Compactness and large deviations

Chiranjib Mukherjee
Courant Institute, NYU

Based on Joint projects with S.R. S. Varadhan (New York), Erwin Bolthausen (Zurich) and Wolfgang König (Berlin)

Eindhoven, March, 2016

A weak LDP for occupation measures

Rate function is Legendre dual of principle eigenvalue

- We have a d-dimensional Brownian motion $(\beta_t)_t$, $d \geq 2$.

A weak LDP for occupation measures

Rate function is Legendre dual of principle eigenvalue

- We have a d-dimensional Brownian motion $(\beta_t)_t$, $d \geq 2$.
- For a continuous function V in a bounded domain B, expect integrals to grow exponentially:

$$\mathbb{E}\left[\exp\left\{ \underbrace{\int_0^t V(\beta_s)\mathrm{d}s}_{=t\int V(x)L_t(\mathrm{d}x)} \right\}\right] \sim \exp\{t\,\underbrace{\lambda(V)}_{>0}\}$$

A weak LDP for occupation measures

Rate function is Legendre dual of principle eigenvalue

- We have a d-dimensional Brownian motion $(\beta_t)_t$, $d \geq 2$.
- For a continuous function V in a bounded domain B, expect integrals to grow exponentially:

$$\mathbb{E}\Big[\exp\Big\{\underbrace{\int_0^t V(\beta_s)\mathrm{d}s}_{=t\int V(x)L_t(\mathrm{d}x)}\Big\}\Big] \sim \exp\{t\underbrace{\lambda(V)}_{>0}\}$$

- Important object: $L_t = \frac{1}{t}\int_0^t \delta_{\beta_s}\mathrm{d}s$ (time spent on Borel sets).

A weak LDP for occupation measures

Rate function is Legendre dual of principle eigenvalue

- We have a d-dimensional Brownian motion $(\beta_t)_t$, $d \geq 2$.
- For a continuous function V in a bounded domain B, expect integrals to grow exponentially:

$$\mathbb{E}\Big[\exp\Big\{ \underbrace{\int_0^t V(\beta_s)\mathrm{d}s}_{=t\int V(x)L_t(\mathrm{d}x)} \Big\}\Big] \sim \exp\{t\underbrace{\lambda(V)}_{>0}\}$$

- Important object: $L_t = \frac{1}{t}\int_0^t \delta_{\beta_s}\mathrm{d}s$ (time spent on Borel sets).
- Equivalently: then, exponential decay of probabilities:

$$\mathbb{P}(L_t \simeq f^2\mathrm{d}x \text{ on } B) \sim \exp\Big\{-tI(f^2)\Big\} \quad \|f\|_2 = 1, f \in H_0^1(B)$$

$I(f^2) = \frac{1}{2}||\nabla f||_2^2$ Donsker-Varadhan rate function.

Theory suffers from the lack of full LDP

- Crucial: theory needs compact state spaces.

Theory suffers from the lack of full LDP

- Crucial: theory needs compact state spaces.
- Workaround: For non-compact spaces, sometimes, one point compactification of $\mathbb{R}^d$ is enough (or, wrap your BM around a torus).

Theory suffers from the lack of full LDP

- Crucial: theory needs compact state spaces.
- Workaround: For non-compact spaces, sometimes, one point compactification of $\mathbb{R}^d$ is enough (or, wrap your BM around a torus).
- Success depends on problem and limited scope.

Theory suffers from the lack of full LDP

- Crucial: theory needs compact state spaces.
- Workaround: For non-compact spaces, sometimes, one point compactification of $\mathbb{R}^d$ is enough (or, wrap your BM around a torus).
- Success depends on problem and limited scope.
- Here is a problem where it does not work.

The mean-field polaron problem

Physical problems often need statements on the whole space

- Statistical mechanics:

The mean-field polaron problem

Physical problems often need statements on the whole space

- Statistical mechanics: For $V \in C_0(\mathbb{R}^d)$ (think of $V(x) = \frac{1}{|x|}$ in $d = 3$),

The mean-field polaron problem

Physical problems often need statements on the whole space

- Statistical mechanics: For $V \in C_0(\mathbb{R}^d)$ (think of $V(x) = \frac{1}{|x|}$ in $d = 3$), still exponential growth??

$$Z_t = \mathbb{E}\Big[\exp\Big\{\underbrace{\frac{1}{t}\int_0^t\int_0^t \mathrm{d}r\mathrm{d}s\, V(\beta_s - \beta_r)}_{=t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(x-y)L_t(\mathrm{d}x)L_t(\mathrm{d}y)}\Big\}\Big]$$

The mean-field polaron problem

Physical problems often need statements on the whole space

- Statistical mechanics: For $V \in C_0(\mathbb{R}^d)$ (think of $V(x) = \frac{1}{|x|}$ in $d = 3$), still exponential growth??

$$Z_t = \mathbb{E}\Big[\exp\Big\{\underbrace{\frac{1}{t}\int_0^t\int_0^t \mathrm{d}r\mathrm{d}s\ V(\beta_s - \beta_r)}_{=t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(x-y)L_t(\mathrm{d}x)L_t(\mathrm{d}y)}\Big\}\Big] \sim e^{t\rho}?\ \ \rho > 0.$$

The mean-field polaron problem

Physical problems often need statements on the whole space

- Statistical mechanics: For $V \in C_0(\mathbb{R}^d)$ (think of $V(x) = \frac{1}{|x|}$ in $d = 3$), still exponential growth??

$$Z_t = \mathbb{E}\left[\exp\left\{\underbrace{\frac{1}{t}\int_0^t\int_0^t \mathrm{d}r\mathrm{d}s\ V(\beta_s - \beta_r)}_{=t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(x-y)L_t(\mathrm{d}x)L_t(\mathrm{d}y)}\right\}\right] \sim e^{t\rho}?\ \ \rho > 0.$$

- Even more: Can we say something about the measures

$$d\mathbb{Q}_t = \frac{1}{Z_t}\exp\left\{\dots\right\}d\mathbb{P}?$$

- Starting point: Need full LDP for L_t in $\mathcal{M}_1(\mathbb{R}^d)$.

The mean-field polaron problem

Physical problems often need statements on the whole space

- Statistical mechanics: For $V \in C_0(\mathbb{R}^d)$ (think of $V(x) = \frac{1}{|x|}$ in $d = 3$), still exponential growth??

$$Z_t = \mathbb{E}\Bigg[\exp\Bigg\{\underbrace{\frac{1}{t}\int_0^t\int_0^t \mathrm{d}r\mathrm{d}s\ V(\beta_s - \beta_r)}_{=t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(x-y)L_t(\mathrm{d}x)L_t(\mathrm{d}y)}\Bigg\}\Bigg] \sim e^{t\rho}?\quad \rho > 0.$$

- Even more: Can we say something about the measures

$$d\mathbb{Q}_t = \frac{1}{Z_t}\exp\Big\{\dots\Big\} d\mathbb{P}?$$

- Starting point: Need full LDP for L_t in $\mathcal{M}_1(\mathbb{R}^d)$.
- No full LDP exists, and projection on torus does not save us.

The mean-field polaron problem

Physical problems often need statements on the whole space

- Statistical mechanics: For $V \in C_0(\mathbb{R}^d)$ (think of $V(x) = \frac{1}{|x|}$ in $d = 3$), still exponential growth??

$$Z_t = \mathbb{E}\left[\exp\left\{\underbrace{\frac{1}{t}\int_0^t\int_0^t \mathrm{d}r\mathrm{d}s\ V(\beta_s - \beta_r)}_{=t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(x-y)L_t(\mathrm{d}x)L_t(\mathrm{d}y)}\right\}\right] \sim e^{t\rho}?\quad \rho > 0.$$

- Even more: Can we say something about the measures

$$d\mathbb{Q}_t = \frac{1}{Z_t}\exp\left\{\ldots\right\}d\mathbb{P}?$$

- Starting point: Need full LDP for L_t in $\mathcal{M}_1(\mathbb{R}^d)$.
- No full LDP exists, and projection on torus does not save us. So, need a robust theory of large deviations via general compactification.

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?
- First start with $\mathcal{M}_1(\mathbb{R}^d)$. Not compact under the weak topology.

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?
- First start with $\mathcal{M}_1(\mathbb{R}^d)$. Not compact under the weak topology.
- Why?

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?
- First start with $\mathcal{M}_1(\mathbb{R}^d)$. Not compact under the weak topology.
- Why? Mass may escape and leak out or spread too flat.

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?
- First start with $\mathcal{M}_1(\mathbb{R}^d)$. Not compact under the weak topology.
- Why? Mass may escape and leak out or spread too flat.
- For any sequence $(\mu_n)_n$, locate regions with high accumulation of mass.

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?
- First start with $\mathcal{M}_1(\mathbb{R}^d)$. Not compact under the weak topology.
- Why? Mass may escape and leak out or spread too flat.
- For any sequence $(\mu_n)_n$, locate regions with high accumulation of mass. Recover lumps of masses one by one.
- Set

$$p_n(R) = \overbrace{\sup_{x\in\mathbb{R}^d} \mu_n(B_R + x)}^{\text{mass of biggest lump}} \leq 1$$

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?
- First start with $\mathcal{M}_1(\mathbb{R}^d)$. Not compact under the weak topology.
- Why? Mass may escape and leak out or spread too flat.
- For any sequence $(\mu_n)_n$, locate regions with high accumulation of mass. Recover lumps of masses one by one.
- Set

$$p_n(R) = \overbrace{\sup_{x\in\mathbb{R}^d} \mu_n(B_R + x)}^{\text{mass of biggest lump}} \leq 1$$

$$\longrightarrow p(R) \quad n \to \infty \text{ along a subsequence}$$

Probability measures are not compact

Need to identify regions where mass is accumulated

- What do we want to compactify?
- First start with $\mathcal{M}_1(\mathbb{R}^d)$. Not compact under the weak topology.
- Why? Mass may escape and leak out or spread too flat.
- For any sequence $(\mu_n)_n$, locate regions with high accumulation of mass. Recover lumps of masses one by one.
- Set

$$p_n(R) = \overbrace{\sup_{x\in\mathbb{R}^d} \mu_n(B_R + x)}^{\text{mass of biggest lump}} \leq 1$$

$$\longrightarrow p(R) \quad n \to \infty \text{ along a subsequence}$$

$$\longrightarrow p_1 \in [0,1] \quad R \uparrow \infty$$

Restrict measures on compact regions and shift, weak limit

Peel off the mass recovered. Now work with the leftover. Repeat

Restrict measures on compact regions and shift, weak limit

Peel off the mass recovered. Now work with the leftover. Repeat

- This means, decompose

$$\mu_n = \underbrace{\mu_n\big|_{B_R(x_n)}}_{\alpha_n} + \underbrace{\text{rest}}_{\beta_n} \quad \text{so that}$$

Restrict measures on compact regions and shift, weak limit

Peel off the mass recovered. Now work with the leftover. Repeat

- This means, decompose

$$\mu_n = \underbrace{\mu_n\big|_{B_R(x_n)}}_{\alpha_n} + \underbrace{\text{rest}}_{\beta_n} \quad \text{so that}$$

the shift $\alpha_n \star \delta_{x_n} \Rightarrow p_1\alpha_1$ weakly, along some subsequence.

Restrict measures on compact regions and shift, weak limit

Peel off the mass recovered. Now work with the leftover. Repeat

- This means, decompose

$$\mu_n = \underbrace{\mu_n\big|_{B_R(x_n)}}_{\alpha_n} + \underbrace{\text{rest}}_{\beta_n} \quad \text{so that}$$

the shift $\alpha_n \star \delta_{x_n} \Rightarrow p_1\alpha_1$ weakly, along some subsequence.

- We peel off α_n from μ_n and work with the leftover β_n.

Restrict measures on compact regions and shift, weak limit

Peel off the mass recovered. Now work with the leftover. Repeat

- This means, decompose

$$\mu_n = \underbrace{\mu_n\big|_{B_R(x_n)}}_{\alpha_n} + \underbrace{\mathrm{rest}}_{\beta_n} \quad \text{so that}$$

the shift $\alpha_n \star \delta_{x_n} \Rightarrow p_1 \alpha_1$ weakly, along some subsequence.

- We peel off α_n from μ_n and work with the leftover β_n.
- Repeat the process for β_n,

Restrict measures on compact regions and shift, weak limit

Peel off the mass recovered. Now work with the leftover. Repeat

- This means, decompose

$$\mu_n = \underbrace{\mu_n\big|_{B_R(x_n)}}_{\alpha_n} + \underbrace{\mathrm{rest}}_{\beta_n} \quad \text{so that}$$

the shift $\alpha_n \star \delta_{x_n} \Rightarrow p_1\alpha_1$ weakly, along some subsequence.

- We peel off α_n from μ_n and work with the leftover β_n.
- Repeat the process for β_n, recover mass $p_2 \leq p_1$ so that $\beta_n|_{\text{a ball elsewhere}}$, suitably shifted, $\Rightarrow p_2\alpha_2$ along some further subsequence.

Restrict measures on compact regions and shift, weak limit

Peel off the mass recovered. Now work with the leftover. Repeat

- This means, decompose

$$\mu_n = \underbrace{\mu_n\big|_{B_R(x_n)}}_{\alpha_n} + \underbrace{\text{rest}}_{\beta_n} \quad \text{so that}$$

the shift $\alpha_n \star \delta_{x_n} \Rightarrow p_1 \alpha_1$ weakly, along some subsequence.

- We peel off α_n from μ_n and work with the leftover β_n.
- Repeat the process for β_n, recover mass $p_2 \leq p_1$ so that $\beta_n|_{\text{a ball elsewhere}}$, suitably shifted, $\Rightarrow p_2\alpha_2$ along some further subsequence.
- Continue recursively: $\{\mu_n\}_n$ concentrates on compact pieces of mass $\{p_j\}_j$, which are widely separated, while the rest of the mass $1 - \sum_j p_j$ dissipates. μ_n on these compact pieces, when suitably shifted, converges along subsequences.

Compactification

What is in the compactification? Pairs of (recovered masses, equiv. classes)

Compactification

What is in the compactification? Pairs of (recovered masses, equiv. classes)

- Quotient space of orbits

$$\widetilde{\mathcal{M}} = \mathcal{M}_1 / \sim \quad \text{with } \mu \sim \nu \text{ if } \nu = \mu \star \delta_x$$

Compactification

What is in the compactification? Pairs of (recovered masses, equiv. classes)

- Quotient space of orbits

$$\widetilde{\mathcal{M}} = \mathcal{M}_1 / \sim \quad \text{with } \mu \sim \nu \text{ if } \nu = \mu \star \delta_x$$

-

$$\widetilde{\mathcal{M}}_1 \hookrightarrow \mathbf{M}^\star = \underbrace{\Big\{ \{p_j, \widetilde{\alpha}_j\}_j : p_j \downarrow 0, \sum_j p_j \leq 1, \alpha_j \in \mathcal{M}_{\leq 1} \Big\}}_{\text{all pairs of (mass, equivalence class)}}$$

Compactification

What is in the compactification? Pairs of (recovered masses, equiv. classes)

- Quotient space of orbits

$$\widetilde{\mathcal{M}} = \mathcal{M}_1 / \sim \quad \text{with } \mu \sim \nu \text{ if } \nu = \mu \star \delta_x$$

- $$\widetilde{\mathcal{M}}_1 \hookrightarrow \mathbf{M}^\star = \underbrace{\left\{ \{p_j, \widetilde{\alpha}_j\}_j : p_j \downarrow 0, \sum_j p_j \leq 1, \alpha_j \in \mathcal{M}_{\leq 1} \right\}}_{\text{all pairs of (mass, equivalence class)}}$$

- Given any sequence $(\mu_n)_n$ in $\mathcal{M}_1(\mathbb{R}^d)$, pass to its equivalence class $\widetilde{\mu_n}$ in $\widetilde{\mathcal{M}}_1$. There is a subsequence which converges to some element in $\mathbf{M}^\star$.

Compactification

What is in the compactification? Pairs of (recovered masses, equiv. classes)

- Quotient space of orbits

$$\widetilde{\mathcal{M}} = \mathcal{M}_1 / \sim \quad \text{with } \mu \sim \nu \text{ if } \nu = \mu \star \delta_x$$

-

$$\widetilde{\mathcal{M}}_1 \hookrightarrow \mathbf{M}^\star = \underbrace{\left\{ \{p_j, \widetilde{\alpha}_j\}_j : p_j \downarrow 0, \sum_j p_j \le 1, \alpha_j \in \mathcal{M}_{\le 1} \right\}}_{\text{all pairs of (mass, equivalence class)}}$$

- Given any sequence $(\mu_n)_n$ in $\mathcal{M}_1(\mathbb{R}^d)$, pass to its equivalence class $\widetilde{\mu_n}$ in $\widetilde{\mathcal{M}}_1$. There is a subsequence which converges to some element in $\mathbf{M}^\star$. What does it mean?

Compactification

What is in the compactification? Pairs of (recovered masses, equiv. classes)

- Quotient space of orbits

$$\widetilde{\mathcal{M}} = \mathcal{M}_1/\sim \quad \text{with } \mu \sim \nu \text{ if } \nu = \mu \star \delta_x$$

-

$$\widetilde{\mathcal{M}}_1 \hookrightarrow \mathbf{M}^\star = \underbrace{\Big\{ \{p_j, \widetilde{\alpha}_j\}_j \colon p_j \downarrow 0, \sum_j p_j \leq 1, \alpha_j \in \mathcal{M}_{\leq 1} \Big\}}_{\text{all pairs of (mass, equivalence class)}}$$

- Given any sequence $(\mu_n)_n$ in $\mathcal{M}_1(\mathbb{R}^d)$, pass to its equivalence class $\widetilde{\mu_n}$ in $\widetilde{\mathcal{M}}_1$. There is a subsequence which converges to some element in $\mathbf{M}^\star$. What does it mean?
- Conclude: $\mathbf{M}^\star$ is the compactfication of $\widetilde{\mathcal{M}}_1$.

A key step: A new LDP in a compactified space

Turns out weak LDP for L_T is not good enough

- What can we do with this compact space $\mathbf{M}^\star$?

A key step: A new LDP in a compactified space

Turns out weak LDP for L_T is not good enough

- What can we do with this compact space $\mathbf{M}^\star$?
- Equivalence classes $\widetilde{L_t} \in \mathbf{M}^\star$. This is the sequence we want to work with.

A key step: A new LDP in a compactified space

Turns out weak LDP for L_T is not good enough

- What can we do with this compact space $\mathbf{M}^\star$?
- Equivalence classes $\widetilde{L}_t \in \mathbf{M}^\star$. This is the sequence we want to work with.
- How does $\mathbb{P}\big(\widetilde{L}_t \in \mathbf{C}\big)$ behave? Have upper bound as $t \to \infty$ for all $\mathbf{C} \subset \mathbf{M}^\star \mathrm{closed}$?

A key step: A new LDP in a compactified space

Turns out weak LDP for L_T is not good enough

- What can we do with this compact space $\mathbf{M}^\star$?
- Equivalence classes $\widetilde{L}_t \in \mathbf{M}^\star$. This is the sequence we want to work with.
- How does $\mathbb{P}(\widetilde{L}_t \in \mathbf{C})$ behave? Have upper bound as $t \to \infty$ for all $\mathbf{C} \subset \mathbf{M}^\star \text{closed}$?
- Yes! Compactness helps: Need to prove the upper bound only locally:

A key step: A new LDP in a compactified space

Turns out weak LDP for L_T is not good enough

- What can we do with this compact space $\mathbf{M}^\star$?
- Equivalence classes $\widetilde{L}_t \in \mathbf{M}^\star$. This is the sequence we want to work with.
- How does $\mathbb{P}(\widetilde{L}_t \in \mathbf{C})$ behave? Have upper bound as $t \to \infty$ for all $\mathbf{C} \subset \mathbf{M}^\star \text{closed}$?
- Yes! Compactness helps: Need to prove the upper bound only locally:

$$\mathbb{P}\big(\widetilde{L}_t \simeq \{(p_j, \widetilde{\alpha}_j)_j\}\big)$$

A key step: A new LDP in a compactified space

Turns out weak LDP for L_T is not good enough

- What can we do with this compact space $\mathbf{M}^\star$?
- Equivalence classes $\widetilde{L}_t \in \mathbf{M}^\star$. This is the sequence we want to work with.
- How does $\mathbb{P}\big(\widetilde{L}_t \in \mathbf{C}\big)$ behave? Have upper bound as $t \to \infty$ for all $\mathbf{C} \subset \mathbf{M}^\star \text{closed}$?
- Yes! Compactness helps: Need to prove the upper bound only locally:

$$\mathbb{P}\big(\widetilde{L}_t \simeq \{(p_j, \widetilde{\alpha}_j)_j\}\big)$$

optimal strategy: move "independently" on distant regions

$$\leq \exp\Big\{-\sum_j p_j \underbrace{I(\alpha_j)}\Big\}$$

where $I(\alpha)$ is the Donsker-Varadhan rate function.

Have full LDP on $\mathbf{M}^{\star}$

Our model is shift-invariant: Does not care about equivalence classes!

Theorem (M-Varadhan 2014)

The family of distributions $\widetilde{L}_t$ satisfies a (strong) LDP in the compact space $\mathbf{M}^{\star}$ with rate function

$$\sum_j p_j I(\alpha_j)$$

where $I(\cdot)$ is the Donsker-Varadhan rate function.

Have full LDP on $\mathbf{M}^\star$

Our model is shift-invariant: Does not care about equivalence classes!

Theorem (M-Varadhan 2014)

The family of distributions $\widetilde{L}_t$ satisfies a (strong) LDP in the compact space $\mathbf{M}^\star$ with rate function

$$\sum_j p_j I(\alpha_j)$$

where $I(\cdot)$ is the Donsker-Varadhan rate function.

- Upshot: $Q_t \circ \widetilde{L}_t^{-1}$ converges to $\delta_{\tilde{\mu}_0}$, with $\mu_0 = \psi_0^2$ unique maximizer of the free energy.

Have full LDP on $\mathbf{M}^\star$

Our model is shift-invariant: Does not care about equivalence classes!

Theorem (M-Varadhan 2014)

The family of distributions $\widetilde{L}_t$ satisfies a (strong) LDP in the compact space $\mathbf{M}^\star$ with rate function

$$\sum_j p_j I(\alpha_j)$$

where $I(\cdot)$ is the Donsker-Varadhan rate function.

- Upshot: $Q_t \circ \widetilde{L}_t^{-1}$ converges to $\delta_{\tilde{\mu}_0}$, with $\mu_0 = \psi_0^2$ unique maximizer of the free energy.
- We have a model on a non-compact space. If model is shift-invariant, we can address questions for exponential growth of integrals/ exponential decay of probabilities!

- So L_t under Q_t stays in the infinite tube.

- So L_t under Q_t stays in the infinite tube.
- But the tube is infinitely long!

- So L_t under Q_t stays in the infinite tube.
- But the tube is infinitely long!
- Next step: Tightness: L_t can not fluctuate wildly in the tube. Stays close to starting point.

- So L_t under Q_t stays in the infinite tube.
- But the tube is infinitely long!
- Next step: Tightness: L_t can not fluctuate wildly in the tube. Stays close to starting point.
- Justification requires quite some work. Singularity of the Coulomb potential is a serious problem here.

- So L_t under Q_t stays in the infinite tube.
- But the tube is infinitely long!
- Next step: Tightness: L_t can not fluctuate wildly in the tube. Stays close to starting point.
- Justification requires quite some work. Singularity of the Coulomb potential is a serious problem here.
- Culmination: Mean-field approximation of the Polaron:

Theorem (Bolthausen-König-M 2015)

$$Q_t \circ L_t^{-1} \Rightarrow \frac{\int_{\mathbb{R}^3} dx\, \psi_0(x)\delta_{\theta_x \psi_0^2}}{\int_{\mathbb{R}^3} dx\, \psi_0(x)}$$