\right].$$

- Typical behaviour of μ_N (mean-field limit):
Let $\mu_N(0) \to \nu$ weakly as $N \to \infty$. Then $\{\mu_N(t), t \geq 0\}$ is "close to" the solution to the McKean-Vlasov equation:

$$\dot{\mu}(t) = \Lambda^*(\mu(t))\mu(t),\ \mu(0) = \nu.$$



- Our interest: study of the large time behaviour of μ_N when the above ODE has multiple stable equilibria.

An Example: Interaction in WiFi networks

- N nodes accessing a common wireless medium.
- Interaction among nodes via the distributed MAC protocol.

An Example: Interaction in WiFi networks

- N nodes accessing a common wireless medium.
- Interaction among nodes via the distributed MAC protocol.
- State $X_n^N(t)$ represents aggressiveness of packet transmission.

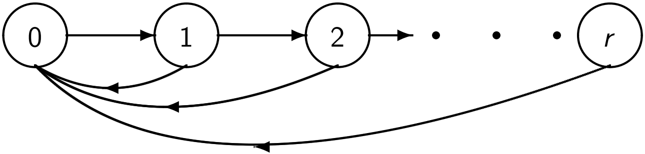


Figure: Set of allowed transitions in WiFi example

- State evolution:
  - Becomes less aggressive after a collision.
  - Moves to the most aggressive state after a successful packet transmission.

A sample path of μ_N in WiFi example

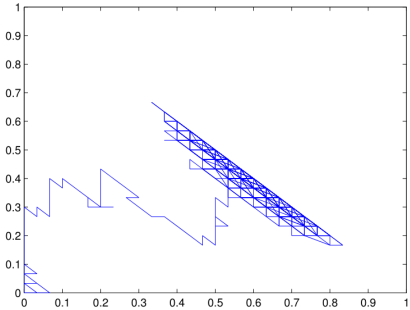


Figure: Evolution of states in a WiFi network under the MAC protocol

- Multiple stable regions in the system. Transition between two stable region occur over large time durations.

A sample path of μ_N in WiFi example

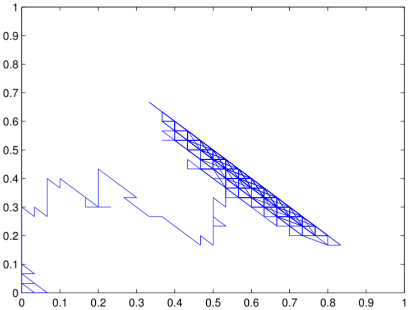


Figure: Evolution of states in a WiFi network under the MAC protocol

- Multiple stable regions in the system. Transition between two stable region occur over large time durations.
- Goals:
  - Study large time behaviour of μ_N
  - Mixing time of μ_N
  - Control μ_N to a desired equilibrium

Large deviations

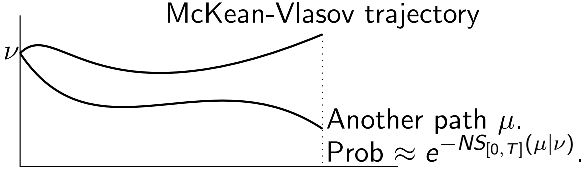
- S: a metric space. $\{X_N\}_{N \geq 1}$ is a sequence of S-valued random variables.
- Roughly, $P(X_N \in A) \sim \exp\{-NI_A\}$ where $I_A = \inf_{x \in A} I(x)$.

Large deviations

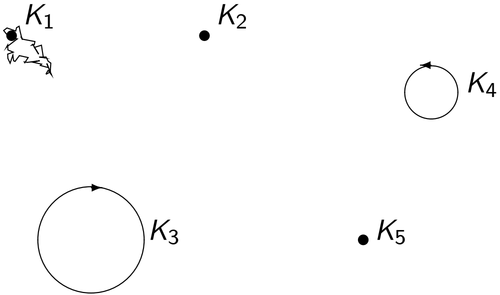
- S: a metric space. $\{X_N\}_{N\geq 1}$ is a sequence of S-valued random variables.
- Roughly, $P(X_N \in A) \sim \exp\{-NI_A\}$ where $I_A = \inf_{x\in A} I(x)$.
- Consider μ_N as a trajectory-valued random variable

Large deviations

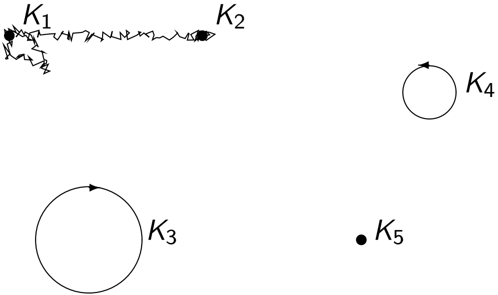
- S: a metric space. $\{X_N\}_{N\geq 1}$ is a sequence of S-valued random variables.
- Roughly, $P(X_N \in A) \sim \exp\{-NI_A\}$ where $I_A = \inf_{x\in A} I(x)$.
- Consider μ_N as a trajectory-valued random variable
- We want probability of deviations of μ_N from its typical trajectory



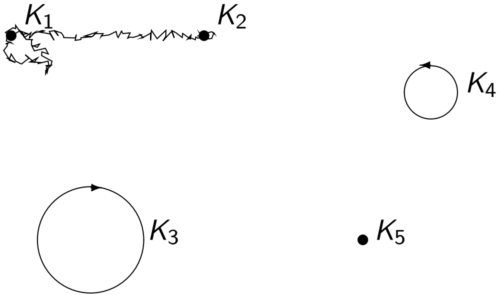
Approximation of μ_N using a discrete chain



Approximation of μ_N using a discrete chain

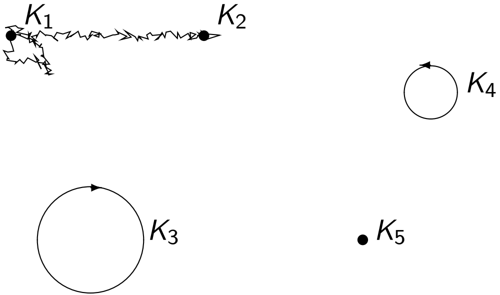


Approximation of μ_N using a discrete chain



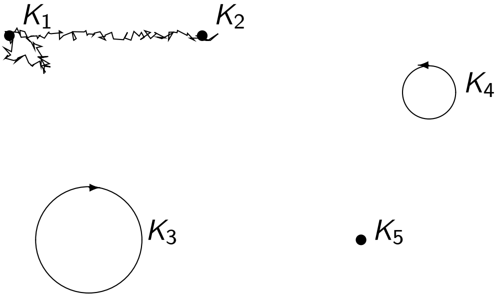
- τ_n: hitting time of μ_N in a given neighbourhood of K_i's.
- Hitting time chain: $Z_n^N = \mu_N(\tau_n)$, $n \geq 1$.

Approximation of μ_N using a discrete chain



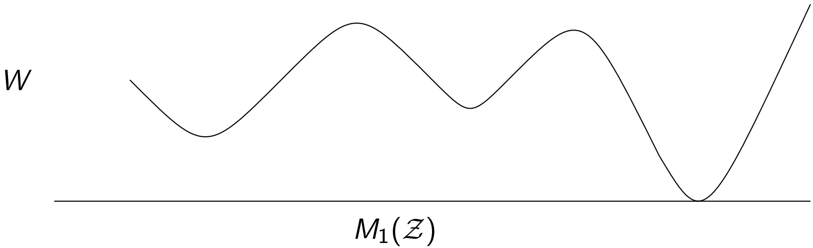
- τ_n: hitting time of μ_N in a given neighbourhood of K_i's.
- Hitting time chain: $Z_n^N = \mu_N(\tau_n)$, $n \geq 1$.
- Freidlin-Wentzell quasipotential $\tilde{V} : M_1(\mathcal{Z}) \times M_1(\mathcal{Z}) \to \mathbb{R}_+$.
    - $\tilde{V}(K_i, K_j)$: cost to go from K_i to K_j.

Approximation of μ_N using a discrete chain



- τ_n: hitting time of μ_N in a given neighbourhood of K_i's.
- Hitting time chain: $Z_n^N = \mu_N(\tau_n),\ n \geq 1$.
- Freidlin-Wentzell quasipotential $\tilde{V} : M_1(\mathcal{Z}) \times M_1(\mathcal{Z}) \to \mathbb{R}_+$.
  - $\tilde{V}(K_i, K_j)$: cost to go from K_i to K_j.
- One-step transition probability of Z^N:
  $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$.

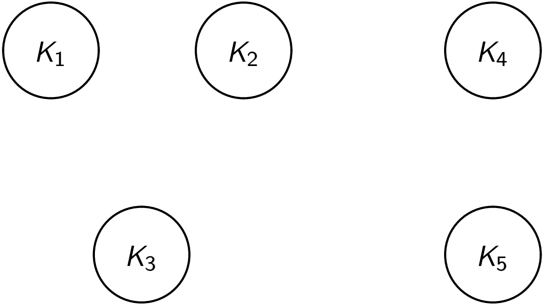
An entropy function



- $W(\xi) = \min_{1 \le i \le l}(W(K_i) + V(K_i, \xi)) - \min_{1 \le i \le l} W_i$.
- Stationary distribution scales like $\wp_N(\xi) \sim \exp\{-NW(\xi)\}$.
- Under stationarity, we find μ_N to be in a neighbourhood of the set of minimisers of W, with very high probability.

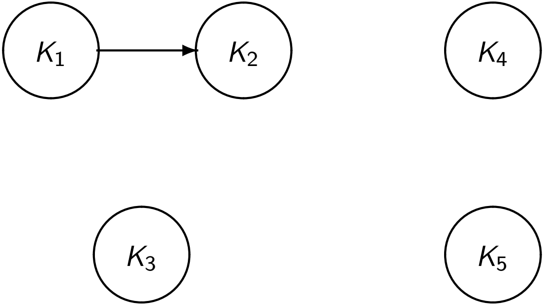
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



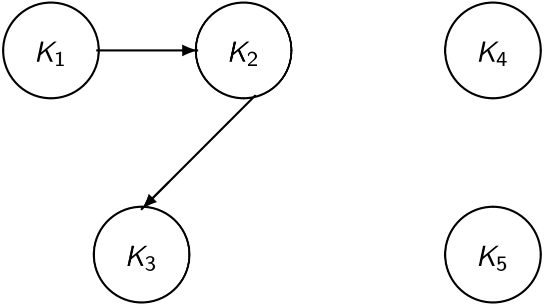
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



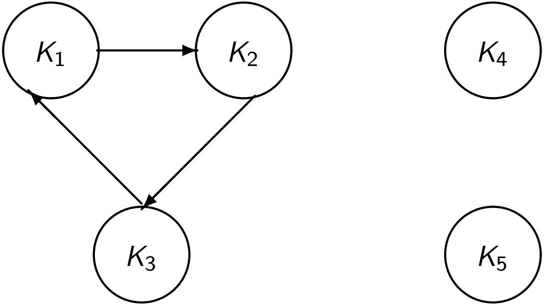
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



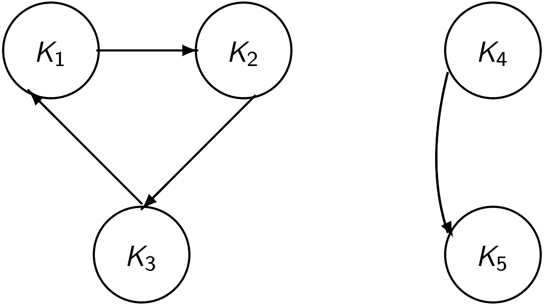
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



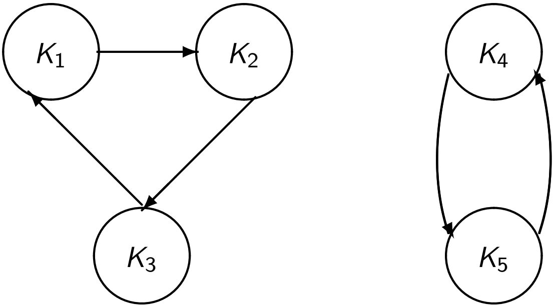
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



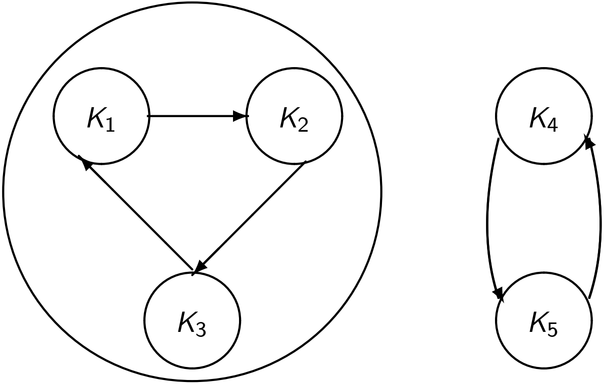
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



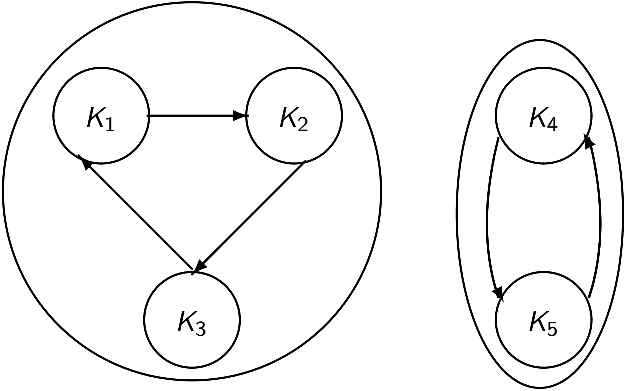
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



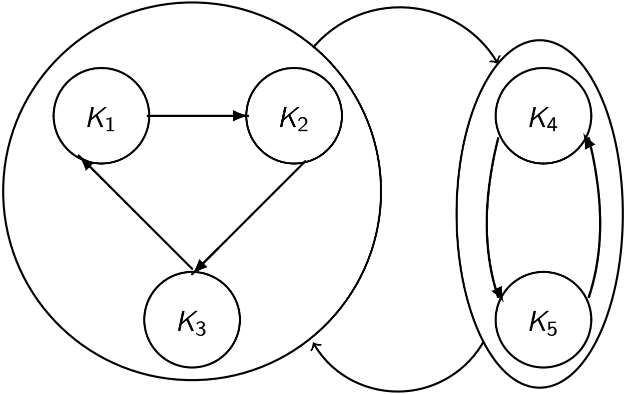
Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



Decomposition into cycles - example

- Recall: $P(K_i, K_j) \sim \exp\{-N\tilde{V}(K_i, K_j)\}$
- Consider $\tilde{V} = \begin{pmatrix} 0 & 4 & 9 & 13 & 12 \\ 7 & 0 & 5 & 10 & 11 \\ 6 & 8 & 0 & 17 & 15 \\ 3 & 6 & 8 & 0 & 2 \\ 5 & 7 & 10 & 3 & 0 \end{pmatrix}$



Large time behaviour

- Cycles are "very stable" subsets of $M_1(\mathcal{Z})$.

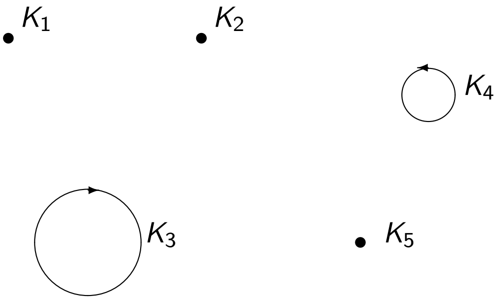
Theorem

Let π_1^k, π_2^k be k-cycles, $\pi_1^k \neq \pi_2^k$, and $K_i \in \pi_1^k$. Let $W = L \setminus \pi_1^k$. Given $\varepsilon > 0$, there exist $\rho > 0$ and $N_0 \geq 1$ such that for all $\rho_1 \leq \rho$, $\nu \in \gamma_i \cap M_1^N(\mathcal{Z})$ and $N \geq N_0$, we have

$$\begin{aligned}\exp\{-N(\tilde{V}(\pi_1^k, \pi_2^k) - \tilde{V}(\pi_1^k) + \varepsilon)\} &\leq P_\nu(\mu_N(\hat{\tau}_W) \in \gamma_{\pi_2^k}) \\ &\leq \exp\{-N(\tilde{V}(\pi_1^k, \pi_2^k) - \tilde{V}(\pi_1^k) - \varepsilon)\}.\end{aligned}$$

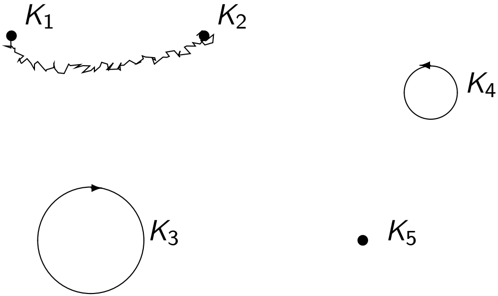
- Lower bound: construct a specific path.
- Upper bound: Use strong Markov property and the uniform LDP of μ_N.

Mixing of μ_N



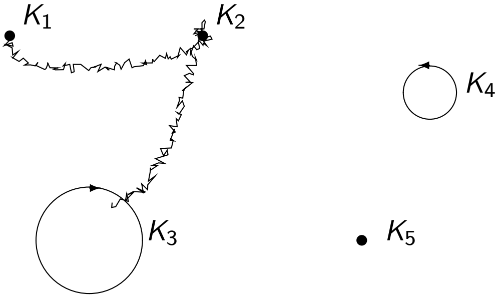
- Recall that $\wp_N(\xi) \sim \exp\{-NW(\xi)\}$.
- Suppose that $W(K_3) = \min_{i \neq 3} W(K_i)$.
- μ_N mixes well if it reaches a neighbourhood of K_3.

Mixing of μ_N



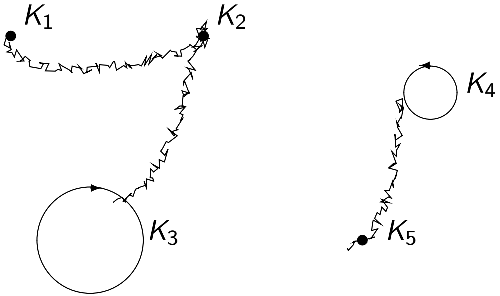
- Recall that $\wp_N(\xi) \sim \exp\{-NW(\xi)\}$.
- Suppose that $W(K_3) = \min_{i \neq 3} W(K_i)$.
- μ_N mixes well if it reaches a neighbourhood of K_3.

Mixing of μ_N



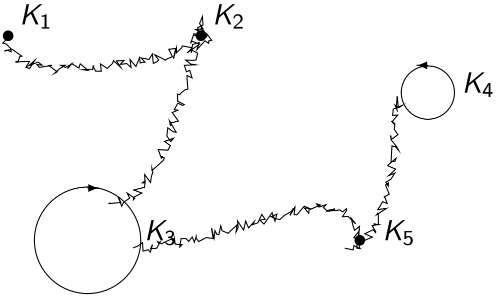
- Recall that $\wp_N(\xi) \sim \exp\{-NW(\xi)\}$.
- Suppose that $W(K_3) = \min_{i \neq 3} W(K_i)$.
- μ_N mixes well if it reaches a neighbourhood of K_3.

Mixing of μ_N



- Recall that $\wp_N(\xi) \sim \exp\{-NW(\xi)\}$.
- Suppose that $W(K_3) = \min_{i \neq 3} W(K_i)$.
- μ_N mixes well if it reaches a neighbourhood of K_3.

Mixing of μ_N



- Recall that $\wp_N(\xi) \sim \exp\{-NW(\xi)\}$.
- Suppose that $W(K_3) = \min_{i \neq 3} W(K_i)$.
- μ_N mixes well if it reaches a neighbourhood of K_3.

Mixing of μ_N

- Define

$$\Lambda = \min\{\tilde{V}(g) : g \in G(i), i \in L\} - \min\{\tilde{V}(g) : g \in G(i,j), i,j \in L, i \neq j\}.$$

Theorem

Given $\delta > 0$, there exist $\varepsilon > 0$ and $N_0 \geq 1$ such that for all $\nu \in M_1^N(\mathcal{Z})$ and $N \geq N_0$

$$|E_\nu(f(\mu_N(T))) - \langle f, \wp_N \rangle| \leq \|f\|_\infty \exp\{-\exp(N\varepsilon)\},$$

where $T = \exp\{N(\Lambda + \delta)\}$ and $f \in B(M_1(\mathcal{Z}))$.

- Proof via the estimates of large time behaviour.

Asymptotics of the second eigenvalue

- If μ_N is reversible (i.e. L^N is self-adjoint in $L^2(\wp_N)$), spectral decomposition of L^N tells us that

$$E_\nu f(\mu_N(t)) = \langle f, \wp_N \rangle + \sum_{k \geq 2} e^{-t\lambda_k^N}(f, u_k^N)u_k^N(\nu),$$

- Mixing time of μ_N is governed by λ_2^N.

Theorem

$$\lim_{N\to\infty} \frac{1}{N} \log \lambda_2^N = -\Lambda.$$

Convergence to a global minimum

- Fix $c > 0$. Start with $N_0 = \min\{n \in \mathbb{N} : \exp\{nc\} - 2 \geq 0\}$ particles.
- Let $t_{N_0} = 0$. Add a particle at times $t_N = \exp\{Nc\} - 2$, $N > N_0$, with a certain state.
- $\bar{\mu}$: the resulting process.

Convergence to a global minimum

- Fix $c > 0$. Start with $N_0 = \min\{n \in \mathbb{N} : \exp\{nc\} - 2 \geq 0\}$ particles.
- Let $t_{N_0} = 0$. Add a particle at times $t_N = \exp\{Nc\} - 2$, $N > N_0$, with a certain state.
- $\bar{\mu}$: the resulting process.
- Small c: particles are added too frequently; $\bar{\mu}$ could get trapped in a local minimum of the entropy function depending on $\bar{\mu}(0)$.

Convergence to a global minimum

- Fix $c > 0$. Start with $N_0 = \min\{n \in \mathbb{N} : \exp\{nc\} - 2 \geq 0\}$ particles.
- Let $t_{N_0} = 0$. Add a particle at times $t_N = \exp\{Nc\} - 2$, $N > N_0$, with a certain state.
- $\bar{\mu}$: the resulting process.
- Small c: particles are added too frequently; $\bar{\mu}$ could get trapped in a local minimum of the entropy function depending on $\bar{\mu}(0)$.
- Large c: sufficient time for exploration, $\bar{\mu}$ converges to a global minimum of the entropy function.

Convergence to a global minimum

- Fix $c > 0$. Start with $N_0 = \min\{n \in \mathbb{N} : \exp\{nc\} - 2 \geq 0\}$ particles.
- Let $t_{N_0} = 0$. Add a particle at times $t_N = \exp\{Nc\} - 2$, $N > N_0$, with a certain state.
- $\bar{\mu}$: the resulting process.
- Small c: particles are added too frequently; $\bar{\mu}$ could get trapped in a local minimum of the entropy function depending on $\bar{\mu}(0)$.
- Large c: sufficient time for exploration, $\bar{\mu}$ converges to a global minimum of the entropy function.

Theorem

For $c > c^*$ and any $\rho_1 > 0$,

$$P_{0,\nu}(\bar{\mu}(t) \in \gamma_{i_0}) \to 1$$

as $t \to \infty$, uniformly for all $\nu \in M_1^{N_0}(\mathcal{Z})$.

Conclusion

- Study of large time behaviour of finite state mean-field interacting particle systems
  - Exit time estimates. Decomposition into cycles.
  - Convergence of μ_N to its invariant measure.
- Scaling of $\lambda_2^N \sim \exp\{-N\Lambda\}$ when μ_N is reversible.
- Convergence of a controlled process to the global minimum of a certain entropy functional.

Reference: arXiv:1909.03805

Thank You