Large deviations of mean-field interacting particle systems in a fast varying environment

Sarath Yasodharan
Joint work with Rajesh Sundaresan

ECE Department, Indian Institute of Science

Bangalore Probability Seminar
03 May 2021

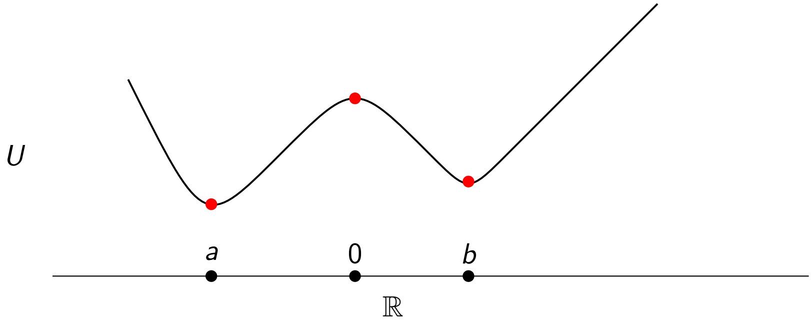
Background and motivation

- Metastability in dynamical systems perturbed by small-noise.

Background and motivation

- Metastability in dynamical systems perturbed by small-noise.
- Consider the SDE

$$dX_t^\varepsilon = -U'(X_t^\varepsilon)dt + \varepsilon dB_t$$

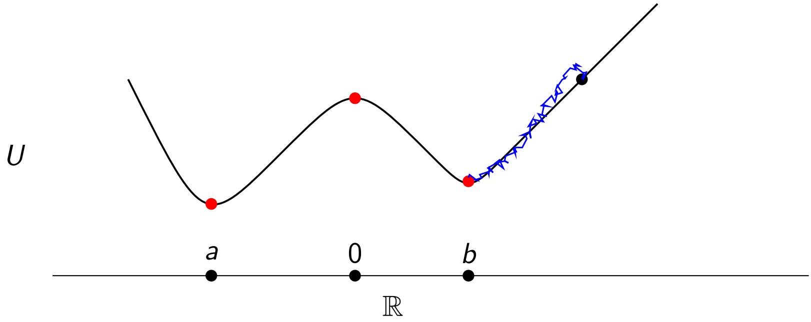


- Metastability: different behaviour at different time scales.
- Interested in quantifying probabilities of rare dynamical transitions.

Background and motivation

- Metastability in dynamical systems perturbed by small-noise.
- Consider the SDE

$$dX_t^{\varepsilon} = -U'(X_t^{\varepsilon})dt + \varepsilon dB_t$$

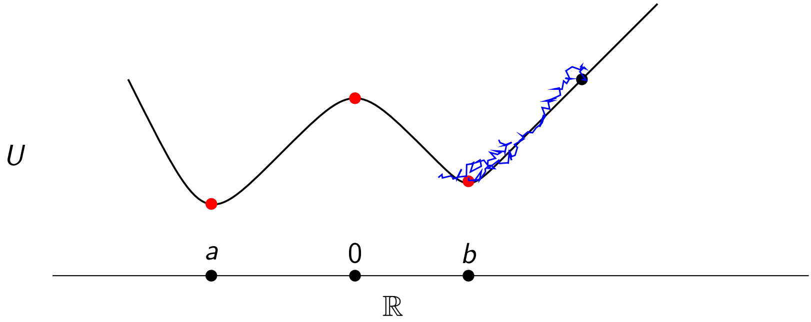


- Metastability: different behaviour at different time scales.
- Interested in quantifying probabilities of rare dynamical transitions.

Background and motivation

- Metastability in dynamical systems perturbed by small-noise.
- Consider the SDE

$$dX_t^\varepsilon = -U'(X_t^\varepsilon)dt + \varepsilon dB_t$$

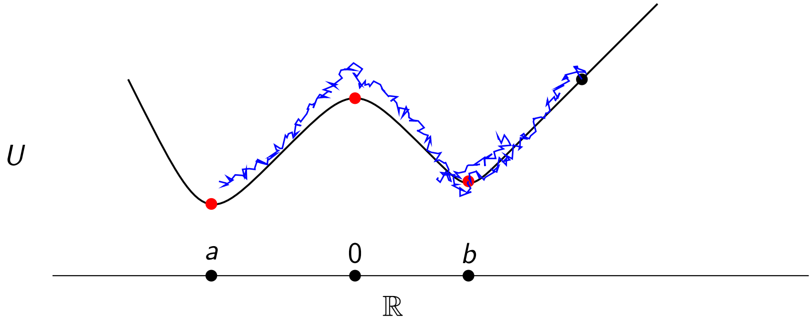


- Metastability: different behaviour at different time scales.
- Interested in quantifying probabilities of rare dynamical transitions.

Background and motivation

- Metastability in dynamical systems perturbed by small-noise.
- Consider the SDE

$$dX_t^\varepsilon = -U'(X_t^\varepsilon)dt + \varepsilon dB_t$$



- Metastability: different behaviour at different time scales.
- Interested in quantifying probabilities of rare dynamical transitions.

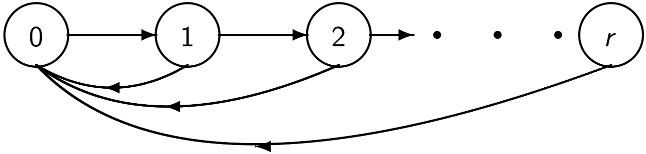
A wireless local area network

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

A wireless local area network

- N nodes accessing a common wireless medium.
- Interaction among nodes via the distributed MAC protocol.
- Channel state: idle, collision, successful transmission

A wireless local area network

- N nodes accessing a common wireless medium.
- Interaction among nodes via the distributed MAC protocol.
- Channel state: idle, collision, successful transmission
- State of a node represents aggressiveness of packet transmission.
- 

- Evolution of the state of a node:
  - Becomes less aggressive after a collision.
  - Moves to the most aggressive state after a successful packet transmission.

A sample path of the macroscopic behaviour

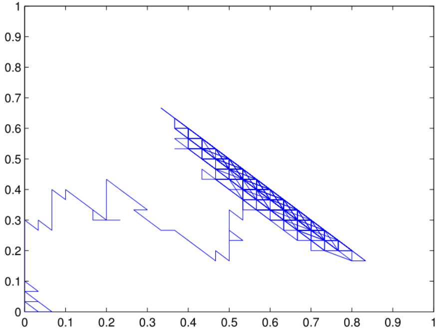


Figure: Evolution of states in a WiFi network under the MAC protocol

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

System model

- N particles and an environment.

System model

- N particles and an environment.
- At time t,
  - The state of the nth particle is $X_n^N(t) \in \mathcal{X}$;
  - The state of the environment is $Y_N(t) \in \mathcal{Y}$.

System model

- N particles and an environment.
- At time t,
  - The state of the nth particle is $X_n^N(t) \in \mathcal{X}$;
  - The state of the environment is $Y_N(t) \in \mathcal{Y}$.
- Certain allowed transitions.
  - Particles: a directed graph $(\mathcal{X}, \mathcal{E}_{\mathcal{X}})$;
  - Environment: a directed graph $(\mathcal{Y}, \mathcal{E}_{\mathcal{Y}})$.

System model

- N particles and an environment.
- At time t,
  - The state of the nth particle is $X_n^N(t) \in \mathcal{X}$;
  - The state of the environment is $Y_N(t) \in \mathcal{Y}$.
- Certain allowed transitions.
  - Particles: a directed graph $(\mathcal{X}, \mathcal{E}_{\mathcal{X}})$;
  - Environment: a directed graph $(\mathcal{Y}, \mathcal{E}_{\mathcal{Y}})$.
- Empirical measure of the system of particles at time t:

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

- We are given functions $\lambda_{x,x'}(\cdot, y)$, $(x, x') \in \mathcal{E}_{\mathcal{X}}$, $y \in \mathcal{Y}$ and $\gamma_{y,y'}(\cdot)$, $(y, y') \in \mathcal{E}_{\mathcal{Y}}$ on $M_1(\mathcal{X})$.

System model

- N particles and an environment.
- At time t,
  - The state of the nth particle is $X_n^N(t) \in \mathcal{X}$;
  - The state of the environment is $Y_N(t) \in \mathcal{Y}$.
- Certain allowed transitions.
  - Particles: a directed graph $(\mathcal{X}, \mathcal{E}_{\mathcal{X}})$;
  - Environment: a directed graph $(\mathcal{Y}, \mathcal{E}_{\mathcal{Y}})$.
- Empirical measure of the system of particles at time t:

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

- We are given functions $\lambda_{x,x'}(\cdot, y)$, $(x, x') \in \mathcal{E}_{\mathcal{X}}$, $y \in \mathcal{Y}$ and $\gamma_{y,y'}(\cdot)$, $(y, y') \in \mathcal{E}_{\mathcal{Y}}$ on $M_1(\mathcal{X})$.
- Markovian evolution at time t:
  - Particles: $x \to x'$ at rate $\lambda_{x,x'}(\mu_N(t), Y_N(t))$;

System model

- N particles and an environment.
- At time t,
  - The state of the nth particle is $X_n^N(t) \in \mathcal{X}$;
  - The state of the environment is $Y_N(t) \in \mathcal{Y}$.
- Certain allowed transitions.
  - Particles: a directed graph $(\mathcal{X}, \mathcal{E}_{\mathcal{X}})$;
  - Environment: a directed graph $(\mathcal{Y}, \mathcal{E}_{\mathcal{Y}})$.
- Empirical measure of the system of particles at time t:

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

- We are given functions $\lambda_{x,x'}(\cdot, y)$, $(x, x') \in \mathcal{E}_{\mathcal{X}}$, $y \in \mathcal{Y}$ and $\gamma_{y,y'}(\cdot)$, $(y, y') \in \mathcal{E}_{\mathcal{Y}}$ on $M_1(\mathcal{X})$.
- Markovian evolution at time t:
  - Particles: $x \to x'$ at rate $\lambda_{x,x'}(\mu_N(t), Y_N(t))$;
  - Environment: $y \to y'$ at rate $N\gamma_{y,y'}(\mu_N(t))$.

System model

- (μ_N, Y_N) is a Markov process with infinitesimal generator

$$f \mapsto \sum_{(x,x')\in\mathcal{E}_{\mathcal{X}}} N\xi(x)\lambda_{x,x'}(\xi,y)\left[f\left(\xi+\frac{\delta_{x'}}{N}-\frac{\delta_x}{N},y\right)-f(\xi,y)\right]$$
$$+N\sum_{y':(y,y')\in\mathcal{E}_{\mathcal{Y}}}(f(\xi,y')-f(\xi,y))\gamma_{y,y'}(\xi),$$

$(\xi, y) \in M_1^N(\mathcal{X}) \times \mathcal{Y}$.

System model

- (μ_N, Y_N) is a Markov process with infinitesimal generator

$$f \mapsto \sum_{(x,x')\in\mathcal{E}_{\mathcal{X}}} N\xi(x)\lambda_{x,x'}(\xi,y)\left[f\left(\xi+\frac{\delta_{x'}}{N}-\frac{\delta_x}{N},y\right)-f(\xi,y)\right] + N\sum_{y':(y,y')\in\mathcal{E}_{\mathcal{Y}}}(f(\xi,y')-f(\xi,y))\gamma_{y,y'}(\xi),$$

  $(\xi, y) \in M_1^N(\mathcal{X}) \times \mathcal{Y}$.
- A "fully coupled" two time scale process.

System model

- (μ_N, Y_N) is a Markov process with infinitesimal generator

$$f \mapsto \sum_{(x,x')\in\mathcal{E}_{\mathcal{X}}} N\xi(x)\lambda_{x,x'}(\xi,y)\left[f\left(\xi+\frac{\delta_{x'}}{N}-\frac{\delta_x}{N},y\right)-f(\xi,y)\right]$$
$$+N\sum_{y':(y,y')\in\mathcal{E}_{\mathcal{Y}}}(f(\xi,y')-f(\xi,y))\gamma_{y,y'}(\xi),$$

  $(\xi,y)\in M_1^N(\mathcal{X})\times\mathcal{Y}$.
- A "fully coupled" two time scale process.
- Assumptions:
  - The graphs $(\mathcal{X},\mathcal{E}_{\mathcal{X}})$ and $(\mathcal{Y},\mathcal{E}_{\mathcal{Y}})$ are irreducible.

System model

- (μ_N, Y_N) is a Markov process with infinitesimal generator

$$f \mapsto \sum_{(x,x')\in\mathcal{E}_{\mathcal{X}}} N\xi(x)\lambda_{x,x'}(\xi,y)\left[f\left(\xi+\frac{\delta_{x'}}{N}-\frac{\delta_x}{N},y\right)-f(\xi,y)\right] \\ +N\sum_{y':(y,y')\in\mathcal{E}_{\mathcal{Y}}}(f(\xi,y')-f(\xi,y))\gamma_{y,y'}(\xi),$$

$(\xi, y) \in M_1^N(\mathcal{X}) \times \mathcal{Y}$.

- A "fully coupled" two time scale process.
- Assumptions:
  - The graphs $(\mathcal{X}, \mathcal{E}_{\mathcal{X}})$ and $(\mathcal{Y}, \mathcal{E}_{\mathcal{Y}})$ are irreducible.
  - The functions $\lambda_{x,x'}(\cdot, y)$ are Lipschitz continuous and $\inf_\xi \lambda_{x,x'}(\xi, y) > 0$ for all $(x,x') \in \mathcal{E}_{\mathcal{X}}$ and $y \in \mathcal{Y}$.

System model

- (μ_N, Y_N) is a Markov process with infinitesimal generator

$$f \mapsto \sum_{(x,x')\in\mathcal{E}_\mathcal{X}} N\xi(x)\lambda_{x,x'}(\xi,y)\left[f\left(\xi+\frac{\delta_{x'}}{N}-\frac{\delta_x}{N},y\right)-f(\xi,y)\right]$$
$$+N\sum_{y':(y,y')\in\mathcal{E}_\mathcal{Y}}(f(\xi,y')-f(\xi,y))\gamma_{y,y'}(\xi),$$

  $(\xi,y)\in M_1^N(\mathcal{X})\times\mathcal{Y}$.
- A "fully coupled" two time scale process.
- Assumptions:
  - The graphs $(\mathcal{X},\mathcal{E}_\mathcal{X})$ and $(\mathcal{Y},\mathcal{E}_\mathcal{Y})$ are irreducible.
  - The functions $\lambda_{x,x'}(\cdot,y)$ are Lipschitz continuous and $\inf_\xi \lambda_{x,x'}(\xi,y)>0$ for all $(x,x')\in\mathcal{E}_\mathcal{X}$ and $y\in\mathcal{Y}$.
  - The functions $\gamma_{y,y'}(\cdot)$ are continuous and $\inf_\xi \gamma_{y,y'}(\xi)>0$ for all $(y,y')\in\mathcal{E}_\mathcal{Y}$.

The occupation measure process

- Fix a time duration $T > 0$.
- View μ_N as a random element of $D([0,T], M_1(\mathcal{X}))$.

The occupation measure process

- Fix a time duration $T > 0$.
- View μ_N as a random element of $D([0, T], M_1(\mathcal{X}))$.
- Consider the occupation measure of the fast environment:

$$\theta_N(t)(\cdot) := \int_0^t 1_{\{Y_N(s)\in\cdot\}} ds,\ 0 \leq t \leq T.$$

- θ_N is a random element of $D_\uparrow([0, T], M(\mathcal{Y}))$, the set of θ such that $\theta_t - \theta_s \in M(\mathcal{Y})$ and $\theta_t(\mathcal{Y}) = t$ for $0 \leq s \leq t \leq T$.

The occupation measure process

- Fix a time duration $T > 0$.
- View μ_N as a random element of $D([0, T], M_1(\mathcal{X}))$.
- Consider the occupation measure of the fast environment:

$$\theta_N(t)(\cdot) := \int_0^t 1_{\{Y_N(s) \in \cdot\}} ds, \; 0 \leq t \leq T.$$

- θ_N is a random element of $D_{\uparrow}([0, T], M(\mathcal{Y}))$, the set of θ such that $\theta_t - \theta_s \in M(\mathcal{Y})$ and $\theta_t(\mathcal{Y}) = t$ for $0 \leq s \leq t \leq T$.
- $\theta \in D_{\uparrow}([0, T], M(\mathcal{Y}))$ is also viewed as a measure on $[0, T] \times \mathcal{Y}$ and obeys the disintegration $\theta(dydt) = m_t(dy)dt$ where $m_t \in M_1(\mathcal{Y})$.

The occupation measure process

- Fix a time duration $T > 0$.
- View μ_N as a random element of $D([0,T], M_1(\mathcal{X}))$.
- Consider the occupation measure of the fast environment:

$$\theta_N(t)(\cdot) := \int_0^t 1_{\{Y_N(s)\in\cdot\}} ds,\ 0 \le t \le T.$$

- θ_N is a random element of $D_\uparrow([0,T], M(\mathcal{Y}))$, the set of θ such that $\theta_t - \theta_s \in M(\mathcal{Y})$ and $\theta_t(\mathcal{Y}) = t$ for $0 \le s \le t \le T$.
- $\theta \in D_\uparrow([0,T], M(\mathcal{Y}))$ is also viewed as a measure on $[0,T] \times \mathcal{Y}$ and obeys the disintegration $\theta(dydt) = m_t(dy)dt$ where $m_t \in M_1(\mathcal{Y})$.
- We consider the process (μ_N, θ_N) with sample paths in $D([0,T], M_1(\mathcal{X})) \times D_\uparrow([0,T], M(\mathcal{Y}))$.

The averaging principle

- Suppose we freeze $\mu_N(t)$ to be ξ.

The averaging principle

- Suppose we freeze $\mu_N(t)$ to be ξ. Then for large N,
  - The Y_N process would quickly equilibrate to π_ξ, the unique invariant probability measure of

$$L_\xi g(y) := \sum_{y':(y,y')\in\mathcal{E}_\mathcal{Y}} (g(y') - g(y))\gamma_{y,y'}(\xi), y \in \mathcal{Y}.$$

The averaging principle

- Suppose we freeze $\mu_N(t)$ to be ξ. Then for large N,
  - The Y_N process would quickly equilibrate to π_ξ, the unique invariant probability measure of

$$L_\xi g(y) := \sum_{y':(y,y')\in\mathcal{E}_\mathcal{Y}} (g(y') - g(y))\gamma_{y,y'}(\xi), y \in \mathcal{Y}.$$

  - For a particle, an (x, x') transition occurs at rate $\sum_{y\in\mathcal{Y}} \lambda_{x,x'}(\xi, y)\pi_\xi(y) =: \bar{\lambda}_{x,x'}(\xi, \pi_\xi)$.

The averaging principle

- Suppose we freeze $\mu_N(t)$ to be ξ. Then for large N,
  - The Y_N process would quickly equilibrate to π_ξ, the unique invariant probability measure of

$$L_\xi g(y) := \sum_{y':(y,y')\in\mathcal{E}_\mathcal{Y}} (g(y') - g(y))\gamma_{y,y'}(\xi), y \in \mathcal{Y}.$$

  - For a particle, an (x, x') transition occurs at rate $\sum_{y\in\mathcal{Y}} \lambda_{x,x'}(\xi, y)\pi_\xi(y) =: \bar{\lambda}_{x,x'}(\xi, \pi_\xi)$.

Theorem (Bordenave et al. 2009)

Suppose that $\mu_N(0) \to \nu$ in $M_1(\mathcal{X})$. Then μ_N converges in probability, in $D([0, T], M_1(\mathcal{X}))$, to the solution to the ODE

$$\dot{\mu}_t = \bar{\Lambda}^*_{\mu_t,\pi_{\mu_t}}\mu_t,\ 0 \le t \le T,\ \mu_0 = \nu.$$

where $\bar{\Lambda}_{\mu_t,\pi_{\mu_t}}(x, x') = \bar{\lambda}_{x,x'}(\mu_t, \pi_{\mu_t})$.

The averaging principle

- Suppose we freeze $\mu_N(t)$ to be ξ. Then for large N,
  - The Y_N process would quickly equilibrate to π_ξ, the unique invariant probability measure of

$$L_\xi g(y) := \sum_{y':(y,y')\in\mathcal{E}_\mathcal{Y}} (g(y') - g(y))\gamma_{y,y'}(\xi), y \in \mathcal{Y}.$$

  - For a particle, an (x, x') transition occurs at rate $\sum_{y\in\mathcal{Y}} \lambda_{x,x'}(\xi, y)\pi_\xi(y) =: \bar{\lambda}_{x,x'}(\xi, \pi_\xi)$.

Theorem (Bordenave et al. 2009)

Suppose that $\mu_N(0) \to \nu$ in $M_1(\mathcal{X})$. Then μ_N converges in probability, in $D([0, T], M_1(\mathcal{X}))$, to the solution to the ODE

$$\dot{\mu}_t = \bar{\Lambda}^*_{\mu_t,\pi_{\mu_t}}\mu_t,\ 0 \le t \le T,\ \mu_0 = \nu.$$

where $\bar{\Lambda}_{\mu_t,\pi_{\mu_t}}(x, x') = \bar{\lambda}_{x,x'}(\mu_t, \pi_{\mu_t})$.

- μ_N is a small random perturbation of the above ODE. We study fluctuations of (μ_N, θ_N).

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 $I_A = \inf_{x\in A} I(x)$.
- $\{X_N\}_{N\geq 1}$ is said to satisfy the large deviation principle (LDP) with rate function $I : S \to [0, +\infty]$ if
  - for each $M > 0$, $\{x \in S : I(x) \leq M\}$ is a compact subset of S;
  - for each open set $G \subset S$,

$$\liminf_{N\to\infty} \frac{1}{N} \log P(X_N \in G) \geq -\inf_{x\in G} I(x);$$

  - for each closed set $F \subset S$,

$$\limsup_{N\to\infty} \frac{1}{N} \log P(X_N \in F) \leq -\inf_{x\in F} I(x).$$

Large deviations: contraction principle

- S, T are metric spaces. $f : S \to T$ is continuous.
- $\{X_N\}$s are S-valued random variables. Define $Y_N = f(X_N)$.

Theorem (Contraction Principle)

If $\{X_N\}$ satisfies the LDP with rate function I, then $\{Y_N\}$ satisfies the LDP with rate function

$$J(y) = \inf_{x \in S : y = f(x)} I(x).$$

Large deviations: contraction principle

- S, T are metric spaces. $f : S \to T$ is continuous.
- $\{X_N\}$s are S-valued random variables. Define $Y_N = f(X_N)$.

Theorem (Contraction Principle)

If $\{X_N\}$ satisfies the LDP with rate function I, then $\{Y_N\}$ satisfies the LDP with rate function

$$J(y) = \inf_{x \in S : y = f(x)} I(x).$$

- Compactness of level sets:
  $\{y \in T : J(y) \leq M\} = f(\{x \in S : I(x) \leq M\})$.
- Upper and lower bounds:
  $P(Y_N \in A) = P(X_N \in f^{-1}(A))$.

Main result

Theorem

Suppose that $\{\mu_N(0)\}_{N\geq 1}$ *satisfies the LDP on* $M_1(\mathcal{X})$ *with rate function* I_0*. Then the sequence* $\{(\mu_N(t), \theta_N(t)), 0 \leq t \leq T\}_{N\geq 1}$ *satisfies the LDP on* $D([0,T], M_1(\mathcal{X})) \times D_{\uparrow}([0,T], M(\mathcal{Y}))$ *with rate function*

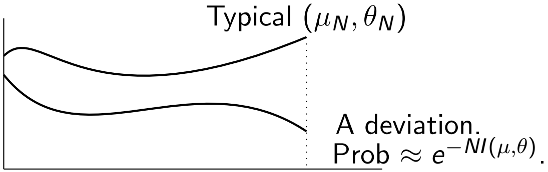
$$I(\mu, \theta) := I_0(\mu(0)) + J(\mu, \theta).$$

Main result

Theorem

Suppose that $\{\mu_N(0)\}_{N\geq 1}$ *satisfies the LDP on* $M_1(\mathcal{X})$ *with rate function* I_0*. Then the sequence* $\{(\mu_N(t), \theta_N(t)), 0 \leq t \leq T\}_{N\geq 1}$ *satisfies the LDP on* $D([0,T], M_1(\mathcal{X})) \times D_{\uparrow}([0,T], M(\mathcal{Y}))$ *with rate function*

$$I(\mu, \theta) := I_0(\mu(0)) + J(\mu, \theta).$$



The rate function J

$$
\begin{aligned}
J(\mu,\theta) := \int_{[0,T]} \Bigg\{ \sup_{\alpha\in\mathbb{R}^{|\mathcal{X}|}} \Bigg(& \langle \alpha, (\dot{\mu}_t - \bar{\Lambda}^*_{\mu_t,m_t}\mu_t)\rangle \\
& - \sum_{(x,x')\in\mathcal{E}_\mathcal{X}} \tau(\alpha(x')-\alpha(x))\bar{\lambda}_{x,x'}(\mu_t,m_t)\mu_t(x)\Bigg) \\
+ \sup_{g\in\mathbb{R}^{|\mathcal{Y}|}} \sum_{y\in\mathcal{Y}} \Bigg(& -L_{\mu_t} g(y) \\
& - \sum_{y':(y,y')\in\mathcal{E}_\mathcal{Y}} \tau(g(y')-g(y))\gamma_{y,y'}(\mu_t)\Bigg) m_t(y) \Bigg\}\, dt
\end{aligned}
$$

whenever the mapping $[0,T] \ni t \mapsto \mu_t \in M_1(\mathcal{X})$ is absolutely continuous, where $\theta(dtdy) = m_t(dy)dt$, and $J(\mu,\theta) = +\infty$ otherwise.

- $\tau(u) = e^u - u - 1, u \in \mathbb{R}$.

Some remarks about the rate function

- $J(\mu, \theta) \geq 0$ with equality iff (μ, θ) satisfies the mean-field limit.
- Two parts. The mean-field part (slow component) and occupation measure part (fast component).
  - For the slow component, the average of the fast variable appears.
  - For the fast component, the slow variable is frozen.
- For occupation measure of Markov processes, the canonical form of the rate function is $\int_{[0,T]} \sup_{h>0} \sum_{\mathcal{Y}} -\frac{L_{\mu_t} h(y)}{h(y)} m_t(y) dt$ (Donsker and Varadhan, 1973). This can be obtained by taking $h = e^g$.

Large deviations of μ_N

Corollary

$\{\mu_N\}$ satisfies the LDP on $D([0,T], M_1(\mathcal{X}))$ with rate function

$$\mu \mapsto I_0(\mu_0) + \inf_\theta J(\mu, \theta).$$

Large deviations of μ_N

Corollary

$\{\mu_N\}$ satisfies the LDP on $D([0,T], M_1(\mathcal{X}))$ with rate function

$$\mu \mapsto I_0(\mu_0) + \inf_\theta J(\mu, \theta).$$

- Follows from contraction principle since the mapping $(\mu, \theta) \mapsto \mu$ is continuous.

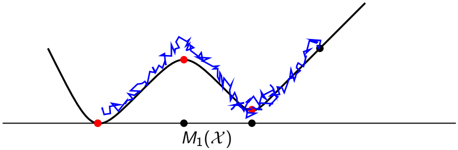
Large deviations of μ_N

Corollary

$\{\mu_N\}$ satisfies the LDP on $D([0,T], M_1(\mathcal{X}))$ with rate function

$$\mu \mapsto I_0(\mu_0) + \inf_\theta J(\mu, \theta).$$

- Follows from contraction principle since the mapping $(\mu, \theta) \mapsto \mu$ is continuous.
- Can quantify metastable transitions, mean exit time from a domain etc.



Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).

Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).
  - Weak convergence: Budhiraja et al. (2018).

Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).
  - Weak convergence: Budhiraja et al. (2018).
  - Semigroup: Kumar and Popovic (2017), Kraaij and Schlottke (2020).

Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).
  - Weak convergence: Budhiraja et al. (2018).
  - Semigroup: Kumar and Popovic (2017), Kraaij and Schlottke (2020).
- We use the method of stochastic exponentials (Pulahskii 2016, 1994).
  - Show exponential tightness. This gives a subsequential LDP.

Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).
  - Weak convergence: Budhiraja et al. (2018).
  - Semigroup: Kumar and Popovic (2017), Kraaij and Schlottke (2020).
- We use the method of stochastic exponentials (Pulahskii 2016, 1994).
  - Show exponential tightness. This gives a subsequential LDP.
  - Get a condition for any subsequential rate function (in terms of an exponential martingale).

Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).
  - Weak convergence: Budhiraja et al. (2018).
  - Semigroup: Kumar and Popovic (2017), Kraaij and Schlottke (2020).
- We use the method of stochastic exponentials (Pulahskii 2016, 1994).
  - Show exponential tightness. This gives a subsequential LDP.
  - Get a condition for any subsequential rate function (in terms of an exponential martingale).
  - Identify the subsequential rate function on "nice" elements of the space.

Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).
  - Weak convergence: Budhiraja et al. (2018).
  - Semigroup: Kumar and Popovic (2017), Kraaij and Schlottke (2020).
- We use the method of stochastic exponentials (Pulahskii 2016, 1994).
  - Show exponential tightness. This gives a subsequential LDP.
  - Get a condition for any subsequential rate function (in terms of an exponential martingale).
  - Identify the subsequential rate function on "nice" elements of the space.
  - Extend to the whole space using suitable approximations.

Outline of the proof

- Usual techniques:
  - Discretisation and change of measure: Freidlin and Wentzell (1984), Liptser (1996), Veretennikov (1999), Dawson and Gärtner (1987), Léonard (1995).
  - Weak convergence: Budhiraja et al. (2018).
  - Semigroup: Kumar and Popovic (2017), Kraaij and Schlottke (2020).
- We use the method of stochastic exponentials (Pulahskii 2016, 1994).
  - Show exponential tightness. This gives a subsequential LDP.
  - Get a condition for any subsequential rate function (in terms of an exponential martingale).
  - Identify the subsequential rate function on "nice" elements of the space.
  - Extend to the whole space using suitable approximations.
  - Unique identification any subsequential rate function (regardless of the subsequence) implies the LDP.
- Also used in the context of invariant measure LDP, Borkar and Sundaresan (2012).

Exponential tightness

Theorem

The sequence $\{(\mu_N(t), \theta_N(t)), t \in [0, T]\}_{N \geq 1}$ *is exponentially tight in* $D([0, T], M_1(\mathcal{X})) \times D_{\uparrow}([0, T], M(\mathcal{Y}))$*, i.e., given any* $M > 0$*, there exists a compact set* K_M *such that*

$$\limsup_{N \to \infty} \frac{1}{N} \log P\left(\{(\mu_N(t), \theta_N(t)), 0 \leq t \leq T\} \notin K_M\right) \leq -M.$$

Exponential tightness

Theorem

The sequence $\{(\mu_N(t), \theta_N(t)), t \in [0, T]\}_{N \geq 1}$ *is exponentially tight in* $D([0, T], M_1(\mathcal{X})) \times D_{\uparrow}([0, T], M(\mathcal{Y}))$*, i.e., given any* $M > 0$*, there exists a compact set* K_M *such that*

$$\limsup_{N\to\infty} \frac{1}{N} \log P\left(\{(\mu_N(t), \theta_N(t)), 0 \leq t \leq T\} \notin K_M\right) \leq -M.$$

For $\beta > 0$ and $\alpha \in \mathbb{R}^{|\mathcal{X}|}$, with $X_{N,t} = \langle \alpha, \mu_N(t) \rangle$,

$$\exp\Bigg\{ N\Bigg(\beta X_{N,t} - \beta X_{N,0} - \beta \int_0^t \Phi_{Y_{N,s}} f(\mu_{N,s}) ds - \int_0^t \sum_{(x,x')} \tau(\beta(\alpha(x') - \alpha(x))) \lambda_{x,x'}(\mu_{N,s}, Y_{N,s}) \mu_{N,s}(x) ds \Bigg) \Bigg\}, t \geq 0,$$

is an exponential martingale. Use Doob's inequality and a condition for exponential tightness in $D([0, T], \mathbb{R})$ (Puhalskii, 1994).

An equation for the subsequential rate function

- Let $\{(\mu_{N_k}, \theta_{N_k})\}_{k\geq 1}$ be a subsequence that satisfies the LDP with rate function $\tilde{I}$.

An equation for the subsequential rate function

- Let $\{(\mu_{N_k}, \theta_{N_k})\}_{k\geq 1}$ be a subsequence that satisfies the LDP with rate function $\tilde{I}$.
- Let $\alpha : [0, T] \times M_1(\mathcal{X}) \to \mathbb{R}^{|\mathcal{X}|}$ and $g : [0, T] \times M_1(\mathcal{X}) \times \mathcal{Y} \to \mathbb{R}$ be bounded measurable, and continuous on $M_1(\mathcal{X})$.

An equation for the subsequential rate function

- Let $\{(\mu_{N_k}, \theta_{N_k})\}_{k\geq 1}$ be a subsequence that satisfies the LDP with rate function $\tilde{I}$.
- Let $\alpha : [0,T] \times M_1(\mathcal{X}) \to \mathbb{R}^{|\mathcal{X}|}$ and $g : [0,T] \times M_1(\mathcal{X}) \times \mathcal{Y} \to \mathbb{R}$ be bounded measurable, and continuous on $M_1(\mathcal{X})$.
- Define $U_t^{\alpha,g}(\mu,\theta)$ by

$$\int_{[0,t]} \Big\{ \langle \alpha_s(\mu_s), \dot{\mu}_s - \bar{\Lambda}^*_{\mu_s,m_s}\mu_s \rangle - \sum_{(x,x')} \tau(\alpha_s(\mu_s)(x') - \alpha_s(\mu_s)(x))\bar{\lambda}_{x,x'}(\mu_s,m_s)\mu_s(x) + \sum_y \Big(-L_{\mu_s} g_s(\mu_s,\cdot)(y) - \sum_{y:(y,y')\in\mathcal{E}_\mathcal{Y}} \tau(g_s(\mu_s,y') - g_s(\mu_s,y))\gamma_{y,y'}(\mu_s) \Big) m_s(y) \Big\} ds.$$

An equation for the subsequential rate function

- We can show that, for each α and g,

$$\sup_{(\mu,\theta)\in\Gamma} \left(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)\right) = 0, \tag{1}$$

where Γ is the set of (μ,θ) such that $t \mapsto \mu_t$ absolutely continuous.

An equation for the subsequential rate function

- We can show that, for each α and g,

$$\sup_{(\mu,\theta)\in\Gamma} (U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0, \qquad (1)$$

  where Γ is the set of (μ,θ) such that $t \mapsto \mu_t$ absolutely continuous.
- On one hand, for a smaller class of α and g,

$$E\exp\{NU_T^{\alpha,g}(\mu_N,\theta_N) + V_T^g(\mu_N,Y_N)\} = 1,$$

  where V_T^g is $O(1)$ a.s.

An equation for the subsequential rate function

- We can show that, for each α and g,

$$\sup_{(\mu,\theta)\in\Gamma} (U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0, \qquad (1)$$

  where Γ is the set of (μ,θ) such that $t \mapsto \mu_t$ absolutely continuous.

- On one hand, for a smaller class of α and g,

$$E\exp\{NU_T^{\alpha,g}(\mu_N,\theta_N) + V_T^g(\mu_N, Y_N)\} = 1,$$

  where V_T^g is $O(1)$ a.s.

- On the other hand, Varadhan's lemma tells us that

$$\lim_{k\to\infty} \frac{1}{N_k} \log E\exp\{N_k U_T^{\alpha,g}(\mu_{N_k},\theta_{N_k}) + V_T^g(\mu_{N_k}, Y_{N_k})\}$$
$$= \sup_{(\mu,\theta)} (U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta))$$

An equation for the subsequential rate function

- We can show that, for each α and g,

$$\sup_{(\mu,\theta)\in\Gamma} (U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0, \tag{1}$$

  where Γ is the set of (μ,θ) such that $t \mapsto \mu_t$ absolutely continuous.
- On one hand, for a smaller class of α and g,

$$E \exp\{NU_T^{\alpha,g}(\mu_N,\theta_N) + V_T^g(\mu_N, Y_N)\} = 1,$$

  where V_T^g is $O(1)$ a.s.
- On the other hand, Varadhan's lemma tells us that

$$\lim_{k\to\infty} \frac{1}{N_k} \log E \exp\{N_k U_T^{\alpha,g}(\mu_{N_k},\theta_{N_k}) + V_T^g(\mu_{N_k}, Y_{N_k})\}$$
$$= \sup_{(\mu,\theta)} (U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta))$$

  This can be extended to (1).

An equation for the subsequential rate function

- We can show that, for each α and g,

$$\sup_{(\mu,\theta)\in\Gamma}(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0, \tag{1}$$

  where Γ is the set of (μ,θ) such that $t \mapsto \mu_t$ absolutely continuous.
- On one hand, for a smaller class of α and g,

$$E\exp\{NU_T^{\alpha,g}(\mu_N,\theta_N) + V_T^g(\mu_N,Y_N)\} = 1,$$

  where V_T^g is $O(1)$ a.s.
- On the other hand, Varadhan's lemma tells us that

$$\lim_{k\to\infty}\frac{1}{N_k}\log E\exp\{N_kU_T^{\alpha,g}(\mu_{N_k},\theta_{N_k}) + V_T^g(\mu_{N_k},Y_{N_k})\}$$
$$= \sup_{(\mu,\theta)}(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta))$$

  This can be extended to (1).
- Moreover, the supremum in (1) is attained.

A candidate rate function

- Recall that $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0$.

A candidate rate function

- Recall that $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0$.
- A natural candidate for the rate function

$$I^*(\mu,\theta) = \sup_{\alpha,g} U_T^{\alpha,g}(\mu,\theta).$$

A candidate rate function

- Recall that $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0$.
- A natural candidate for the rate function

$$I^*(\mu,\theta) = \sup_{\alpha,g} U_T^{\alpha,g}(\mu,\theta).$$

- It can be shown that $I^* = J$.

A candidate rate function

- Recall that $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0$.
- A natural candidate for the rate function

$$I^*(\mu,\theta) = \sup_{\alpha,g} U_T^{\alpha,g}(\mu,\theta).$$

- It can be shown that $I^* = J$.
- Note that $\tilde{I} \geq I^*$ on Γ. Outside Γ, I^* can be shown to be $+\infty$.

A candidate rate function

- Recall that $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\alpha,g}(\mu,\theta) - \tilde{I}(\mu,\theta)) = 0$.
- A natural candidate for the rate function

$$I^*(\mu,\theta) = \sup_{\alpha,g} U_T^{\alpha,g}(\mu,\theta).$$

- It can be shown that $I^* = J$.
- Note that $\tilde{I} \geq I^*$ on Γ. Outside Γ, I^* can be shown to be $+\infty$.
- Goal: show that $\tilde{I} \leq I^*$ whenever $I^* < +\infty$. Once this is established, the LDP follows.

Identification of $\tilde{I}$ on "nice" elements

- Suppose $(\hat{\mu}, \hat{\theta})$ is such that $I^*(\hat{\mu}, \hat{\theta}) < +\infty$, and
  - $\inf_{t \in [0,T]} \min_{x \in \mathcal{X}} \hat{\mu}_t(x) > 0$,
  - the mapping $[0, T] \ni t \mapsto \hat{\mu}_t \in M_1(\mathcal{X})$ is Lipschitz continuous,
  - $\inf_{t \in [0,T]} \min_{y \in \mathcal{Y}} \hat{m}_t(y) > 0$ where $\hat{\theta}(dydt) = \hat{m}_t(dy)dt$.

Identification of $\tilde{I}$ on "nice" elements

- Suppose $(\hat{\mu}, \hat{\theta})$ is such that $I^*(\hat{\mu}, \hat{\theta}) < +\infty$, and
  - $\inf_{t \in [0,T]} \min_{x \in \mathcal{X}} \hat{\mu}_t(x) > 0$,
  - the mapping $[0, T] \ni t \mapsto \hat{\mu}_t \in M_1(\mathcal{X})$ is Lipschitz continuous,
  - $\inf_{t \in [0,T]} \min_{y \in \mathcal{Y}} \hat{m}_t(y) > 0$ where $\hat{\theta}(dydt) = \hat{m}_t(dy)dt$.
- Then, there exists $(\hat{\alpha}, \hat{g})$ that attains $\sup_{\alpha, g} U_T^{\alpha, g}(\hat{\mu}, \hat{\theta})$.

Identification of $\tilde{I}$ on "nice" elements

- Suppose $(\hat{\mu}, \hat{\theta})$ is such that $I^*(\hat{\mu}, \hat{\theta}) < +\infty$, and
  - $\inf_{t\in[0,T]} \min_{x\in\mathcal{X}} \hat{\mu}_t(x) > 0$,
  - the mapping $[0, T] \ni t \mapsto \hat{\mu}_t \in M_1(\mathcal{X})$ is Lipschitz continuous,
  - $\inf_{t\in[0,T]} \min_{y\in\mathcal{Y}} \hat{m}_t(y) > 0$ where $\hat{\theta}(dydt) = \hat{m}_t(dy)dt$.
- Then, there exists $(\hat{\alpha}, \hat{g})$ that attains $\sup_{\alpha,g} U_T^{\alpha,g}(\hat{\mu}, \hat{\theta})$.
- With this $(\hat{\alpha}, \hat{g})$, get $(\tilde{\mu}, \tilde{\theta})$ that attains the supremum in $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\hat{\alpha},\hat{g}}(\mu, \theta) - \tilde{I}(\mu, \theta)) = 0$.
- Hence, $I^*(\tilde{\mu}, \tilde{\theta}) \geq U_T^{\hat{\alpha},\hat{g}}(\tilde{\mu}, \tilde{\theta}) = \tilde{I}(\tilde{\mu}, \tilde{\theta})$.

Identification of $\tilde{I}$ on "nice" elements

- Suppose $(\hat{\mu}, \hat{\theta})$ is such that $I^*(\hat{\mu}, \hat{\theta}) < +\infty$, and
  - $\inf_{t\in[0,T]} \min_{x\in\mathcal{X}} \hat{\mu}_t(x) > 0$,
  - the mapping $[0, T] \ni t \mapsto \hat{\mu}_t \in M_1(\mathcal{X})$ is Lipschitz continuous,
  - $\inf_{t\in[0,T]} \min_{y\in\mathcal{Y}} \hat{m}_t(y) > 0$ where $\hat{\theta}(dydt) = \hat{m}_t(dy)dt$.
- Then, there exists $(\hat{\alpha}, \hat{g})$ that attains $\sup_{\alpha,g} U_T^{\alpha,g}(\hat{\mu}, \hat{\theta})$.
- With this $(\hat{\alpha}, \hat{g})$, get $(\tilde{\mu}, \tilde{\theta})$ that attains the supremum in $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\hat{\alpha},\hat{g}}(\mu, \theta) - \tilde{I}(\mu, \theta)) = 0$.
- Hence, $I^*(\tilde{\mu}, \tilde{\theta}) \geq U_T^{\hat{\alpha},\hat{g}}(\tilde{\mu}, \tilde{\theta}) = \tilde{I}(\tilde{\mu}, \tilde{\theta})$.
- Since $I^* \leq \tilde{I}$, we get $I^*(\tilde{\mu}, \tilde{\theta}) = \tilde{I}(\tilde{\mu}, \tilde{\theta})$.
- Show that $(\tilde{\mu}, \tilde{\theta}) = (\hat{\mu}, \hat{\theta})$.

Identification of $\tilde{I}$ on "nice" elements

- Suppose $(\hat{\mu}, \hat{\theta})$ is such that $I^*(\hat{\mu}, \hat{\theta}) < +\infty$, and
  - $\inf_{t\in[0,T]} \min_{x\in\mathcal{X}} \hat{\mu}_t(x) > 0$,
  - the mapping $[0, T] \ni t \mapsto \hat{\mu}_t \in M_1(\mathcal{X})$ is Lipschitz continuous,
  - $\inf_{t\in[0,T]} \min_{y\in\mathcal{Y}} \hat{m}_t(y) > 0$ where $\hat{\theta}(dydt) = \hat{m}_t(dy)dt$.
- Then, there exists $(\hat{\alpha}, \hat{g})$ that attains $\sup_{\alpha,g} U_T^{\alpha,g}(\hat{\mu}, \hat{\theta})$.
- With this $(\hat{\alpha}, \hat{g})$, get $(\tilde{\mu}, \tilde{\theta})$ that attains the supremum in $\sup_{(\mu,\theta)\in\Gamma}(U_T^{\hat{\alpha},\hat{g}}(\mu, \theta) - \tilde{I}(\mu, \theta)) = 0$.
- Hence, $I^*(\tilde{\mu}, \tilde{\theta}) \geq U_T^{\hat{\alpha},\hat{g}}(\tilde{\mu}, \tilde{\theta}) = \tilde{I}(\tilde{\mu}, \tilde{\theta})$.
- Since $I^* \leq \tilde{I}$, we get $I^*(\tilde{\mu}, \tilde{\theta}) = \tilde{I}(\tilde{\mu}, \tilde{\theta})$.
- Show that $(\tilde{\mu}, \tilde{\theta}) = (\hat{\mu}, \hat{\theta})$.
- It follows that $\tilde{I}(\hat{\mu}, \hat{\theta}) = I^*(\hat{\mu}, \hat{\theta})$.

Approximation procedure

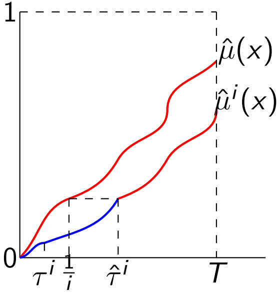
- For general elements $(\hat{\mu}, \hat{\theta})$, $(\hat{\alpha}, \hat{g})$ may not exist.

Approximation procedure

- For general elements $(\hat{\mu}, \hat{\theta})$, $(\hat{\alpha}, \hat{g})$ may not exist.
- Produce $(\hat{\mu}_i, \hat{\theta}_i)$ that are "nice", and satisfy
  - $(\hat{\mu}_i, \hat{\theta}_i) \to (\hat{\mu}, \hat{\theta})$ as $i \to \infty$,
  - $\tilde{I} = I^*$ on $(\hat{\mu}_i, \hat{\theta}_i)$ for all i,
  - $I^*(\hat{\mu}_i, \hat{\theta}_i) \to I^*(\hat{\mu}, \hat{\theta})$ as $i \to \infty$.
- It then follows that $\tilde{I} = I^*$ on $(\hat{\mu}, \hat{\theta})$.

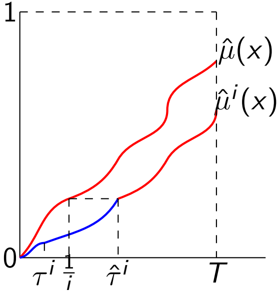
Approximation procedure

- For general elements $(\hat{\mu}, \hat{\theta})$, $(\hat{\alpha}, \hat{g})$ may not exist.
- Produce $(\hat{\mu}_i, \hat{\theta}_i)$ that are "nice", and satisfy
  - $(\hat{\mu}_i, \hat{\theta}_i) \to (\hat{\mu}, \hat{\theta})$ as $i \to \infty$,
  - $\tilde{I} = I^*$ on $(\hat{\mu}_i, \hat{\theta}_i)$ for all i,
  - $I^*(\hat{\mu}_i, \hat{\theta}_i) \to I^*(\hat{\mu}, \hat{\theta})$ as $i \to \infty$.
- It then follows that $\tilde{I} = I^*$ on $(\hat{\mu}, \hat{\theta})$.
- Relaxation of $\inf_{t\in[0,T]} \min_{x\in\mathcal{X}} \hat{\mu}_t(x) > 0$:



Approximation procedure

- For general elements $(\hat{\mu}, \hat{\theta})$, $(\hat{\alpha}, \hat{g})$ may not exist.
- Produce $(\hat{\mu}_i, \hat{\theta}_i)$ that are "nice", and satisfy
  - $(\hat{\mu}_i, \hat{\theta}_i) \to (\hat{\mu}, \hat{\theta})$ as $i \to \infty$,
  - $\tilde{I} = I^*$ on $(\hat{\mu}_i, \hat{\theta}_i)$ for all i,
  - $I^*(\hat{\mu}_i, \hat{\theta}_i) \to I^*(\hat{\mu}, \hat{\theta})$ as $i \to \infty$.
- It then follows that $\tilde{I} = I^*$ on $(\hat{\mu}, \hat{\theta})$.
- Relaxation of $\inf_{t \in [0,T]} \min_{x \in \mathcal{X}} \hat{\mu}_t(x) > 0$:



- Other conditions are relaxed using suitable approximations. We finally get $\tilde{I} = I^*$ for all elements.

Summary and future directions

- We show the LDP for (μ_N, θ_N).

Summary and future directions

- We show the LDP for (μ_N, θ_N).
- Future directions
  - Countable state space for the particles and the environment
  - Diminishing rates

Summary and future directions

- We show the LDP for (μ_N, θ_N).
- Future directions
  - Countable state space for the particles and the environment
  - Diminishing rates

Reference: arXiv: 2008.06855

Thank You