Large deviations for singularly interacting diffusions

Mario Maurelli[1]
Joint work with
Jasper Hoeksema, Thomas Holding, Oliver Tse

Università degli Studi di Milano

Saint-Etienne, May 17, 2022

[1] work partly supported by Newton International Fellowship and PRIN project

McKean-Vlasov SDE on $\mathbb{R}^d$:

$$d\bar{X} = b(\bar{X}, \mathrm{Law}(\bar{X}))dt + dW$$

Associated interacting particle system:

$$dX^i = b(X^i, \frac{1}{N}\sum_{j=1}^{N} \delta_{X^j})dt + dW^i, \quad i = 1, \dots N$$

Singular interaction drift b, e.g. $(d \geq 2)$:

$$b(x, \mu) = \int \varphi(x-y)\mu(dy), \quad \varphi(x) \approx |x|^\alpha, \alpha > -1.$$

We will show, in the limit $N \to \infty$:

- **large deviation principle** for interacting particle system
- **convergence** to McKean-Vlasov SDE

McKean-Vlasov SDEs: SDEs where the drift depends on the law of the solution:

$$d\bar{X} = b(\bar{X}, \mathrm{Law}(\bar{X}))dt + dW$$

Associated mean-field interacting particle system:

$$dX^{i,N} = b(X^{i,N}, \frac{1}{N}\sum_{j=1}^{N} \delta_{X^{j,N}})dt + dW^i, \quad i = 1, \dots N$$

Theorem (e.g. Sznitman 91)

If b is Lipschitz and bounded in the 1-Wasserstein topology, then

- *existence and uniqueness of MV SDE*
- *convergence of interacting particle system to MV SDE*

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{X^{i,N}} \to Law(\bar{X}) \text{ as } N \to \infty$$

As a consequence, if $\mu \neq \text{Law}(\bar{X})$, then $P(\frac{1}{N}\sum_{i=1}^{N} \delta_{X^{i,N}} \approx \mu) \to 0$ (as $N \to \infty$). Question: how fast? large deviations:

Theorem (e.g. Budhiraja-Dupuis-Fischer 12)

If b is Lipschitz and bounded, then LDP (on $\mathcal{P}(C([0,T])^d)$) with (speed N and) rate function $R(\mu \mid Law(X^\mu))$

$$P\left(\frac{1}{N}\sum_{i=1}^{N} \delta_{X^{i,N}} \approx \mu\right) \approx \exp[-NR(\mu \mid Law(X^\mu))]$$

where R is the relative entropy and X^μ is defined below:

$$R(\mu \mid \nu) = \int \frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu} d\nu$$
$$dX^\mu = b(X^\mu, \mu)dt + dW$$

Question: what happens for non smooth b?
Two cases:

- *Gibbs 2-point interaction*:

$$b(x,\mu) = \int \varphi(x-y)\mu(dy)$$

  typically φ singular at $x-y=0$ (problems when particles meet or come close)

- *non-Gibbs interaction*:

$$b(x,\mu) = \psi\left(x,\mu,\int \varphi(x-y)\mu(dy)\right)$$

  ψ smooth, typically φ singular at $x-y=0$

Types of singularity (for interaction kernel φ):

- Subcritical: on small scale, drift is much smaller than diffusion

$$\varphi(x) \approx |x|^\alpha \text{ with } \alpha > -1$$

- Critical: on small scale, drift is comparable to diffusion: in 2D (Gibbs case):

$$\varphi(x) = -\frac{x}{|x|^2} \text{ 2D Keller-Segel}$$
$$\varphi(x) = \frac{x^\perp}{|x|^2} \text{ 2D Navier-Stokes}$$

Interacting particle system:

$$dX^{i,N} = b(X^{i,N}, \frac{1}{N}\sum_{j=1}^{N}\delta_{X^{j,N}})dt + dW^i, \quad i = 1, \dots N$$

where

$$b(x,\mu) = \int \varphi(x-y)\mu(dy) \text{ (Gibbs case)}$$

$$b(x,\mu) = \psi\left(x, \mu, \int \varphi(x-y)\mu(dy)\right) \text{ (non-Gibbs case)}$$

Theorem ((part 1))

Assume (subcritical case) that $\varphi \in L^q_t(L^p_x)$ *with*

$$\frac{d}{p} + \frac{2}{q} < 1, \quad p, q \geq 2$$

Then an LDP holds for $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}} \in \mathcal{P}(C([0,T])^d)$ *with weak topology, with rate function* $R(\mu \mid Law(X^\mu))$*:*

$$P\left(\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}} \approx \mu\right) \approx \exp[-NR(\mu \mid Law(X^\mu))]$$

with

$$dX^\mu = b(X^\mu, \mu)dt + dW$$

Theorem ((part 2))

Moreover, $\frac{1}{N}\sum_{i=1}^{N}\delta_{X^{i,N}}$ converges (in probability) to the unique $Law(\bar{X})$ solution to MV SDE.

This includes a large class of subcritical drifts:

- Gibbs and non-Gibbs cases
- $\varphi(x) \approx |x|^{\alpha}$ for $\alpha > -1$ ($\alpha > -1/2$ for $d = 1$)

Some known results on convergence (no claim of completeness):

- Jabin-Wang 18: Gibbs subcritical and critical case $\varphi \in W^{-1,\infty}$ with $\mathrm{div}\varphi \in W^{-1,\infty}$
- Godinho-Quininao 15: Gibbs subcritical case $\varphi = -\nabla|x|^{\alpha+1}$, $\alpha > -1$
- Tomasevic 20: Gibbs subcritical case $\varphi \in L^q_t(L^p_x)$, $d/p + 2/q < 1$ (+ a continuity condition)
- Fournier-Jourdain 17: 2D Keller-Segel
- Fournier-Hauray-Mischler 14: 2D Navier-Stokes
- Lacker 18: non-Gibbs subcritical case, b bounded and continuous in the measure argument
- Jabir 19: non-Gibbs subcritical case, assuming weak uniqueness + an integrability condition
- Hao-Röckner-Zhang 22: non-Gibbs subcritical case, general

Some known results on large deviations (no claim of completeness):

- Dawson-Gärtner 87, DelMoral-Zajic 03, Budhiraja-Dupuis-Fischer 12: LDP assuming continuous drift;
- Lacker 18: subscritical case large deviation upper bound, b bounded and continuous in the measure argument
- Fontbona 04: critical Gibbs 1D case, $\varphi(x) = 1/x$ (repulsive).

Notation:

$$L^{W,N} = \frac{1}{N}\sum_{j=1}^{N}\delta_{W^j} \text{ (independent (Brownian) particles)}$$

$$L^{X,N} = \frac{1}{N}\sum_{j=1}^{N}\delta_{X^{j,N}} \text{ (dependent particles)}$$

Formal idea of the proof (from at least DaiPra-denHollander 96):

1. Independent case: Sanov theorem: LDP for $L^{W,N}$ (W^i i.i.d.)
$$P(L^{W,N} \approx \mu) \approx e^{-NI(\mu)}$$
with $I(\mu) = R(\mu \mid \mathrm{Law}(W^1))$
2. Density for the dependent case: Girsanov theorem
$$\mathrm{Law}(X^{1,N}, \dots X^{N,N}) = e^{\cdots} \mathrm{Law}(W^1, \dots W^N)$$
3. Mean field interaction:
$$\mathrm{Law}(X^{1,N}, \dots X^{N,N}) = e^{-N\mathcal{E}(L^{W,N})} \mathrm{Law}(W^1, \dots W^N)$$
for a suitable energy $\mathcal{E}$; as a consequence
$$\mathrm{Law}(L^{X,N}) = e^{-N\mathcal{E}(L^{W,N})} \mathrm{Law}(L^{W,N})$$
4. LDP for change of measure: Varadhan lemma:
$$P(L^{X,N} \approx \mu) \approx e^{-N\mathcal{E}(\mu)} P(L^{W,N} \approx \mu) \approx e^{-N(\mathcal{E}+I)(\mu)}$$
Formally $\mathcal{E}(\mu) + I(\mu) = R(\mu \mid \mathrm{Law}(X^\mu))$.

The energy $\mathcal{E}$:

$$\mathcal{E}(L^{W,N}) = \frac{1}{N}\sum_{i=1}^{N}\left[-\int_0^T b(W^i, L^{W,N})dW^i + \frac{1}{2}\int_0^T |b(W^i, L^{W,N})|^2 dt\right]$$
$$= \int \left[-\int_0^T b(\gamma, L^{W,N})d\gamma + \frac{1}{2}\int_0^T |b(\gamma, L^{W,N})|^2 dt\right] L^{W,N}(d\gamma)$$

Gibbs case ($b(x,\mu) = \int \varphi(x-y)d\mu$):

$$\mathcal{E}(\mu) = \int\int\int V(\gamma^1,\gamma^2,\gamma^3)d\mu^{\otimes 3}(\gamma^1,\gamma^2,\gamma^3),$$
$$V(\gamma^1,\gamma^2,\gamma^3) = -\int_0^T \varphi(\gamma^1_t-\gamma^2_t)d\gamma^1_t + \frac{1}{2}\int_0^T \varphi(\gamma^1_t-\gamma^2_t)\cdot\varphi(\gamma^1_t-\gamma^3_t)dt$$

Rigorously:

- Varadhan lemma asks the energy $\mathcal{E}$ to be continuous: problem when φ is singular
- the energy $\mathcal{E}$ involves a stochastic integral, which is a non-smooth function of the path: $\mathcal{E}$ is not continuous even for smooth drifts (but an LDP is known for smooth b)
- Girsanov theorem requires suitable sufficient conditions (e.g. Novikov)

Idea: proceed by approximation:

1. take φ_λ smooth bounded approximations of the interaction kernel φ;
2. LDP with interaction kernel φ_λ holds by classical result;
3. let $\lambda \to 0$ and try to get an LDP with φ.

Notation (φ interaction kernel):

- b drift

$$b(x,\mu) = \int \varphi(x-y)\mu(dy) \text{ (Gibbs case)}$$

$$b(x,\mu) = \psi\left(x,\mu,\int \varphi(x-y)\mu(dy)\right) \text{ (non-Gibbs case)}$$

- $\mathcal{E}$ energy (log-density in Girsanov):

$$\mathcal{E}(L^{W,N}) = \int \left[-\int_0^T b(\gamma, L^{W,N})d\gamma + \frac{1}{2}\int_0^T |b(\gamma, L^{W,N})|^2 dt\right] L^{W,N}$$

- in the Gibbs case, V potential:

$$\mathcal{E}(\mu) = \int\int\int V(\gamma^1,\gamma^2,\gamma^3)d\mu^{\otimes 3}(\gamma^1,\gamma^2,\gamma^3),$$

$$V(\gamma^1,\gamma^2,\gamma^3) = -\int_0^T \varphi(\gamma^1-\gamma^2)d\gamma^1 + \frac{1}{2}\int_0^T \varphi(\gamma^1-\gamma^2)\cdot\varphi(\gamma^1-$$

b_λ, $\mathcal{E}_\lambda$, V_λ defined similarly with φ_λ.

Three main points:

1. **Extended Varadhan lemma**: For general change of measures, find sufficient conditions on $\mathcal{E}_\lambda$ to infer the LDP for $\mathcal{E}$ –> exponentially good approximation $\mathcal{E}_\lambda(L^{W,N})$.
2. **Gibbs measures with singular potential**: In the context of Gibbs energy

$$\mathcal{E}(\mu) = \int V(\gamma^1, \dots \gamma^k) d\mu^{\otimes k}(\gamma^1, \dots \gamma^k)$$

   find sufficient conditions of V and its approximation V_λ to meet the exponentially good approximation for $\mathcal{E}_\lambda(L^{W,N})$.
   Gibbs-like case: In the context of non-Gibbs energy, reduce the exponentially good approx property (condition on N particles) to a condition on k particles.
3. **MV SDEs with singular drift**: In the context of MV SDE, find sufficient conditions on φ and its approximation φ_λ to meet the conditions of Gibbs or Gibbs-like case.

Extended Varadhan lemma: change of measure with discontinuous log-densities:
Here $L^{W,N}$ is a generic random variable on a Polish space S, $\mathcal{E}$ and $\mathcal{E}_\lambda$ are functions on S.

Theorem

Assume that

- $L^{W,N}$ *satisfies an LDP*
- *for* $\lambda > 0$, $\mathcal{E}_\lambda$ *induces an LDP (i.e.* $e^{-N\mathcal{E}_\lambda} Law(L^{W,N})$ *satisfies an LDP)*
- *exponentially good approximation: for every* $\beta > 0$,

$$\limsup_N \frac{1}{N} \log E[e^{N\beta|\mathcal{E}-\mathcal{E}_\lambda|(L^{W,N})}] \to 0 \text{ as } \lambda \to 0$$

Then $\mathcal{E}$ *induces an LDP (i.e.* $e^{-N\mathcal{E}} Law(L^{W,N})$ *satisfies an LDP).*

Remark: we do not require $\mathcal{E}_\lambda$ to be continuous, but only to induce an LDP.
The proof is based on convex analysis tools.
Other known results (no claim of completeness):

- Dembo-Zeitouni 10: exponential approximation
- Eichelsbacher-Schmock 02, Liu-Wu 20: in context of singular Gibbs measure, LDP in a stronger topology (by exponential approximation)
- DelMoral-Zajic 03: another extension of Varadhan lemma

In Gibbs case, exponentially good approximation means

$$\limsup_N \frac{1}{N} \log E[\exp[N\beta \frac{1}{N^3} \sum_{1\le i,j,k\le N} |V - V_\lambda|(W^i, W^j, W^k)]] \to 0$$

Question: which conditions on V guarantee the exponentially good approximation?

LDP for Gibbs measures with singular potential V

Here $W^1, \dots W^N$ are i.i.d. random variables on a Polish space χ and

$$L^{W,N} = \frac{1}{N} \sum_{i=1}^{N} \delta_{W^i}$$

are random variables on $\mathcal{P}(\chi)$ endowed with weak topology.

Sanov theorem: $L^{W,N}$ satisfy an LDP with rate function $I(\mu) = R(\mu \mid \mathrm{Law}(W^1))$.

Interaction:

$$\mathcal{E}(\mu) = \int V(\gamma^1, \gamma^2, \gamma^3) d\mu^{\otimes 3}$$

V possibly singular.
V_λ approximations of V, $\mathcal{E}_\lambda$ defined as $\mathcal{E}$ but replacing V with V_λ.

Proposition

Assume that

- *for* $\lambda > 0$, $\mathcal{E}_\lambda$ *induces an LDP (i.e.* $e^{-N\mathcal{E}_\lambda} Law(L^{W,N})$ *satisfies an LDP)*
- *for every* $\beta > 0$,

$$\log E[\exp[\int_S \beta |V - V_\lambda|(W^1, W^2, W^3)]] \to 0 \text{ as } \lambda \to 0$$

Then $\mathcal{E}$ *induces an LDP (i.e.* $e^{-N\mathcal{E}} Law(L^{W,N})$ *satisfies an LDP).*

Recall exponentially good approximation (here two point interaction for simplicity):

$$\limsup_N \frac{1}{N} \log E[\exp[N\beta \frac{1}{N^2} \sum_{1\le i,j\le N} |V - V_\lambda|(W^i, W^j)]] \to 0$$

Idea of the proof: Hoeffding decomposition: rearrange the terms $|V - V_\lambda|(W^i, W^j)$ to get averages of $\approx N$ indipendent terms:

$$\frac{1}{N^2} \sum_{1\le i,j\le N} f(W^i, W^j) \approx \frac{1}{N!} \sum_{\sigma\in S_N} \frac{1}{[N/2]} \sum_{i=1}^{[N/2]} f(W^{\sigma(2i-1)}, W^{\sigma(2i)})$$

$$\frac{1}{N^2}\sum_{1\le i,j\le N} f(W^i,W^j) \approx \frac{1}{N!}\sum_{\sigma\in S_N}\frac{1}{[N/2]}\sum_{i=1}^{[N/2]} f(W^{\sigma(2i-1)},W^{\sigma(2i)})$$

Then, first by Jensen inequality then by independence,

$$\begin{aligned}
&E[\exp[N\beta\frac{1}{N^2}\sum_{1\le i,j\le N}|V-V_\lambda|(W^i,W^j)]]\\
&\le C\frac{1}{N!}\sum_{\sigma\in S_N} E[\exp[N\beta\frac{1}{[N/2]}\sum_{i=1}^{[N/2]}|V-V_\lambda|(W^{\sigma(2i-1)},W^{\sigma(2i)})]]\\
&\le C\frac{1}{N!}\sum_{\sigma\in S_N}\prod_{i=1}^{[N/2]} E[\exp[c\beta|V-V_\lambda|(W^{\sigma(2i-1)},W^{\sigma(2i)})]]\\
&= CE[\exp[c\beta|V-V_\lambda|(W^1,W^2)]]^{[N/2]}
\end{aligned}$$

Other known results (no claim of completeness):

- Eichelsbacher-Schmock 02: LDP in a stronger topology, where $\mathcal{E}$ is continuous
- Liu-Wu 20: LDP in a stronger topology, result very close to ours (even more general)
- Berman 18: critical singularity when state space is $\mathbb{R}^d$ (+ other assumptions)

Gibbs-like framework: $\mathcal{E}$ no more of the form $\int V d\mu^{\otimes k}$, but ...:

Proposition

Assume that

- *for* $\lambda > 0$, $\mathcal{E}_\lambda$ *induces an LDP (i.e.* $e^{-N\mathcal{E}_\lambda} Law(L^{W,N})$ *satisfies an LDP)*
- *there exist* V, V_λ *such that, for every* $\beta > 0$,

$$\log E[e^{N\beta|\mathcal{E}-\mathcal{E}_\lambda|(L^{W,N})}] \leq C \log E[e^{N\frac{c_\beta}{N^3}\sum_{1\leq i,j,k\leq N}|V-V_\lambda|(W^i,W^j,W^k)}]$$

- *for every* $\beta > 0$,

$$\log E[\exp[\int_S \beta|V - V_\lambda|(W^1, W^2, W^3)]] \to 0 \text{ as } \lambda \to 0$$

Then $\mathcal{E}$ *induces an LDP (i.e.* $e^{-N\mathcal{E}} Law(L^{W,N})$ *satisfies an LDP).*

In MV SDEs, Gibbs case ($b(x,\mu) = \int \varphi(x-y)\mu(dy)$),

$$V(W^1,W^2,W^3)$$
$$= -\int_0^T \varphi(W^1,W^2)dW^1 + \frac{1}{2}\int_0^T \varphi(W^1-W^2)\cdot\varphi(W^1-W^3)dt$$

Question: which condition on φ meet the assumption (below) of the Gibbs case?

$$\log E[\exp[\int_S \beta|V-V_\lambda|(W^1,W^2,W^3)]] \to 0$$

In MV SDEs, non-Gibbs case:

$$\mathcal{E}(L^{W,N}) = \int \left[-\int_0^T b(\gamma,L^{W,N})d\gamma + \frac{1}{2}\int_0^T |b(\gamma,L^{W,N})|^2 dt\right] L^{W,N}(d\gamma)$$

Question: can one show a Gibbs-like structure? bound on

$$\frac{1}{N}\log E[e^{-N\beta|\mathcal{E}-\mathcal{E}_\lambda|(L^{W,N})}]$$

In all cases we must control stochastic exponentials

LDP for MV SDEs with singular drift

Interacting particle system:

$$dX^{i,N} = b(X^{i,N}, \frac{1}{N}\sum_{j=1}^{N} \delta_{X^{j,N}})dt + dW^i, \quad i = 1, \dots N$$

where

$$b(x,\mu) = \int \varphi(x-y)\mu(dy) \text{ (Gibbs case)}$$

$$b(x,\mu) = \psi\left(x, \mu, \int \varphi(x-y)\mu(dy)\right) \text{ (non-Gibbs case)}$$

Theorem

Assume that, for every $\beta > 0$,

$$\log E\left[\exp\left[\beta\int_0^T |\varphi - \varphi_\lambda|^2(W^1, W^2)dt\right]\right] \to 0 \text{ as } \lambda \to 0$$

Then $L^{X,N}$ *satisfies an LDP with rate function* $R(\mu \mid Law(X^\mu))$.

Idea of the proof, Gibbs case ($b(x,\mu) = \int \varphi(x-y)d\mu$):
Recall:

$$V(W^1,W^2,W^3)$$
$$= -\int_0^T \varphi(W^1-W^2)dW^1 + \frac{1}{2}\int_0^T \varphi(W^1-W^2)\cdot\varphi(W^1-W^3)dt$$

Use **exponential martingale bounds** to show that

$$\log E[\exp[\int \beta|V-V_\lambda|(W^1,W^2,W^3)]]$$
$$\leq C\log E[\exp[c_\beta \int_0^T |\varphi-\varphi_\lambda|^2(W^1,W^2)dt]]$$

Thanks to the assumption on the RHS, apply the criterion for Gibbs measures with singular potential.

Idea of the proof, non-Gibbs (general) case:
Recall:

$$\mathcal{E}(L^{W,N}) = \int \left[-\int_0^T b(\gamma, L^{W,N}) d\gamma + \frac{1}{2} \int_0^T |b(\gamma, L^{W,N})|^2 dt \right] L^{W,N}(d\gamma)$$

Also here use exponential martingale bounds to show that

$$\frac{1}{N} \log E[e^{-N\beta|\mathcal{E}-\mathcal{E}_\lambda|(L^{W,N})}] \leq C \frac{1}{N} \log E[\exp[-N c_\beta \frac{1}{N^2} \sum_{i,j=1}^{N} \int_0^T |\varphi - \varphi_\lambda|^2 (W^i, W^j) dt]]$$

and apply the Gibbs-like framework.

Proposition

Sufficient condition to guarantee assumption on φ: $\varphi \in L^q_t(L^p_x)$ with $d/p + 2/q < 1$.

The proof is based on:

- Khasminskii Lemma

$$\sup_{x,y\in\mathbb{R}^d} E[\int_0^T |\varphi|^2(x+W^1, y+W^2)dt] < \infty$$
$$\Rightarrow E[\exp \int_0^T |\varphi|^2(W^1, W^2)dt] < \infty$$

- heat kernel bounds to show that

$$E[\int_0^T |\varphi|^2(x+W^1, y+W^2)dt] < \infty$$

Uniqueness for the McKean-Vlasov SDE: idea of the proof in the Gibbs case:
Call $F : \mathcal{P}(C([0,T]^d)) \ni \mu \mapsto \mathrm{Law}(X^\mu) \in \mathcal{P}(C([0,T]^d))$, where

$$dX^\mu = b(X^\mu, \mathrm{Law}(X^\mu))dt + dW$$

Show that F is a contraction (for small T) on $\{\mu \mid \left\| \frac{d\mu}{d\mathrm{Law}(W)} \right\|_{L^m} < \infty\}$.

$$E^1[|\frac{d\mu}{d\mathrm{Law}(W)} - \frac{d\nu}{d\mathrm{Law}(W)}|^m(W^1)]$$

$$\leq E^1[\left|\int\int_0^T \varphi(W^1,y)dt(\mu-\nu)(dy)\right|^{2m+}]^{1/m+}$$

$$\leq E^1\left[\left|E^2[\int_0^T \varphi(W^1,W^2)dt(\frac{d\mu}{d\mathrm{Law}(W)} - \frac{d\nu}{d\mathrm{Law}(W)})(W^2)]\right|^{2m+}\right]$$

$$\leq E\left[\left(\int_0^T |\varphi(W^1,W^2)|^2 dt\right)^{m+}\right]^{1/m+} \cdot$$

$$\cdot E^2[|\frac{d\mu}{d\mathrm{Law}(W)} - \frac{d\nu}{d\mathrm{Law}(W)}|^2(W^2)]^m$$

Results:

- LDP and convergence in Gibbs and non-Gibbs cases with singular interaction kernel in the subcritical class (by Itô calculus and Khasminskii lemma)
- extended Varadhan lemma (by convex integration tools)
- LDP for singular Gibbs measures (by Hoeffding decomposition)

Possible perspectives:

- LDP for some drifts in the critical class?
- other approaches (suited e.g. for degenerate noise)?
- application to metastability?

Thank you!