# BARRIERS OF THE MCKEAN–VLASOV ENERGY VIA A MOUNTAIN PASS THEOREM IN THE SPACE OF PROBABILITY MEASURES

RISHABH S. GVALANI AND ANDRÉ SCHLICHTING

Abstract. We show that the empirical process associated to a system of weakly interacting diffusion processes exhibits a form of noise-induced metastability. The result is based on an analysis of the associated McKean–Vlasov free energy, which for suitable attractive interaction potentials has at least two distinct global minimisers at the critical parameter value $\beta = \beta_c$. On the torus, one of these states is the spatially homogeneous constant state and the other is a clustered state. We show that a third critical point exists at this value. As a result, we obtain that the probability of transition of the empirical process from the constant state scales like $\exp(-N\Delta)$, with $\Delta$ the energy gap at $\beta = \beta_c$. The proof is based on a version of the mountain pass theorem for lower semicontinuous and $\lambda$-geodesically convex functionals on the space of probability measures $\mathcal{P}(M)$ equipped with the $W_2$ Wasserstein metric, where $M$ is a Riemannian manifold or $\mathbb{R}^d$.

## 1. Introduction

In recent years a lot a progress has been made in understanding the convergence of interacting particle systems to their hydrodynamic or mean-field limits at the level of the convergence of gradient flows (cf. [ADPZ11, ADPZ13, DPZ13, Fat16, EFLS16, FS16, KJZ18]). These results lead to dissipative systems which are driven by some macroscopic free energy with respect to some metric. This gradient flow structure allows for a characterisation of the stationary states of the system in terms of the critical points and minimisers of the free energy. Hence, the free energy landscape and the underlying metric encodes some of the dynamic properties of the system. In many of the applications the free energy is usually a lower semicontinuous function with the space of probability measures $\mathcal{P}(M)$ as its domain. Here $\mathcal{P}(M)$ represents the distribution of particle positions on some base manifold $M$. The appropriate metric for the gradient flow is usually the

*Key words and phrases.* Free energy barrier, Large deviations, McKean–Vlasov equation, mountain pass theorem, optimal transport, space of probability measures.

RSG is funded by an Imperial College President's PhD Scholarship, partially through EPSRC Award Ref. 1676118. RSG also acknowledges the hospitality of RWTH Aachen University. AS acknowledges the hospitality of Imperial College London. Part of this work was carried out at the workshop "Nonlocal differential equations in collective behaviour" held at the American Institute of Mathematics, San José and at the "Junior Trimester Programme in Kinetic Theory" held at the Hausdorff Research Institute for Mathematics, Bonn. RSG and AS are grateful to both institutes for their hospitality.

2-Wasserstein metric and its variants. For example, in [BB18] the authors derive a local mean-field model as the gradient flow of the macroscopic free energy w.r.t a modified Wasserstein metric.

For macroscopic models originating from interacting particle systems, the free energy can exhibit multiple local minima corresponding to distinguished stationary states of the macroscopic system. In this case one may ask about typical transition times and transition states between two such distinct states in the presence of noise. A common example of this is a classical particle moving on $\mathbb{R}^d$ along the gradient of some potential $V \in C^2(\mathbb{R}^d;\mathbb{R})$, i.e.,

$$\dot{x}(t) = -\nabla V(x)\,, \tag{1.1}$$

with $x(0) = x_0 \in \mathbb{R}^d$. Let us assume that $V$ has exactly two distinct global minima $x_1, x_2 \in \mathbb{R}^d$, which are also the stationary points of (1.1). If one considers these to be the states of interest, then a relevant question is how does the particle transition from one to the other under the influence of noise. To understand this one considers the stochastic differential equation

$$dX_t = -\nabla V(X_t)\,\mathrm{d}t + \sqrt{2\beta^{-1}}dB_t\,, \tag{1.2}$$

where $B_t$ is a $\mathbb{R}^d$-valued Wiener process and $\beta > 0$ is a parameter representing the strength of the noise in the system. In the setting of the above SDE, the question can be reframed as follows: given $X(0) = x_1$, what is the probability that in some finite time $T > 0$, we have that $X(T) = x_2$. This question is answered, at least for $\beta \gg 1$, by the Freidlin–Wentzell theorem. In particular it tells us that the family of processes $\{X_t^\beta\} \in C([0,T])$ with $X_0 = x_1$ satisfy a large deviations principle with good rate function $S : C([0,T];\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$ given by

$$S(f) = \frac{1}{2}\int_0^T |\dot{f}(t) + \nabla V(f(t))|^2\,\mathrm{d}t\,,$$

whenever the above integral is finite and $+\infty$ otherwise. As a consequence of the above result we have for any closed and measurable $\Gamma \subset C([0,T];\mathbb{R}^d)$ that

$$\limsup_{\beta\to+\infty} \beta^{-1}\log \mathbb{P}(X_t^\beta \in \Gamma) \leq -\inf_{f\in\Gamma} S(f)\,.$$

If we pick $\Gamma = \{f \in C([0,T];\mathbb{R}^d) : f(0) = x_1, f(T) = x_2\}$, we obtain an upper bound on the probability that the process reaches $x_2$ given that it starts at $x_1$. Setting $T^* = \arg\max_{t\in[0,T]}(V(f(t)) - V(f(0)))$, we can obtain the following lower bound for $f \in \Gamma$,

$$\begin{aligned} S(f) &\geq \frac{1}{2}\int_0^{T^*} |\dot{f}(t) + \nabla V(f(t))|^2\,\mathrm{d}t \geq \int_0^{T^*} \dot{f}(t)\cdot\nabla V(f(t))\,\mathrm{d}t \\ &= V(f(T^*)) - V(f(0)) \geq \inf_{f\in\Gamma}(V(f(T^*)) - V(f(0))) =: c - V(f(0))\,. \end{aligned}$$

It turns out that $c > 0$ is in fact a critical value of $V$, i.e., there exists $x_3 \in \mathbb{R}^d$ such that $V(x_3) = c$ and $\nabla V(x_3) = 0$. The reader will recognise this as the finite-dimensional version of the well-known mountain pass theorem. Setting $\Delta := V(x_3) - V(x_1)$, we see that for $\beta$ sufficiently large

$$\mathbb{P}(X_t^\beta \in \Gamma) \lesssim \exp(-\beta\Delta)\,.$$

Thus the probability of the process reaching the new phase/state in time $T > 0$ goes exponentially with $\beta$ with the rate given by the difference between the energies of the saddle point and the initial phase. Thus we can see that the process finds the path of least resistance to reach the new phase in agreement with the fundamental tenet of large deviations theory that "*an unlikely event will happen in the most likely of the possible unlikely ways*". These transitions correspond to the phenomenon of noise-induced metastability, i.e., the process is stable around $x_1$ for $\beta \gg 1$ but there is a finite probability of it transitioning to $x_2$.

The purpose of this paper is to obtain results in a similar flavour but in an infinite-dimensional setting. Specifically, we are interested in understanding how similar phenomena, i.e., noise-induced transitions, occur in systems governed by the Wasserstein gradient flow of some free energy $I$, especially those that arise as mean-field limits of interacting particle systems. We consider the following system of $N$ interacting stochastic differential equations on $\mathbb{T}_L^d$ (the $d$-dimensional torus of side length $L > 0$)

$$dX_t^i = -\frac{1}{N}\sum_{j=1}^N \nabla W(X_t^i - X_t^j)\,\mathrm{d}t + \sqrt{2\beta^{-1}}dB_t^i$$

$$\mathrm{Law}(\overline{X}_0^N) = \prod_{i=1}^N \nu(x_i) \quad \overline{X}_t^N = \left(X_t^1, \dots, X_t^N\right)$$

where $\beta > 0$ is a parameter, $W \in C^2(\mathbb{T}_L^d)$ is an interaction potential which is even along every coordinate, and $B_t^i$ are $\mathbb{T}_L^d$-valued independent Wiener processes. Let $\mu^{(N)}(t) := N^{-1}\sum_{i=1}^N \delta_{X_i^t}$, then it is well known (cf. [Szn91]) that $\mu^{(N)}(t)$ as a measure-valued random variable converges in law to $\mu = \mu(x,t)$ for each $t > 0$, where $\mu$ is a weak solution of the following PDE

$$\partial_t \mu = \nabla \cdot (\mu \nabla(\beta^{-1}\log\mu + W \star \mu)) \qquad \text{with} \qquad \mu(x,0) = \nu(x)\,.$$

The above PDE is commonly referred to as the McKean–Vlasov equation and can be rewritten as $W_2$-gradient flow

$$\partial_t \mu = \nabla \cdot \left(\mu \nabla \frac{\delta I}{\delta \mu}\right),$$

where $I : \mathcal{P}(\mathbb{T}_L^d) \to \mathbb{R} \cup \{+\infty\}$ is the associated free energy. Its domain is the space of absolutely continuous measures and for those it is given by

$$I(\mu) = \beta^{-1} \int \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}x}\right) \mathrm{d}\mu + \frac{1}{2} \iint W(x-y) \, \mathrm{d}\mu(y) \, \mathrm{d}\mu(x) \,, \tag{1.3}$$

where $\frac{\mathrm{d}\mu}{\mathrm{d}x}$ denotes the density of $\mu$ w.r.t the Lebesgue measure $\mathrm{d}x$ on $\mathbb{T}_L^d$. The first term in (1.3) is referred to as the entropy and the second as the interaction energy. The function $I$ is referred to as the free energy of the system. The balance between entropy and interaction energy in terms of $\beta$ determines what the minimisers of $I$ look like. For $\beta$ smaller than some critical value $\beta_c$, the normalised Lebesgue measure, $\mu_0(\mathrm{d}x) = \mathrm{d}x/L^d$ is the unique minimiser of the free energy. Above the value $\beta_c$ a new minimiser of the free energy appears which is not $\mu_0$ emerges. This phenomenon in which there is a change in structure of the set of minimisers of $I$ is called a phase transition and is observed in many models from the physical sciences [LP66, Sin82, Daw83, Shi87, GP18, FV18].

This operator $\nabla \cdot (\mu \nabla \frac{\delta}{\delta \mu}(\cdot))$ can be formally thought of as a gradient in the space of probability measures on $\mathbb{T}_L^d$ equipped with the $W_2$ mass transportation distance defined as follows

$$W_2^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \iint_{\mathbb{T}_L^d} d(x,y)^2 \, \mathrm{d}\pi(x,y) \, \mathrm{d}x$$

where $\Pi(\mu, \nu)$ is the set of all couplings between $\mu$ and $\nu$ and $d(x, y)$ is the distance on $\mathbb{T}_L^d$. As mentioned earlier $\mathcal{P}(\mathbb{T}_L^d)$ equipped with $W_2$ is a complete, separable, metric space. For $\mu, \nu$ absolutely continuous w.r.t $\mathrm{d}x$ the definition of the metric can be recast into the form discussed in Theorem 3.2. This notion of a gradient flow can be made rigorous and is an extremely active field of research, where as the present work relies on quite classical results from [CEMS01, McC01, McC97, AGS08]. Indeed, the solutions of the McKean–Vlasov PDE can be thought of as curves of maximal slope of the McKean–Vlasov energy $I$ w.r.t $W_2$ (see [DS10]). Comparing this with the toy model discussed further up in the introduction, we see that the PDE has a gradient structure in $W_2$ and so the functional $I$ will play a similar role to the potential $V$ in (1.1). The distinct phases/states are then characterised by the global minima of the functional $I$ over $\mathcal{P}(\mathbb{T}_L^d)$. The role of the SDE in (1.2) is then played by the empirical process $\mu^{(N)}$ and that of the parameter $\beta$ is played by $N$.

Understanding such transitions requires two ingredients, a version of the mountain pass theorem in the space of probability measures $\mathcal{P}(M)$ equipped with the Wasserstein metric and an appropriate large deviations principle for the underlying particle system. We focus on the first ingredient noting that that the second ingredient is usually application-specific. Our main result in this direction is as follows.

**Theorem 1.1.** *Let $I : \mathcal{P}(M) \to \mathbb{R} \cup \{+\infty\}$ be a proper, l.s.c, and $\lambda$-geodesically convex functional and assume that if $\mu \in D(I)$, then $\mu \ll \mathrm{vol}$. Suppose $\mu, \nu \in \mathcal{P}(M) \cap D(I)$, $\Gamma$ is the set of all continuous curves $\gamma : [0,1] \to \mathcal{P}(M)$ (where $\mathcal{P}(M)$ is equipped with the 2-Wasserstein metric, $W_2$) with $\gamma(0) = \mu$ and $\gamma(1) = \nu$, and the function $\Upsilon : \Gamma \to \mathbb{R}$ is defined by:*

$$\Upsilon(\gamma) = \max_{t\in[0,1]} I(\gamma(t))\,.$$

*Let $c = \inf_{\gamma\in\Gamma} \Upsilon(\gamma)$ and $c_1 = \max\{I(\mu), I(\nu)\}$. If $c > c_1$ and $I$ satisfies* (**MPS**) *(see Assumption* (2.2)*), then there exists a $c' \geq c$ such that $c'$ is a critical value of $I$, i.e., there exists a $\eta \in \mathcal{P}(M)$ with $I(\eta) = c'$ such that $|\partial I|(\eta) = |dI|(\eta) = 0$ (see Definition* 2.1 *and Definition* 4.5*).*

The proof utilises the notion of the weak metric slope $|dI|$ first introduced in [Kat94]. The main advantage over previous results in this direction is that we can apply the result to l.s.c functionals on $\mathcal{P}(M)$ as long as they are $\lambda$-convex by working with the extension of the function to its epigraph based on ideas discussed by Degiovanni and Marzocchi [DM94] originating from work in [DGMT80]. In fact for $\lambda$-convex functionals one can identify the usual (strong) metric slope $|\partial I|$ and $|dI|$. We focus on the case in which the metric is $W_2$ although the results generalise for $W_p$ or other variants of the metric.

By relying on results from [CGPS18], we establish conditions on the interaction potential $W$ such that two distinct minimisers $\{\mu_0, \overline{\mu}\}$ of the free energy $I$ (1.3) exist. Hereby $\mu_0$ is the uniform state and $\overline{\mu}$ is a clustered state. After having done so, we can apply Theorem 1.1 to obtain the following result.

**Theorem 1.2.** *Assume $W$ and $\beta_c$ are such that there exist at least two distinct global minimisers $\{\mu_0, \overline{\mu}\}$ of $I$. Then, there exists $\mu^* \in \mathcal{P}(\mathbb{T}^d_L)$ distinct from $\mu_0$ and $\overline{\mu}$ such that $|\partial I|(\mu^*) = |dI|(\mu^*) = 0$. Additionally, $I(\mu^*) = c'$ where $c'$ is characterised by*

$$c' \geq c = \inf_{\gamma\in\Gamma} \max_{t\in[0,1]} I(\gamma(t))\,,$$

*where $\Gamma = \{C([0,1]; \mathcal{P}(\mathbb{T}^d_L)) : \gamma(0) = \mu_0, \gamma(1) = \overline{\mu}\}$.*

The fact that the free energy functional $I$ has an energy barrier at $\beta = \beta_c$ allows us to study escape probabilities for the underlying particle system using results which were first proved by Dawson and Gärtner [DG87]. We refer the reader to [ADPZ11, Rey18, GPY13] for further discussions of the connections between large deviations theory and theory of gradient flows.

**Theorem 1.3.** *Assume $W$ and $\beta_c$ are such that there exist at least two distinct minimisers $\{\mu_0, \overline{\mu}\}$ of $I$. It follows then that the underlying empirical process $\mu^{(N)} \in \mathcal{C}_T$ with initial*

*i.i.d uniformly distributed particles satisfies*

$$\mathbb{P}(\mu^N(T) \in \overline{B}_\varepsilon^{W_2}(\overline{\mu}), \mu^{(N)}(0) = \mu_0^{(N)}) \lesssim \exp\Big(-N(\Delta - O(\varepsilon^2))\Big)$$

*for $N$ sufficiently large, where $\overline{B}_\varepsilon^{W_2}(\overline{\mu})$ is the closed ball of size $\varepsilon > 0$ around $\overline{\mu}$ in $W_2$, $\Delta := I(\mu^*) - I(\mu_0)$ and $\mu^*$ is the critical point defined in Theorem 5.7.*

The result above says that the probability of the empirical process reaching the clustered state $\overline{\mu}$ from the uniform state $\mu_0$ in time $T > 0$ becomes exponentially small as the number of particles increases as long as the system is at a discontinuous transition point.

**Remark 1.4.** *Even though the Dawson–Gärtner large deviations principle provides an exponential lower bound on the above probability, it is not clear how this can be compared to the energy barrier* $\exp(-N(\Delta - O(\varepsilon^2)))$ *for a general model. However, such a lower bound could be obtained, for example, in the following setting: for all $\varepsilon > 0$, there exist two points in $\mu_0^*, \overline{\mu}^*$ in a neighborhood of $\mu^*$ such that $\mu_0^*$ is connected to $\mu_0$ and $\overline{\mu}^*$ is connected to $\overline{\mu}$ through a heteroclinic orbit under the flow of the McKean–Vlasov PDE. However, it is unlikely that such heteroclinic connections exist at this level of generality in the choice of $W$. A first step in this direction would be the characterisation of $\mu^*$ for specific choices of $W$.*

**Outline.** The paper is organised as follows: In Section 2 we introduce the notion of the weak metric slope and metric critical points that we will use throughout the paper and a version of the mountain pass theorem due to Katriel that holds for continuous functions on metric spaces. In Section 3, we briefly recall some results due to McCann on optimal transport on Riemmanian manifolds. In Section 4, we compare the notion of weak metric slope with the notion of (strong) metric slope used in the gradient flows community and show that, under the assumption of $\lambda$-convexity of $I$, the two are equivalent. We conclude the section by proving Theorem 1.1. We conclude with Section 5 in which discuss a specific application of the result: the McKean–Vlasov model. We state and extend some results from [CGPS18] on the structure of the set of minimisers of $I$ and their phase transitions. We proceed by showing the existence of mountain pass at the point of discontinuous phase transition thus proving Theorem 1.2. Finally, we introduce the precise form of the large deviations principle due to Dawson and Gärtner and complete the proof of Theorem 1.3. This implies a form of noise-induced metastability for the underlying particle system.

## 2. Critical points in metric spaces

We will assume throughout this section that $(\mathcal{X}, d)$ is a complete metric space. We start with the definition of the weak metric slope for some real-valued continuous function

defined on $\mathcal{X}$. The original notion comes from Ioffe and Schwartzman [IS96] who provided the definition in the Banach space setting.

**Definition 2.1** ($\delta$-regular points, weak metric slope and critical points [Kat94])**.** *Let $x \in \mathcal{X}$, and $I : \mathcal{X} \to \mathbb{R}$ be a continuous function defined in a neighbourhood of $x$. Given $\delta > 0$, $x$ is said to be a $\delta$-regular point of $I$ if there is a neighbourhood $U$ of $x$, a constant $\alpha > 0$, and a continuous mapping $\psi : U \times [0, \alpha] \to \mathcal{X}$ such that for all $(u, t) \in U \times [0, \alpha]$, it holds:*

*(1) $d(\psi(u, t), u) \leq t$.*
*(2) $I(u) - I(\psi(u, t)) \geq \delta t$.*

*$\psi$ is called a $\delta$-regularity mapping for $I$ at $x$. $x$ is called a* regular point *of $I$ if there exists a $\delta$-regularity mapping $\psi$ for some $\delta > 0$.*

*The* weak metric slope *of $I$ at $x$ is given by the extended real number*

$$|dI|(x) = \sup\{\delta \in (0, \infty) : I \text{ is } \delta\text{-regular at } x\}.$$

*If $x$ is not $\delta$-regular for any $\delta > 0$, then $x$ is called a* critical point *of $I$ with $|dI|(x) = 0$.*

**Assumption 2.2** (Weak Palais–Smale condition **(MPS)**)**.** *A function $I : \mathcal{X} \to \mathbb{R}$ is said to satisfy the weak metric Palais–Smale condition **(MPS)** if any* Palais sequence*, that is $\{u_n\}_{n\in\mathbb{N}} \in \mathcal{X}$ with $I(u_n) \to c \in \mathbb{R}$ and $|dI|(u_n) \to 0$, possesses a convergent subsequence.*

Given this notion, we have the following generalisation of the Ambrosetti–Rabinowitz mountain pass theorem due to Katriel [Kat94].

**Theorem 2.3.** *Let $\mathcal{X}$ be a path-connected metric space and $I : \mathcal{X} \to \mathbb{R}$ be continuous. For $u_0, u_1 \in X$ let $\Gamma$ be the set of all continuous curves $\gamma : [0, 1] \to X$ with $\gamma(0) = u_0$ and $\gamma(1) = u_1$, and the function $\Upsilon : \Gamma \to \mathbb{R}$ is given by*

$$\Upsilon(\gamma) = \max_{t\in[0,1]} I(\gamma(t)).$$

*Let $c = \inf_{\gamma\in\Gamma} \Upsilon(\gamma)$ and $c_1 = \max\{I(u_0), I(u_1)\}$. If $c > c_1$ and $I$ satisfies* (**MPS**)*, then $c$ is a critical value of $I$.*

For the application considered in this paper, we need a working definition of the weak slope if $I$ is only lower semicontinuous. Consider the example $I : \mathbb{R} \to \mathbb{R}$ with $I(x) = x + 1$ for $x < 0$ and $I(x) = x$ for $x \geq 0$. Then $I$ is l.s.c and it is easy to verify, that $f$ has a critical point at $x = 0$ in the sense of Definition 2.1. However, this seems to be in some sense pathological as $I$ has a jump at $x = 0$. We like to use a theory of critical points for l.s.c. functionals which disregards such pathological points. Degiovanni and Marzocchi [DM94] using notions developed in [DGMT80] suggested the following generalisations.

**Definition 2.4** (Extension to the epigraph)**.** *Let $I : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be a proper l.s.c functional and denote by* $\mathrm{epi}(I) = \{(u, \xi) \in \mathcal{X} \times \mathbb{R} : I(u) \leq \xi\}$ *its epigraph. Then, its epigraph extension $\mathcal{G}_I : \mathrm{epi}(I) \to \mathbb{R} \cup \{+\infty\}$ is defined by*

$$\mathcal{G}_I(u, \xi) = \xi, \qquad (u, \xi) \in \mathrm{epi}(I)\,.$$

*Additionally,* $\mathrm{epi}(I)$ *is equipped with the metric* $d_{\mathrm{epi}}((u, \xi), (v, \zeta)) = \sqrt{d(u, v)^2 + |\xi - \zeta|^2}$.

It is now straightforward to check that $\mathcal{G}_I$ is a continuous function with respect to $d_{\mathrm{epi}}$ and that $|d\mathcal{G}_I|(u, \xi) \leq 1$ for all $(u, \xi) \in \mathrm{epi}(I)$. It turns out, that the notion of the weak slope for l.s.c functions on $\mathcal{X}$ based on the epigraph extension is suitable for applications to mountain pass theorems in metric spaces.

**Definition 2.5.** *Let $I : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be a proper l.s.c function. Define its domain as*

$$D(I) := \{x \in \mathcal{X} : I(x) < +\infty\}\,.$$

*Then for any $x \in D(I)$, we define its* weak metric slope *at $x$ by*

$$|dI(x)| = \begin{cases} \frac{|d\mathcal{G}_I(x, I(x))|}{\sqrt{1 - |d\mathcal{G}_I(x, I(x))|^2}} & \text{if } |d\mathcal{G}_I(x, I(x))| < 1 \\ +\infty & \text{if } |d\mathcal{G}_I(x, I(x))| = 1\,. \end{cases}$$

*Again, $x \in D(I)$ is called* critical point *of $I$ if $(x, I(x)) \in \mathrm{epi}(I)$ is a critical point of $|d\mathcal{G}_I|(u, I(u))$.*

In the case when $I$ is continuous the above definition is equivalent to Definition 2.1. Indeed, it holds by [DM94, Proposition 2.3], that in this case

$$|d\mathcal{G}_I(x, I(x))| = \begin{cases} \frac{|dI(x)|}{\sqrt{1 + |dI(x)|^2}} & \text{if } |dI(x)| < \infty \\ 1 & \text{if } |dI(x)| = \infty \end{cases} \quad \text{and} \quad |d\mathcal{G}_I(x, \xi)| = 1 \text{ if } I(x) < \xi\,.$$

Hence, the Definition 2.5 is a generalisation of the weak metric slope from Definition 2.1 to lower semicontinuous functionals. However, this definition is, in general, hard to verify. For this reason, we state without a proof a result from [DM94] that povides a lower bound on $|dI|$.

**Proposition 2.6** ([DM94, Proposition 2.5])**.** *Let $I : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be a proper, l.s.c functional and for $b \in \mathbb{R}$ let $D(I)_b = \{x \in D(I) : I \leq b\}$. If for some $x \in D(I)$ there exist constants $\delta > 0, b > I(x), \alpha > 0$, a neighbourhood $U$ of $x$, and a mapping $\Psi : (U \cap D(I)_b) \times [0, \alpha] \to \mathcal{X}$ such that for all $(u, t) \in U \cap D(I)_b \times [0, \alpha]$ it holds that*

$$d(\Psi(u, t), u) \leq t \quad \textit{and} \quad I(u) - I(\Psi(u, t)) \geq \delta\, t\,.$$

*Then, $|dI|(x) \geq \delta$.*

We return to the previous example $I : \mathbb{R} \to \mathbb{R}$ with $I(x) = x + 1$ for $x < 0$ and $I(x) = x$ for $x \geq 0$. In regard of Definition 2.1, we pick $U$ to be the ball of size $\delta$ around $(0, 0)$ in

epi($f$). Choosing the map $\Phi((x,\xi),t) = ((x+t,\xi),t)$, we have that $|d\mathcal{G}_f|(0,0) = 1$ and thus $|df|(0) = +\infty$. Thus the new definition captures the fact that $f$ has a jump at $x = 0$ and correctly does not classify it as a criticial point. Although we can apply the mountain pass Theorem 2.3 to the function $\mathcal{G}_I$, we do not know if the critical point we obtain is of the form $(x, I(x))$, that is we have no information about how $|d\mathcal{G}_I|$ behaves away from the boundary of epi($I$). Degiovanni and Marzocchi [DM94] provide some intuition in the case in which $I$ is a functional defined on a Banach space and consists of convex l.s.c part plus a $C^1$ perturbation. The critical point in this case is defined relative to the metric generated by the norm. This problem has to be treated differently in our case. Before discussing this in further detail, we first cover some preliminaries on optimal transport.

## 3. Optimal transport on manifolds

Let $M$ be a complete, connected Riemannian manifold equipped with a metric given in local coordinates by $g_{ij}$. We denote the geodesic distance between $x, y \in M$ by $d(x,y)$ and the Riemannian volume element by $\mathrm{dvol}(x) = \sqrt{\det g_{ij}(x)}\,\mathrm{d}x$ in local coordinates. For $x \in M$, we denote the inner product on the tangent space $T_xM$ by $\langle\cdot,\cdot\rangle$. Let $c(x,y) := d(x,y)^2/2$ denote the cost function. We now introduce the following definition following [CEMS01].

**Definition 3.1.** *Let $A, B$ be compact subsets of $M$. The set $\mathcal{I}^c(A,B)$ of* $c$-concave functions *is the set of functions $\phi : A \to \mathbb{R} \cup \{-\infty\}$ not identically $-\infty$, for which there exists a function $\psi : B \to \mathbb{R} \cup \{-\infty\}$ such that*

$$\phi(x) = \inf_{y \in B} c(x,y) - \psi(y), \quad \forall x \in A\,.$$

*We refer to $\phi$ as the c-transform of $\psi$ and abbreviate it as $\phi = \psi^c$.*

We have the following main result on the well-posedness of the Monge problem from [McC01].

**Theorem 3.2.** *Let $M$ be a complete Riemannian manifold. Fix two Borel probability measures $\mu \ll \mathrm{vol}$ and $\nu$ on $M$ and two compact subsets $A, B \subset M$ containing the supports of $\mu$ and $\nu$, respectively. Then there exists a $\phi \in \mathcal{I}^c(A,B)$ such that the map*

$$F(x) := \exp_x(-\nabla\phi(x)) \qquad \text{is a pushforward of } \mu \text{ to } \nu.$$

*Furthermore, $F$ is the unique minimiser of the quadratic cost $\int_M c(x, G(x))\,\mathrm{d}\mu(x)$ among all Borel maps $G : M \to M$ pushing $\mu$ forward to $\nu$ apart from variations on sets of $\mu$-measure zero. It follows then that the $W_2$ transportation distance between $\mu$ and $\nu$ takes the following form*

$$W_2^2(\mu,\nu) = \int_M d(x, F(x))^2\,\mathrm{d}\mu(x)\,.$$

The natural extension on McCann's notion of displacement interpolation [McC97] to the manifold setting is given in the following definition.

**Definition 3.3** (Optimal interpolant)**.** *Let $M$ be a complete Riemannian manifold. Fix two Borel probability measures $\mu \ll \mathrm{vol}$ and $\nu$ on $M$ and two compact subsets $A, B \subset M$ containing the supports of $\mu$ and $\nu$, respectively. We define the optimal interpolant to be the map $t \mapsto \mu_t$ for $t \in [0,1]$ such that $\mu_t = (F_t)_{\#}\mu$ and $F_t = \exp_x(-t\nabla\phi(x))$. Here $\phi \in \mathcal{I}^c(\overline{A}, B)$ is the so-called Kantorovich potential from Theorem* 3.2*.*

We are finally in a position to conclude this section with the following results from [CEMS01] about the properties of the optimal interpolant.

**Lemma 3.4.** *Let $M$ be a complete Riemannian manifold. Fix two Borel probability measures $\mu \ll \mathrm{vol}$ and $\nu$ on $M$ and two compact subsets $A, B \subset M$ containing the supports of $\mu$ and $\nu$, respectively. Then the following two results hold*

- *(a) (Optimality of the optimal interpolant.) The map $F_t$ defined in Definition* 3.3 *is the minimiser of the quadratic cost between $\mu_t$ and $\mu$ among all maps pushing forward $\mu$ to $\mu_t$ for all $t \in [0,1]$.*
- *(b) (Absolute continuity of the interpolant.) If $\mu$ and $\nu$ are compactly supported absolutely continuous w.r.t the Riemannian volume, then so is $\mu_t$ for all $t \in [0,1]$.*

## 4. A mountain pass theorem in $\mathcal{P}(M)$

We now turn to the question of obtaining a notion of mountain passes for l.s.c functions. We fix our metric space to be $\mathcal{X} = \mathcal{P}(M)$, where $M$ is now a compact complete connected Riemannian manifold or $\mathbb{R}^d$, and we equip it with the $d = W_2$ transportation distance which makes it a complete, separable metric space [Vil08]. The functionals under consideration satisfy a geodesic $\lambda$-convexity assumption introduced in the following definition.

**Definition 4.1** (Geodesic $\lambda$-convexity)**.** *A proper l.s.c function $I : \mathcal{P}(M) \to \mathbb{R} \cup \{+\infty\}$ is said to be* $\lambda$-geodesically convex *for some $\lambda \in \mathbb{R}$, if for any $\mu, \nu \in \mathcal{P}(M) \cap D(I)$ it holds that $I(\mu_t)$, where $\mu_t$ is the optimal interpolant defined in Definition* 3.3*, satisfies*

$$I(\mu_t) \leq (1-t)I(\mu) + tI(\nu) - \frac{\lambda}{2}t(1-t)W_2^2(\mu,\nu) \quad \forall t \in [0,1].$$

The following Lemma, which proof is similar in spirit to [DM94, Theorem 3.13], shows that the weak metric slope of $\mathcal{G}_I$ is non-zero for geodesically $\lambda$-convex functionals away from the boundary of the epigraph. In particular, any critical point of $\mathcal{G}_I$, if present, lies on the boundary of the epigraph.

**Lemma 4.2.** *Let $I : \mathcal{P}(M) \to \mathbb{R} \cup \{+\infty\}$ be a proper, l.s.c, and $\lambda$-geodesically convex function and assume all $\mu \in D(I)$ satisfy $\mu \ll \mathrm{vol}$. Then, it holds*

$$|d\mathcal{G}_I|(\mu, \xi) = 1 \qquad \text{if } \xi > I(\mu)\,.$$

*In particular, any critical point of $\mathcal{G}_I$, if it exists, lies on $\partial\,\mathrm{epi}(I)$.*

*Proof.* Let $(\mu', \xi) \in \mathrm{epi}(I)$ be such that $\xi = I(\mu') + 2\varepsilon$ for some $\varepsilon > 0$. We define for any $\delta > 0$ the map $\Psi : B_\delta^{d_{\mathrm{epi}}}(\mu', \xi) \times [0, \varepsilon] \to \mathrm{epi}(I)$ as follows

$$\Psi\big((\mu, \alpha), t\big) = \Big(\mu_{\frac{t}{\Lambda}}, \alpha - \frac{t}{\Lambda}\Big(\alpha - \frac{|\lambda|}{2} W_2^2(\mu, \mu') - I(\mu')\Big)\Big)\,, \tag{4.1}$$

where

$$\Lambda = \sqrt{W_2^2(\mu, \mu') + \Big|\Big(\alpha - \frac{|\lambda|}{2} W_2^2(\mu, \mu') - I(\mu')\Big)\Big|^2}$$

and $\mu_{\frac{t}{\Lambda}}$ is the optimal interpolant between $\mu'$ and $\mu$. Since $\xi \geq I(\mu') + 2\varepsilon$, we find $\delta_0 = \delta_0(\varepsilon)$ such that for all $\delta \in (0, \delta_0)$ it holds

$$\varepsilon \leq \xi - I(\mu') - \frac{|\lambda|}{2}\delta^2 - 2\delta\,. \tag{4.2}$$

The above estimate yields $\Lambda \geq \varepsilon$ implying $\frac{t}{\Lambda} \in [0, 1]$ and so $\mu_{\frac{t}{\Lambda}}$ is well defined, provided that $\delta \in (0, \delta_0)$. We also have from Lemma 3.4 (a) that

$$d_{\mathrm{epi}}\Big(\Psi\big((\mu, \alpha), t\big), (\mu, \alpha)\Big) = t\,.$$

Thus the map $\Psi$ satisfies condition (1) of Definition 2.1. We have to check that $\Psi\big((\mu, \alpha), t\big) \in \mathrm{epi}(I)$. Indeed we have from the definition of $\lambda$-geodesic convexity

$$\begin{aligned} I(\mu_{\frac{t}{\Lambda}}) &\leq I(\mu) + \frac{t}{\Lambda}(I(\mu') - I(\mu)) - \frac{\lambda}{2}\frac{t}{\Lambda}\Big(1 - \frac{t}{\Lambda}\Big) W_2^2(\mu, \mu') \\ &\leq \alpha - \frac{t}{\Lambda}(\alpha - I(\mu')) + \frac{|\lambda|}{2}\frac{t}{\Lambda} W_2^2(\mu, \mu') \\ &= \alpha - \frac{t}{\Lambda}\Big(\alpha - \frac{|\lambda|}{2} W_2^2(\mu, \mu') - I(\mu')\Big)\,. \end{aligned}$$

Finally, we can proceed from (4.1) to the following estimate

$$\begin{aligned} \mathcal{G}_I(\Psi((\mu, \alpha), t)) &= \alpha - \frac{t}{\Lambda}\Big(\alpha - \frac{|\lambda|}{2} W_2^2(\mu, \mu') - I(\mu')\Big) \\ &\leq \mathcal{G}_I((\mu, \alpha)) - t\frac{\xi - I(\mu') - \delta - \delta^2\frac{|\lambda|}{2}}{\sqrt{\delta^2 + \Big|\xi - I(\mu') + \delta + \delta^2\frac{|\lambda|}{2}\Big|^2}}\,. \end{aligned}$$

Thanks to (4.2), we can make $\delta$ aribitrarily small and obtain that $|d\mathcal{G}_I|(\mu', \xi) \geq 1$ from Definition 2.1 (2). Since $|d\mathcal{G}_I|(\mu', \xi) \leq 1$ by Definition 2.4, the result follows. □

Having showed that the weak metric slope of $\mathcal{G}_I$ is a constant equal to one away from the boundary of the epigraph, we investigate how the critical points of $I$ defined through the weak metric slope relate to other relevant notions. Specifically, we compare it to the notion of critical point derived from the strong metric slope used in theory of gradient flows [AGS08]. This theory makes rigorous the notion of the Wasserstein gradient discussed in the introduction. We briefly introduce some terminology. Let $I : \mathcal{P}(M) \to \mathbb{R} \cup \{+\infty\}$ be a proper, l.s.c, and $\lambda$-geodesically convex.

**Definition 4.3** (Absolutely continuous curves)**.** *A curve $\mu : (a, b) \subset \mathbb{R} \to \mathcal{P}(M)$ is said to belong to $AC^p_{loc}(a, b; \mathcal{P}(M))$ for some $p \in [1, +\infty]$ if there exists $m \in L^p_{loc}(a, b)$ such that*

$$W_2(\mu_s, \mu_t) \leq \int_s^t m(r)\,\mathrm{d}r, \quad a < s \leq t < b\,. \tag{4.3}$$

*If $p = 1$, then $\mu$ is said to be an absolutely continuous curve.*

**Theorem 4.4** (Metric derivative)**.** *If $\mu : (a, b) \to \mathcal{P}(M)$ is an absolutely continuous curve then the limit*

$$|\mu'|(t) = \lim_{s\to t} \frac{W_2(\mu_s, \mu_t)}{|t - s|}\,,$$

*exists for a.e. $t$ and is called the metric derivative of $\mu$. Additionally, $|\mu'| \in L^1(a, b)$ and is admissible as an $m$ in* (4.3)*. In fact it is the minimal admissible $m$, i.e.,*

$$|\mu'|(t) \leq m(t)$$

*for $t$ a.e. where $m$ satisfies* (4.3)*.*

Now we introduce the notion of the (strong) metric slope.

**Definition 4.5** (Metric slope)**.** *The metric slope $|\partial I|$ of $I$ at $\mu \in \mathcal{P}(M)$ is defined as*

$$|\partial I|(\mu) = \begin{cases} \limsup_{\nu\to\mu} \frac{(I(\mu)-I(\nu))_+}{W_2(\mu,\nu)} & \mu \in D(I) \\ +\infty & \text{otherwise}\,. \end{cases}$$

Finally, we are in a position to define the notion of a curve of maximal slope.

**Definition 4.6** (Curves of maximal slope)**.** *We say that a curve $\mu \in AC^2(0, +\infty; \mathcal{P}(M))$ is a curve of maximal slope of the function $I$ if the following energy dissipation inequality is satisfied*

$$\frac{1}{2}\int_s^t |\mu'|^2(r)\,\mathrm{d}r + \frac{1}{2}\int_s^t |\partial\phi|^2(\mu_r)\,\mathrm{d}r \leq I(\mu_s) - I(\mu_t)\,,$$

*for all $0 < s \leq t < +\infty$. We say that $\mu$ is a stationary curve of maximal slope if it is a curve of maximal slope and $\mu_t = \mu_s$ for all $s, t \in (0, +\infty)$.*

We have the following straightforward corollary.

**Corollary 4.7.** *A curve of maximal slope $\mu$ of a function $I$ is stationary if and only if $|\partial I|(\mu) = 0$.*

Using all these notions we can finally compare the weak metric slope with the metric slope.

**Proposition 4.8** (Equivalence of the two notions of slope)**.** *Let $I : \mathcal{P}(M) \to \mathbb{R} \cup \{+\infty\}$ be a proper, l.s.c, and $\lambda$-geodesically convex functional. Assume further that if $\mu \in D(I)$, then $\mu \ll \mathrm{vol}$. Then for $\mu \in D(I)$ we have that $|dI|(\mu) = |\partial I|(\mu)$.*

*Proof.* We first show that $|dI|(\mu) \leq |\partial I|(\mu)$. Let $\mathcal{G}_I$ be the continuous extension to the epigraph and let $\Psi$ be a $\delta$-regularity mapping for the point $(\mu, I(\mu))$ with $\mu \in D(I)$, that is by Definition 2.1

$$\frac{I(\mu) - \Psi((\mu, I(\mu)), t)}{d_{\mathrm{epi}}((\mu, I(\mu)), \Psi((\mu, I(\mu)), t))} \geq \delta\,.$$

At the same time, we can choose $\Psi((\mu, I(\mu)), t)$ as the approximating sequence in Definition 4.5 of the strong metric slope and obtain

$$|\partial \mathcal{G}_I|(\mu, I(\mu)) \geq \delta\,.$$

Taking the supremum over all such $\delta$, we have the bound $|\partial \mathcal{G}_I|(\mu, I(\mu)) \geq |d\mathcal{G}_I|(\mu, I(\mu))$. This yields only the comparison of the two different slopes of the epigraph extension $\mathcal{G}_I$. To obtain the comparison of the slopes of the functional $I$ itself, we first assume that $\mu \in D(I)$ is not a local minima, and $|\partial I|(\mu) < +\infty$. Then there exists a sequence $(\nu_n, I(\nu_n)) \in \mathrm{epi}(I)$ such that it converges to $(\mu, I(\mu))$ and such that $I(\nu_n) \leq I(\mu)$ for all $n$ sufficiently large. Using this as the approximating sequence in Definition 4.5, we obtain

$$\begin{aligned}
|\partial \mathcal{G}_I|(\mu, I(\mu)) &= \limsup_{(\nu_n, I(\nu_n)) \to (\mu, I(\mu))} \frac{(I(\mu) - I(\nu_n))_+}{\sqrt{|I(\mu) - I(\nu_n)|^2 + W_2(\mu, \nu_n)^2}} \\
&= \limsup_{(\nu_n, I(\nu_n)) \to (\mu, I(\mu))} \frac{(I(\mu) - I(\nu_n))_+}{\sqrt{(I(\mu) - I(\nu_n))_+^2 + W_2(\mu, \nu_n)^2}} \\
&= \frac{|\partial I|(\mu)}{\sqrt{1 + |\partial I|(\mu)^2}}
\end{aligned}$$

When $\mu \notin D(I)$ or $\mu$ is local minima both $|\partial I|(\mu)$ and $|\partial \mathcal{G}_I|(\mu)$ are $+\infty$ and 0 respectively. Using Definition 2.5 of the weak metric slope, we have $|dI|(\mu) \leq |\partial I|(\mu)$.

To prove the other inequality, we first assume that $|\partial I|(\mu) =: \varepsilon_0 > 0$. Then, for any $\varepsilon \in (0, \varepsilon_0)$ exists $\delta = \delta(\varepsilon) > 0$ by Definition 4.5 such that there exists $\nu \in B_\delta(\mu)$ with

$$I(\nu) < I(\mu) - \varepsilon W_2(\mu, \nu)\,.$$

Choose such a $\nu$ and set $\delta' = W_2(\mu,\nu) < \delta$. Since $I$ is l.s.c, we find for any $n \in \mathbb{N}, n \geq 2$ some $\alpha = \alpha(\delta', n) > 0$ such that for all $\eta \in B_\alpha(\mu)$ it holds that

$$I(\mu) - I(\eta) \leq \frac{\delta'}{n}.$$

We define $\alpha' = \min\{\alpha, \delta'/n\}$ and define a map $\Psi : B_{\alpha'}(\mu) \times [0, \alpha'] \to \mathcal{P}(M)$ as follows

$$\Psi(\eta, t) = \Big(F_{\frac{t}{W_2(\nu,\eta)}}\Big)_{\#}\eta$$

where $F$ is the optimal map between $\eta$ and $\nu$ from Theorem 3.2. We have from the definition of $\alpha'$ by the triangle inequality

$$\frac{n-1}{n}\delta' \leq -W_2(\mu,\eta) + W_2(\mu,\nu) \leq W_2(\nu,\eta) \leq W_2(\mu,\eta) + W_2(\mu,\nu) \leq \frac{n+1}{n}\delta'.$$

Thus it follows that $0 \leq t/W_2(\eta,\nu) \leq 1$. Also, by construction holds $W_2(\eta, \Psi(\eta,t)) = t$. Now, by the $\lambda$-convexity of $I$, we obtain the following estimate

$$\begin{aligned} I(\Psi(\eta,t)) &\leq I(\eta) + \frac{t}{W_2(\eta,\nu)}(I(\nu) - I(\eta)) + \frac{|\lambda|}{2}t\Big(1 - \frac{t}{W_2(\eta,\nu)}\Big)W_2(\eta,\nu) \\ &\leq I(\eta) + \frac{t}{W_2(\eta,\nu)}(I(\nu) - I(\mu)) + \frac{t}{W_2(\eta,\nu)}(I(\mu) - I(\eta)) + \frac{|\lambda|}{2}\, t\, \delta' \frac{n+1}{n} \\ &< I(\eta) + t\Big(-\varepsilon\Big(\frac{n}{n+1}\Big) + \frac{\delta'}{n-1} + \delta'|\lambda|\Big). \end{aligned}$$

We can pick $\delta' > 0$ to be as small and $n$ as large as we want and conclude $I(\Psi(\eta,t)) \leq I(\eta) - \varepsilon t$. It follows from Proposition 2.6 that $|dI|(\mu) \geq \varepsilon$. Thus, since $\varepsilon \in (0, \varepsilon_0)$ is arbitrary, we have that $|dI|(\mu) \geq \varepsilon_0 = |\partial I|(\mu)$ for all positive values. For the case in which $|dI|(\mu) = 0$, assume that $|\partial I|(\mu) = \varepsilon_0 > 0$. But we have shown that $|\partial I|(u) = |\partial I|(\mu) = \varepsilon_0 > 0$ which would be a contradiction. Thus we have $|\partial I|(\mu) = |\partial I|(\mu)$. □

**Proposition 4.9.** *Let $I : \mathcal{P}(M) \to \mathbb{R} \cup \{+\infty\}$ be a proper, l.s.c, and $\lambda$-geodesically convex functional. Then* epi$(I)$ *is complete and path-connected.*

*Proof.* We note that $\mathcal{P}(M) \times \mathbb{R}$ is complete. Also, given any convergent sequence $(\mu_n, \xi_n) \in$ epi$(I)$ converging to some $(\mu, c) \in \mathcal{P}(M) \times \mathbb{R}$ we have that $I(\mu) \leq \liminf I(\mu_n) \leq \liminf \xi_n = c$. Thus epi$(I) \subset \mathcal{P}(M) \times \mathbb{R}$ is closed and thus complete.

Let $(\mu, \alpha), (\nu, \beta) \in$ epi$(I)$. Then $(\mu_t, (1-t)\alpha + t\beta - \frac{\lambda}{2}t(1-t)W_2^2(\mu,\nu))$ is a continuous path (w.r.t $d_{\text{epi}}$) between them which lies entirely in epi$(I)$. □

We conclude this section with the proof of Theorem 1.1.

*Proof of Theorem* 1.1. Denote by $\Gamma_{\text{epi}}$ the set of all continuous curves $\gamma_{\text{epi}} : [0,1] \to \text{epi}(I)$ with $\gamma_{\text{epi}}(0) = (\mu, I(\mu))$ and $\gamma_{\text{epi}}(1) = (\nu, I(\nu))$. We can identify any $\gamma_{\text{epi}} \in \Gamma_{\text{epi}}$ with a $\widetilde{\gamma} \in \Gamma$ by projecting onto the first factor, i.e., $(t \mapsto (\mu_t, \xi)) \mapsto (t \mapsto \mu_t)$. Because of the

definition of the epigraph and $\mathcal{G}_I$, we have that

$$\inf_{\gamma_{\mathrm{epi}}\in\Gamma_{\mathrm{epi}}}\max_{t\in[0,1]}\mathcal{G}_I(\gamma_{\mathrm{epi}}(t))\geq\inf_{\gamma_{\mathrm{epi}}\in\Gamma_{\mathrm{epi}}}\max_{t\in[0,1]}I(\widetilde{\gamma}(t))\geq\inf_{\gamma\in\Gamma}\max_{t\in[0,1]}I(\gamma(t))=c\,.$$

Also we have that $c_1=\max\{I(\mu),I(\nu)\}=\max\{\mathcal{G}_I(\mu,I(\mu)),\mathcal{G}_I(\nu,I(\nu))\}$ and from the above inequality that $c':=\inf_{\gamma_{\mathrm{epi}}\in\Gamma_{\mathrm{epi}}}\max_{t\in[0,1]}\mathcal{G}_I(\gamma_{\mathrm{epi}}(t))\geq c>c_1$. Furthermore, if $I$ satisfies (**MPS**), it follows that $\mathcal{G}_I$ satisfies it as well. Let $(\mu_n,\xi_n)$ be a Palais sequence. Since $|d\mathcal{G}_I|(\mu_n,\xi_n)\to 0$ it follows from Lemma 4.2 that for $n$ large enough the sequence must be of the form $(\mu_n,I(\mu_n))$ and that $|dI|(\mu_n)\to 0$. Since $\mathcal{G}_I((\mu_n,I(\mu_n)))=I(\mu_n)\to c$, it follows that $\mu_n$ is a Palais sequence for $I$. Thus we can construct a subsequence which converges to some $\mu^*\in\mathcal{P}(M)$ and by extension to $(\mu^*,c)\in\mathrm{epi}(I)$. Finally we can apply Theorem 2.3 to $\mathcal{G}_I$ to extract the existence of a critical point $(\eta,c')\in\mathrm{epi}(I)$ such that $|d\mathcal{G}_I|(\eta,c')=0$ and $\mathcal{G}_I((\eta,c'))=c'=\inf_{\gamma_{\mathrm{epi}}\in\Gamma_{\mathrm{epi}}}\Upsilon(\gamma)$. However, the contraposition of Lemma 4.2 implies that $c'=I(\eta)$ if $|d\mathcal{G}_I|(\eta,c')<1$, from which it follows that $|dI|(\eta)=|d\mathcal{G}_I|(\eta,I(\eta))=0$. Thus $\eta$ is critical point of $I$ with critical value $c'$. Also, since $I$ is $\lambda$-geodesically convex it follows from Proposition 4.8 that $|\partial I|(\eta)=0$. □

## 5. Application to the McKean–Vlasov model

The first part of the section analyses the free energy $I$ (1.3) and provides parameter values $\beta$ at which it has two distinct minimisers and then apply Theorem 1.1 to arrive at Theorem 1.2. In the second part, we formulate the large deviations results and complete the proof of Theorem 1.3. Let us recall some of the main definitions and results about the free energy functional from [CP10] and [CGPS18].

**Definition 5.1** (Transition point)**.** *A parameter value $\beta_c>0$ is said to be a* transition point *of $I$ from the uniform measure $\mu_0(\mathrm{d}x)=\mathrm{d}x/L^d$ if it satisfies the following conditions:*

*(1) For $0<\beta<\beta_c$, $\mu_0$ is the unique minimiser of $I$ .*
*(2) For $\beta=\beta_c$, $\mu_0$ is a minimiser of $I$ .*
*(3) For $\beta>\beta_c$, there exists $\mu_\beta\in\mathcal{P}(\mathbb{T}^d_L)\setminus\{\mu_0\}$, such that $\mu_\beta$ is a minimiser of $I$ .*

*Additionally, a transition point $\beta_c>0$ is said to be a* continuous transition point *of $I$ if:*

*(1) For $\beta=\beta_c$, $\mu_0$ is the unique minimiser of $I$ .*
*(2) Given any family of minimisers $\{\mu_\beta|\beta>\beta_c\}$, we have that*

$$\limsup_{\beta\downarrow\beta_c}\|\mu_\beta-\mu_0\|_{TV}=0\,.$$

*A transition point $\beta_c$ which is not continuous is said to be* discontinuous.

In thermodynamics, continuous phase transitions correspond to second-order ones similar to the those seen in the theory of magnetisation and spin systems [Daw83, Shi87,

GP18], whereas discontinuous phase transitions correspond to first-order ones similar to the ones observed in nucleation processes or phase transformation from liquid to vapor [LP66].

One can show that if $\beta_c$ is discontinuous, i.e., if either one of the above two conditions are violated, then it must be the case that both of them are violated. The original statement of these definitions and the proof of the above statement can be found in [CP10]. We summarise the main results about the free energy functional in the theorem below. The proofs can be found in [CP10] and [CGPS18, Theorems 5.11 and 5.19]. The conditions are expressed in terms of the Fourier coefficients of $W$ denoted by

(5.1)
$$\hat{W}(k) = \int_{\mathbb{T}_L^d} e_k(x) W(x)\,\mathrm{d}x \quad \text{with} \quad e_k = L^{-d/2}\exp\left(\frac{2\pi\, i}{L}\, k\cdot x\right) \quad \text{for} \quad k\in\mathbb{Z}^d\,.$$

**Definition 5.2** (Stable potential)**.** *A function $W \in L^2(\mathbb{T}_L^d)$ is said to be $H$-*stable*, denoted by $W \in \mathbb{H}_s$, if it has non-negative Fourier coefficients, that is $\hat{W}(k) \geq 0$ for all $k \in \mathbb{Z}^d$. If $W$ is not $H$-stable, is called $H$-*unstable *and denoted by $W \in \mathbb{H}_s^c$.*

**Theorem 5.3.** *Assume that $W \in C^2(\mathbb{T}_L^d)$ and $\beta > 0$.*

(*a*) *The free energy function $I : \mathcal{P}(\mathbb{T}_L^d) \to \mathbb{R}\cup\{+\infty\}$ always has a minimiser $\mu \in \mathcal{P}(\mathbb{T}_L^d)$ such that $\mu \ll \mathrm{d}x$ and $\frac{\mathrm{d}\mu}{\mathrm{d}x} > 0$ for all $x \in \mathbb{T}_L^d$.*

(*b*) *If $W \in \mathbb{H}_s^c$, that is there exists $k \in \mathbb{Z}^d\setminus\{0\}$ such that $\hat{W}(k) < 0$, then there exists a $\beta_c > 0$ such that $\beta_c$ is a transition point of $I$.*

(*c*) *For $W \in \mathbb{H}_s^c$, the set $K^\delta$ is given for any $\delta > 0$ by*

$$K^\delta := \left\{ k \in \mathbb{Z}^d, k \neq 0 : \hat{W}(k) \leq \min_{k\in\mathbb{Z}^d, k\neq 0} \hat{W}(k) + \delta \right\},$$

*Let $\delta_* > 0$ be the smallest $\delta$ for which there exist distinct $k^a, k^b, k^c \in K^{\delta_*}$ with $k^a = k^b + k^c$, if such points exist, else $\delta_* = \infty$. If $\delta_*$ is sufficiently small, then $\beta_c$ is a discontinuous transition point and $\beta_c < \frac{L^{d/2}}{\min_{k\in\mathbb{Z}^d,k\neq 0}\hat{W}(k)}$.*

(*d*) *Now, let $\{W_n\}_{n\in\mathbb{N}} \in C^2(\mathbb{T}_L^d)\cap\mathbb{H}_s^c$ be a sequence of interaction potentials such that $\delta_* \to 0$ as $n\to\infty$. Assume there exists $N\in\mathbb{N}$ and a positive constant $C > 0$ such that for all $n > N$, $\left|\min_{k\in\mathbb{Z}^d,k\neq 0}\hat{W}_n(k)\right| > C\delta_*^\gamma$ for any $\gamma < \frac{1}{2}$. Then for $n$ sufficiently large, $\beta_c$ is a discontinuous transition point and $\beta_c < \frac{L^{d/2}}{\min_{k\in\mathbb{Z}^d,k\neq 0}\hat{W}_n(k)}$.*

The above result provides conditions when to expect a discontinuous transition point. The case of the discontinuous transition point is particularly interesting for us as it implies the existence of a parameter value $\beta_c$ at which there are two distinct minimisers and hints at a possible scenario in which the mountain pass theorem could be applied. To provide more intuition we show in the following lemma that any potential that under

rescaling localises sufficiently fast but loses mass sufficiently slow will eventually exhibit a discontinuous transition point for the associated free energy $I$.

**Lemma 5.4.** *Let $W \in C^2(\mathbb{T}_L^d)$ be a compactly supported interaction potential with support strictly contained in $\mathbb{T}_L^d$ and $\int_{\mathbb{T}_L^d} W \,\mathrm{d}x < 0$. Assume further that*

$$\int_{\mathbb{T}_L^d} L^{-d/2} W(x) e^{i\frac{2\pi\epsilon k\cdot x}{L}} \,\mathrm{d}x \geq L^{-d/2} \int_{\mathbb{T}_L^d} W \,\mathrm{d}x := -C \text{ for all } k \in \mathbb{Z}^d, \epsilon > 0, \tag{5.2}$$

*Consider the rescaled potential, $W_\epsilon(x) = f(\epsilon) W(x/\epsilon)$ for some $\epsilon > 0$ and positive function $f : \mathbb{R}_+ \to \mathbb{R}_+$. If $\epsilon^\ell \lesssim f(\epsilon) \lesssim \epsilon^m$ as $\epsilon \to 0$ for $m > -d-2$, $\ell \geq -d$, $\ell < \frac{m-d}{2} + 1$ (along with the natural restriction $\ell \geq m$), then for $\epsilon$ small enough, the associated free energy $I$ possesses a discontinuous transition point at some $\beta_c < \frac{L^{d/2}}{\min_{k\in\mathbb{Z}^d, k\neq 0} \hat{W}_\epsilon(k)}$.*

*Proof.* We proceed by checking that the conditions of Theorem 5.3(d) hold for this class of potentials. We first check that for $\epsilon$ small enough, $W_\epsilon \in \mathbb{H}_s^c$. Let $V := \operatorname{supp} W$ and $V_\epsilon := \operatorname{supp} W_\epsilon$. We have for $k \in \mathbb{Z}^d$,

$$\begin{aligned}
\hat{W}_\epsilon(k) &= L^{-d/2} \int_{\mathbb{T}_L^d} W_\epsilon(x) e^{i\frac{2\pi k\cdot x}{L}} \,\mathrm{d}x \\
&= L^{-d/2} f(\epsilon) \int_{V^\epsilon} W(x/\epsilon) e^{i\frac{2\pi k\cdot x}{L}} \,\mathrm{d}x \\
&= L^{-d/2} f(\epsilon) \epsilon^d \int_V W(x) e^{i\frac{2\pi\epsilon k\cdot x}{L}} \,\mathrm{d}x\,. \qquad (5.3)
\end{aligned}$$

Since $e^{i\frac{2\pi\epsilon k\cdot x}{L}} \to 1$ uniformly on $V$ as $\epsilon \to 0$, it follows that eventually $\hat{W}_\epsilon(k) < 0$ for $\epsilon$ sufficiently small since $\int_V W(x) e^{i\frac{2\pi\epsilon k\cdot x}{L}} \,\mathrm{d}x \to \left(\int_{\mathbb{T}_L^d} W \,\mathrm{d}x\right) < 0$. Using (5.2) and (5.3), we can now obtain the following bound

$$\min_{k\in\mathbb{Z}^d, k\neq 0} \hat{W}_\epsilon(k) \geq -C f(\epsilon) \epsilon^d\,.$$

Since $W$ is even along every coordinate we have that

$$\begin{aligned}
\int_{\mathbb{T}_L^d} W(x) e^{i\frac{2\pi\epsilon k\cdot x}{L}} \,\mathrm{d}x &= \int_{\mathbb{T}_L^d} W(x) \cos\left(\frac{2\pi\epsilon k\cdot x}{L}\right) \mathrm{d}x \\
&= \int_{\mathbb{T}_L^d} W(x) \left(1 + \left(\frac{2\pi|k|\epsilon x}{L}\right)^2 + O(\epsilon^4)\right) \mathrm{d}x
\end{aligned}$$

Fix some $k^* \in \mathbb{Z}^d$.The above expansion tells us that we can find some $\epsilon_1$ sufficiently small and some $C_1 > 0$ independent of $\epsilon$ such that

$$\hat{W}_\epsilon(k^*), \hat{W}_\epsilon(2k^*) \leq f(\epsilon)\epsilon^d(-C + C_1\epsilon^2) \text{ for all } \epsilon < \epsilon_1\,.$$

We can thus obtain the following bound

$$-Cf(\epsilon)\epsilon^d \leq \min_{k\in\mathbb{Z}^d, k\neq 0} \hat{W}_\epsilon(k) \leq f(\epsilon)\epsilon^d(-C + C_1\epsilon^2) \text{ for all } \epsilon < \epsilon_1\,. \tag{5.4}$$

Combining the two of them we derive

$$\left.\begin{aligned}\hat{W}_\epsilon(k^*) - \min_{k\in\mathbb{Z}^d, k\neq 0}\hat{W}_\epsilon(k) &\leq C_1 f(\epsilon)\epsilon^{2+d}\\ \hat{W}_\epsilon(2k^*) - \min_{k\in\mathbb{Z}^d, k\neq 0}\hat{W}_\epsilon(k) &\leq C_1 f(\epsilon)\epsilon^{2+d}\end{aligned}\right\} \text{ for all } \epsilon < \epsilon_1\,,$$

which tells us that $k^*, 2k^* \in K^{C_1 f(\epsilon)\epsilon^{2+d}}$ and that $\delta_* \leq C_1 f(\epsilon)\epsilon^{2+d}$. Thus $\delta_* \lesssim \epsilon^{m+d+2}$ and since $m > -d-2$, $\delta_* \to 0$ as $\epsilon \to 0$. Furthermore, using (5.4) we can deduce

$$\left|\min_{k\in\mathbb{Z}^d, k\neq 0}\hat{W}_\epsilon(k)\right| \geq f(\epsilon)\epsilon^d(C - C_1\epsilon^2) \geq C_2\epsilon^{\ell+d}\,.$$

The fact that $\ell \geq -d$ tells us that

$$\left|\min_{k\in\mathbb{Z}^d, k\neq 0}\hat{W}_\epsilon(k)\right| \geq C_3\delta_*^{\frac{\ell+d}{m+d+2}}\,.$$

We now use the assumption that $l < \frac{m-d}{2} + 1$ and apply Theorem 5.3(d), to obtain the desired result. □

Now that we have a set of concrete conditions under which we can expect there to be two distinct minimisers at a particular parameter value, we can try apply the mountain pass theorem. To apply Theorem 1.1, it is sufficient to show that $\mu_0$ is a strict local minima at parameter values $\beta_c$ and that this property is uniform, i.e., we can find a ball $B_r^{W_2}(\mu_0)$ around $\mu_0$ in $W_2$ such that $I(\mu) \geq I(\mu_0)+\delta$ for all $\mu \in \partial B_r^{W_2}(\mu_0)$ and some $\delta > 0$. In order to show this, we need the following comparison of $W_2$ with the homogeneous negative Sobolev space $\dot{H}^{-1}(\mathbb{T}_L^d)$, which we identify with all formal Fourier series of $\mu$ as defined in (5.1) given by $\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\hat{\mu}(k)e_k$ such that $\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\frac{1}{|k|^2}|\hat{\mu}(k)|^2 < \infty$. Note that the functions $\{e_k\}_{k\in\mathbb{Z}^d\setminus\{0\}}$ form an orthogonal basis for $\dot{H}^{-1}(\mathbb{T}_L^d)$ with respect to the inner product defined by duality with the homogeneous space $\dot{H}^1(\mathbb{T}_L^d)$. We have that $\langle\mu, f\rangle_{\dot{H}^{-1},\dot{H}_0^1} = \sum_{k\in\mathbb{Z}^d\setminus\{0\}}\hat{\mu}(k)\hat{f}(k)$. Also, the inner product on $\dot{H}_0^1(\mathbb{T}_L^d)$ is defined as $(f,g)_{\dot{H}_0^1} = \sum_{k\in\mathbb{Z}^d\setminus\{0\}}|k|^2\hat{f}(k)\hat{g}(k)$. It is easy to check then that the Riesz representation of any $\mu \in \dot{H}^{-1}(\mathbb{T}_L^d)$ is given by $\sum_{k\in\mathbb{Z}^d\setminus 0}\frac{1}{|k|^2}\hat{\mu}(k)e_k \in \dot{H}_0^1(\mathbb{T}_L^d)$.

**Lemma 5.5** (Comparison of $\dot{H}^{-1}$ with $W_2$)**.** *Let $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{T}_L^d)\cap L^\infty(\mathbb{T}_L^d)$. Then the following estimate holds*

$$\|\mu_0 - \mu_1\|_{\dot{H}^{-1}(\mathbb{T}_L^d)} \leq \Big(\max\Big[\|\mu_0\|_{L^\infty(\mathbb{T}_L^d)}, \|\mu_1\|_{L^\infty(\mathbb{T}_L^d)}\Big]\Big)^{1/2} W_2(\mu_0,\mu_1)$$

*Proof.* The proof follows the argument in [Loe06, Proposition 2.1]. Let $\mu_t$ be the optimal interpolant between $\mu_0$ and $\mu_1$ from Theorem 3.2. Then by the Benamou–Brenier formulation of the optimal transport problem, there exists a vector field $v_t \in L^2(\mu_t;\mathbb{R}^d)$

such that the pair $(\mu_t, v_t)$ satisfies

$$\partial_t \mu_t + \nabla \cdot (\mu_t v_t) = 0\,,$$

in the sense of distributions. Now consider the sequence of parameterised problems given by

$$\Delta \Psi_t = \mu_t\,.$$

Note that $\|\mu_t\|_{L^\infty(\mathbb{T}^d_L)} \leq \max\Big\{\|\mu_0\|_{L^\infty(\mathbb{T}^d_L)}, \|\mu_1\|_{L^\infty(\mathbb{T}^d_L)}\Big\}$ [Vil08, Corollary 17.19], and thus the above equation has a unique weak solution in $\dot{H}^1(\mathbb{T}^d_L)$ for all $t \in [0,1]$. We know that $\int_{\mathbb{T}^d_L} |v_t|^2 \mu_t \,\mathrm{d}x = W_2^2(\mu_t, \mu_0) = t^2 W_2^2(\mu_0, \mu_1) < \infty$. From this it follows that $\mu_t v_t \in L^2(\mathbb{T}^d_L; \mathbb{R}^d)$ and thus $\nabla \cdot (\mu_t v_t) \in \dot{H}^{-1}(\mathbb{T}^d_L)$. Differentiating w.r.t to $t$ we have

$$\Delta \partial_t \Psi_t = -\nabla \cdot (\mu_t v_t)\,.$$

It follows then that $\partial_t \Psi_t \in \dot{H}^1(\mathbb{T}^d_L)$. Multiplying by $\partial_t \Psi_t$ and integrating by parts with respect to the space variable and then integrating w.r.t to time, we obtain

$$\begin{aligned}\|\nabla \Psi_1 - \nabla \Psi_0\|_{L^2(\mathbb{T}^d_L)} &\leq \|\mu_t\|_\infty^{1/2} W_2(\mu_0, \mu_1) \\ &\leq \Big(\max\Big[\|\mu_0\|_{L^\infty(\mathbb{T}^d_L)}, \|\mu_1\|_{L^\infty(\mathbb{T}^d_L)}\Big]\Big)^{1/2} W_2(\mu_0, \mu_1)\end{aligned}$$

Since $\Psi_t$ is precisely the Riesz representation of $\mu_t$ in $\dot{H}^1(\mathbb{T}^d_L)$, the estimate follows. □

The following lemma now establishes the strictness of a local minima in $W_2$ for discontinuous transitions points.

**Lemma 5.6.** *Assume $W \in \mathbb{H}^c_s \cap C^2(\mathbb{T}^d_L)$ and that $\beta_c$ is a discontinuous transition point. There exists some $r > 0$ such that for $\beta \leq \beta_c$ the measure $\mu_0 = (1/L^d)\,\mathrm{d}x$ is a strict local minima of $I$ and the following estimate holds*

$$I(\mu) \geq I(\mu_0) + \delta$$

*for all $\mu \in \partial B_r^{W_2}(\mu_0)$ for $r$ sufficiently small and some $\delta > 0$.*

*Proof.* By the definition of $\beta_c$ from Definition 5.1, we have that for $\beta \leq \beta_c$, $\mu_0$ is a minimiser of $I$. The proof for $\beta < \beta_c$ is obvious. The idea of the proof is based on the fact that any minimiser of the free energy must be a solution of $T(\mu) = \mu - F(\mu) = 0$ (cf. [CGPS18, Proposition 2.4]), where $F : \dot{H}^{-1}(\mathbb{T}^d_L) \to \dot{H}^{-1}(\mathbb{T}^d_L)$ is the map given by

$$F(\mu) = \exp\left(-\beta W \star \mu - \log \int_{\mathbb{T}^d_L} \exp(-\beta W \star \mu)\,\mathrm{d}x\right).$$

The fact that not all elements of $\mathcal{P}(\mathbb{T}^d_L)$ lie in $\dot{H}^{-1}(\mathbb{T}^d_L)$ does not affect the validity above statement as elements in $\mathcal{P}(\mathbb{T}^d_L) \setminus \dot{H}^{-1}(\mathbb{T}^d_L)$ are not minimisers of the free energy. It is possible to check now that, for $\beta \leq \beta_c$, $DT(\mu_0) : \dot{H}^{-1}(\mathbb{T}^d_L) \to \dot{H}^{-1}(\mathbb{T}^d_L)$ is a bounded,

linear, isomorphism. Indeed, we have that

$$DT(\mu_0)(\eta) = \eta - \beta\mu_0(W \star \eta) - \beta\mu_0^2 \int_{\mathbb{T}_L^d} W \star \eta \, \mathrm{d}x$$

The above operator is bounded on $\dot{H}^{-1}(\mathbb{T}_L^d)$ since $W \in C^2(\mathbb{T}_L^d) \supset \dot{H}^1(\mathbb{T}_L^d)$. Diagonalising $DT(\mu_0)$ using $\{e_k\}_{k\in\mathbb{Z}^d\setminus\{0\}}$, we obtain

$$DT(\mu_0)e_k = (1 - \beta L^{-d/2}\hat{W}(k))e_{-k}\,.$$

It follows then that if $\beta \leq L^{d/2}/\min_{k\in\mathbb{Z}^d\setminus\{0\}} \hat{W}(k)$, then the above map is a bijection. That it is an injection is clear from the fact that if $DT(\mu_0)\eta_1 = DT(\mu_0)\eta_2$ then $\hat{\eta_1}(k) = \hat{\eta_2}(k)$ for all $k \in \mathbb{Z}^d \setminus \{0\}$. It is also surjective since for any $\eta \in \dot{H}^{-1}(\mathbb{T}_L^d)$, we have that $\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{\hat{\eta}(k)}{(1-\beta L^{-d/2}\hat{W}(-k))} e_{-k}$ maps to $\eta$ under $DT(\mu_0)$. We know from Theorem 5.3(c) that $\beta_c$ is lesser than this value and hence the result. Now, by the inverse function theorem, there exists an $\varepsilon$-open ball $B_\varepsilon^{\dot{H}^{-1}}(\mu_0)$ around $\mu_0$ in $\dot{H}^{-1}(\mathbb{T}_L^d)$ such that it is the unique solution of $T(\mu) = 0$ in this ball. This tells us that $\mu_0$ is the unique minimiser of the free energy in $B_\varepsilon^{\dot{H}^{-1}}(\mu_0)$ at $\beta = \beta_c$. Note further that we have the following bounds for all $\mu \in B_\varepsilon^{\dot{H}^{-1}}(\mu_0)$

$$\mu_0 \exp\Big(-2\beta\|W\|_{\dot{H}^1(\mathbb{T}_L^d)}\|\mu\|_{\dot{H}^{-1}(\mathbb{T}_L^d)}\Big) \leq F(\mu) \leq \mu_0 \exp\Big(2\beta\|W\|_{\dot{H}^1(\mathbb{T}_L^d)}\|\mu\|_{\dot{H}^{-1}(\mathbb{T}_L^d)}\Big)\,.$$

Additionally we have that $\|\mu - \mu_0\|_{\dot{H}^{-1}(\mathbb{T}_L^d)} < \varepsilon$ from which it follows that

$$\frac{\mu_0}{C} \leq F(\mu) \leq C\,\mu_0 \qquad \text{with} \qquad C := \exp\Big(2\beta\|W\|_{\dot{H}^1(\mathbb{T}_L^d)}(\|\mu_0\|_{\dot{H}^{-1}(\mathbb{T}_L^d)} + \varepsilon)\Big)\,,$$

for all $\mu \in B_\varepsilon^{\dot{H}^{-1}}(\mu_0)$. Consider the set

$$\mathcal{I} := \left\{\mu \in \dot{H}^{-1}(\mathbb{T}_L^d) \cap \mathcal{P}(\mathbb{T}_L^d) \cap L^1(\mathbb{T}_L^d) : \frac{\mu_0}{C} \leq \mu \leq C\mu_0\right\}.$$

Then, any minimiser of $I$ must lie in $\mathcal{I}$ by construction. Additionally, for all $\mu \in \mathcal{I}$ we have from Lemma 5.5 for some fixed constant $C_0 = C_0(\mu_0, C)$ the bound

$$\|\mu_0 - \mu\|_{\dot{H}^{-1}(\mathbb{T}_L^d)} \leq C_0 W_2(\mu_0, \mu)\,.$$

We can thus pick a ball $B_r^{W_2}(\mu_0)$ sufficiently small such that $\|\mu - \mu_0\|_{\dot{H}^{-1}(\mathbb{T}_L^d)} < \varepsilon$ for all $\mu \in B_r^{W_2}(\mu_0) \cap \mathcal{I}$. It thus follows that we can find a ball in $W_2$ for which $\mu_0$ is the unique minimiser of $I$. Thus $\mu_0$ is a strict local minima in $\mathcal{P}(\mathbb{T}_L^d)$ equipped with the Wasserstein metric.

Consider now the boundary of the ball $\partial B_r^{W_2}(\mu_0)$. This is a compact set and since $I$ is l.s.c we can find a minimiser on this set, say $\mu^*$. Setting $\delta = I(\mu^*) - I(\mu_0) > 0$ the estimate in the lemma follows. □

We can now prove the existence of a mountain pass point.

**Theorem 5.7.** *Assume $W \in \mathbb{H}_s^c \cap C^2(\mathbb{T}_L^d)$ and $\beta_c$ is a discontinuous transition point, i.e., there exist at least two distinct minimisers of $I$ at $\beta = \beta_c$ such that one is $\mu_0$ and the other is some $\overline{\mu} \in \mathcal{P}(\mathbb{T}_L^d)$. It follows then that there exists a $\mu^* \in \mathcal{P}(\mathbb{T}_L^d)$ distinct from $\mu_0$ and $\overline{\mu}$ such that $|\partial I|(\mu^*) = |dI|(\mu^*) = 0$. Additionally, $I(\mu^*) = c'$ where $c'$ is characterised as follows*

$$c' \geq c = \inf_{\gamma \in \Gamma} \max_{t \in [0,1]} I(\gamma(t))\,,$$

*where* $\Gamma = \{C([0,1]; \mathcal{P}(\mathbb{T}_L^d)) : \gamma(0) = \mu_0, \gamma(1) = \overline{\mu}\}$.

*Proof.* We have to check that the conditions of Theorem 1.1 hold. $I$ is proper and l.s.c and since $\|D^2 W\|_{L^\infty(\mathbb{T}_L^d)} \leq C$ it is also $\lambda$-geodesically convex. Additionally, by the definition of $I$, $\mu \in D(I)$ implies $\mu \ll \mathrm{d}x$. The space $\mathcal{P}(\mathbb{T}_L^d)$ is compact and thus $I$ trivially satisfies **(MPS)**. We also know that $c_1 = I(\mu_0)$ and, applying Lemma 5.6, that $c = I(\mu_0) + \delta$. Thus we have that $c > c_1$ and the result follows. □

We state without proof the reformulated version of the main result from [DG87].

**Theorem 5.8.** *Denote by $\mathcal{P}^{(N)}(\mathbb{T}_L^d)$ the space of empirical probability measures on $\mathbb{T}_L^d$, i.e.,*

$$\mathcal{P}^{(N)}(\mathbb{T}_L^d) := \left\{ \mu \in \mathcal{P}(\mathbb{T}_L^d) : \mu = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}, x_i \in \mathbb{T}_L^d \right\}.$$

*Asume that $\mu_0^{(N)} \in \mathcal{P}^{(N)}(\mathbb{T}_L^d)$ is such that there exists $\mu_0 \in \mathcal{P}(\mathbb{T}_L^d)$ with $W_2(\mu_0^{(N)}, \mu_0) \to 0$ as $N \to \infty$. Denote by $\mathcal{C}_T$ the space $C(0,T; \mathcal{P}(\mathbb{T}_L^d))$, equipped with the topology of uniform convergence.*

(a) *For all open subsets $G$ of $\mathcal{C}_T$ holds*

$$\liminf_{N\to\infty} N^{-1} \log \mathbb{P}\Big(\mu^{(N)}(\cdot) \in G, \mu^{(N)}(0) = \mu_0^{(N)}\Big) \geq - \inf_{\mu(\cdot) \in G, \mu(0) = \mu_0} S(\mu(\cdot))\,.$$

(b) *For all closed subsets $F$ of $\mathcal{C}_T$ holds*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\Big(\mu^{(N)}(\cdot) \in F, \mu^{(N)}(0) = \mu_0^{(N)}\Big) \leq - \inf_{\mu(\cdot) \in F, \mu(0) = \mu_0} S(\mu(\cdot))\,,$$

(c) *For each compact subset $K$ of $\mathcal{P}(\mathbb{T}_L^d)$ and $s \geq 0$ is the set*

$$\Phi_K(s) = \{\mu(\cdot) \in \mathcal{C}_T : S(\mu(\cdot)) \leq s, \mu(0) \in K\}\,,$$

*compact.*

*Here $S : \mathcal{C}_T \to \mathbb{R} \cup \{+\infty\}$ is the action or rate functional given for $\mu \in AC^2(0,T; \mathcal{P}(\mathbb{T}_L^d))$ by*

$$S(\mu(\cdot)) = \frac{1}{2} \int_0^T \|\partial_t \mu - \nabla \cdot (\mu(\nabla \log \mu + \nabla W \star \mu))\|^2_{\dot{H}^{-1}(\mathbb{T}_L^d, \mu)} \,\mathrm{d}t\,.$$

*and by $+\infty$ otherwise.*

We are interested in using the above result to understand the probability of the empirical process escaping from the uniform state $\mu_0$ and reaching the clustered state $\overline{\mu}$ in time $T > 0$.

**Theorem 5.9.** *Assume $W \in \mathbb{H}^c_s \cap C^2(\mathbb{T}^d_L)$ and $\beta_c$ is a discontinuous transition point, i.e., there exist at least two distinct minimisers of $I$ at $\beta = \beta_c$ such that one is $\mu_0$ and the other is some $\overline{\mu} \in \mathcal{P}(\mathbb{T}^d_L)$. It follows then that the underlying empirical process $\mu^{(N)} \in \mathcal{C}_T$ with initial i.i.d uniformly distributed particles satisfies*

$$\mathbb{P}(\mu^N(T) \in \overline{B}^{W_2}_\varepsilon(\overline{\mu}), \mu^{(N)}(0) = \mu^{(N)}_0) \lesssim \exp\Big(-N(\Delta - O(\varepsilon^2))\Big)$$

*for $N$ sufficiently large, where $\overline{B}^{W_2}_\varepsilon(\overline{\mu})$ is the closed ball of size $\varepsilon > 0$ around $\overline{\mu}$, $\Delta := I(\mu^*) - I(\mu_0)$, where $\mu^*$ is the critical point defined in Theorem 5.7.*

*Proof.* In order to prove this result we need to relate the rate functional $S$ with the energy functional $I$. We can assume w.l.o.g that $\mu \in AC^2(0,T;\dot{H}^1(\mathbb{T}^d_L) \cap \mathcal{P}(\mathbb{T}^d_L))$ since $S(\mu) = +\infty$ otherwise. It follows then that there exists $\phi \in L^2(0,T;\dot{H}^1(\mathbb{T}^d_L,\mu))$ [San15, Theorem 5.14] such that

$$\partial_t \mu = \nabla \cdot (\mu \nabla \phi)\,,$$

where the above equation is satisfied in $\dot{H}^{-1}(\mathbb{T}^d_L,\mu)$. Thus for any $\mu \in AC^2(0,T;\dot{H}^1(\mathbb{T}^d_L) \cap \mathcal{P}(\mathbb{T}^d_L))$ we can rewrite the rate functional as follows

$$\begin{aligned}
S(\mu) &= \frac{1}{2}\int_0^T \|\partial_t \mu - \nabla\cdot(\mu(\nabla \log \mu + \nabla W \star \mu))\|^2_{\dot{H}^{-1}(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^T \|\phi - \log\mu - W\star\mu\|^2_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^T \|\phi\|^2_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t + \frac{1}{2}\int_0^T \|\log\mu + W\star\mu\|^2_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t \\
&\quad + \int_0^T \langle \log\mu + W\star\mu, \phi\rangle_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^T \|\phi\|^2_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t + \frac{1}{2}\int_0^T \|\log\mu + W\star\mu\|^2_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t \\
&\quad + \int_0^T \langle (\log\mu + W\star\mu), \partial_t\mu\rangle_{\dot{H}^1(\mathbb{T}^d_L,\mu),\dot{H}^{-1}(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t\,.
\end{aligned}$$

We choose the closed subset $F = \{\mu \in \mathcal{C}_T : \mu(T) \in \overline{B}^{W_2}_\varepsilon(\overline{\mu}), \mu(0) = \mu_0\}$ and we set $T^* = \arg\max_{t\in[0,T]}(I(\mu(t)) - I(\mu_0))$ if it is uniquely defined or pick any one if it is not. We can then rewrite the rate functional as follows

$$S(\mu) = \frac{1}{2}\int_0^{T^*} \|\phi\|^2_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t + \frac{1}{2}\int_0^{T^*} \|\log\mu + W\star\mu\|^2_{\dot{H}^1(\mathbb{T}^d_L,\mu)}\,\mathrm{d}t$$

$$
\begin{aligned}
&+ \int_0^{T^*} \langle (\log \mu + W \star \mu), \partial_t \mu \rangle_{\dot{H}^1(\mathbb{T}^d_L,\mu), \dot{H}^{-1}(\mathbb{T}^d_L,\mu)} \, \mathrm{d}t \\
&+ \frac{1}{2} \int_{T^*}^{T} \| \partial_t \mu - \nabla \cdot (\mu(\nabla \log \mu + \nabla W \star \mu)) \|^2_{\dot{H}^{-1}(\mathbb{T}^d_L,\mu)} \, \mathrm{d}t \\
\geq &\int_0^{T^*} \langle (\log \mu + W \star \mu), \partial_t \mu \rangle_{\dot{H}^1(\mathbb{T}^d_L,\mu), \dot{H}^{-1}(\mathbb{T}^d_L,\mu)} \, \mathrm{d}t \\
= &\max_{t \in [0,T]} \left( I(\mu(t)) - I(\mu_0) \right).
\end{aligned}
$$

The estimate implies the lower bound

$$
\inf_{\mu \in F} S(\mu) \geq \inf_{\mu \in F \cap AC^2} \max_{t \in [0,T]} \left( I(\mu(t)) - I(\mu_0) \right).
$$

At this point, we cannot apply Theorem 5.7 directly, since $F$ contains curves with varying endpoints not necessarily critical points. To handle this case, we define

$$
\begin{aligned}
F_{\mathrm{epi}} = \Big\{ &(\mu(\cdot), \xi(\cdot)) \in C([0,T]; \mathrm{epi}(I)) : (\mu(0), \xi(0)) = (\mu_0, I(\mu_0)), \\
&(\mu(T), \xi(T)) \in \textstyle\bigcup_{\mu \in \overline{B}^{W_2}_{\varepsilon}(\overline{\mu})} (\mu, I(\mu)) \Big\}
\end{aligned}
$$

If $\mu \in F \cap AC^2$, then the function $t \mapsto I(\mu(t))$ is absolutely continuous by [AGS08, Section 10.1.2 E.]. Thus the curve $(\mu(\cdot), I(\mu(\cdot)))$ lies in the set $F_{\mathrm{epi}}$ with $\max_{t \in [0,T]}(I(\mu(t)) - I(\mu_0)) = \max_{t \in [0,T]}(\mathcal{G}_I(\mu(t), I(\mu(t))) - \mathcal{G}_I(\mu_0, I(\mu_0)))$. Thus we have that

$$
\inf_{\mu \in F \cap AC^2} \max_{t \in [0,T]} \left( I(\mu(t)) - I(\mu_0) \right) \geq \inf_{\mu \in F_{\mathrm{epi}}} \max_{t \in [0,T]} \left( \mathcal{G}_I(\mu(t), \xi(t)) - \mathcal{G}_I(\mu_0, I(\mu_0)) \right)
$$

Now, we argue that if $\varepsilon$ is small enough the above quantity can be made arbitrarily close to $\Delta$. For doing so, we define $\delta > 0$ such that $\inf_{\mu \in F_{\mathrm{epi}}} \max_{t \in [0,T]} (\mathcal{G}_I(\mu(t), \xi(t)) - \mathcal{G}_I(\mu_0, I(\mu_0))) = \Delta - 2\delta$. Then, we find $(\widetilde{\mu}(t), \xi(t)) \in F_{\mathrm{epi}}$ with $\max_{t \in [0,T]} (\mathcal{G}_I(\widetilde{\mu}(t), \xi(t)) - \mathcal{G}_I(\mu_0, I(\mu_0))) \leq \Delta - \delta$ from which it follows that $I(\widetilde{\mu}(T)) - I(\mu_0) \leq \Delta - \delta$. Let

$$
\begin{aligned}
\Gamma_{\mathrm{epi}} = \Big\{ &(\mu(\cdot), \xi(\cdot)) \in C([0,T], \mathrm{epi}(I)) : (\mu(0), \xi(0)) = (\mu_0, I(\mu_0)), \\
&(\mu(T), \xi(T)) = (\overline{\mu}, I(\overline{\mu})) \Big\} \subset F_{\mathrm{epi}}
\end{aligned}
$$

We know that $\inf_{\mu \in \Gamma_{\mathrm{epi}}} \max_{t \in [0,T]} (\mathcal{G}_I(\mu(t), \xi(t)) - \mathcal{G}_I(\mu_0, I(\mu_0))) = \Delta$. Thus if we take any continuous curve $(\mu(s), \xi(s))$ in $\mathrm{epi}(I)$ from $(\widetilde{\mu}(T), I(\widetilde{\mu}(T)))$ to $(\overline{\mu}, I(\overline{\mu}))$ parametrised by $s \in [0,1]$, $\mathcal{G}_I(\cdot, \cdot)$ must exceed or be equal to $I(\overline{\mu}) + \Delta$ at some $s \in [0,1]$. Indeed, if this would not be the case then we could concatenate $(\widetilde{\mu}(t), \xi(t))$ and $(\mu(s), \xi(s))$ to obtain, after reparametrisation, a new curve $[0,1] \ni t \mapsto (\mu(t), \xi(t))$ in $\mathrm{epi}(I)$ from $(\mu_0, I(\mu_0))$ to $(\overline{\mu}, I(\overline{\mu}))$ such that $\max_{t \in [0,1]} \mathcal{G}_I(\mu(t), \xi(t)) < \Delta$, a contradiction, since this curve is also an element of $\Gamma_{\mathrm{epi}}$.

We pick the curve $(\mu_t, I(\mu_t))$ where $\mu_t$ is the optimal interpolant between $\widetilde{\mu}(T)$ and $\overline{\mu}$ as defined in Definition 3.3. Let $t'$ be the time at which $I(\mu_{t'})$ exceeds $I(\overline{\mu}) + \Delta$. By

$\lambda$-convexity of $I$ we have

$$I(\overline{\mu}) + \Delta \leq I(\mu_{t'}) \leq (1-t')I(\widetilde{\mu}(T)) + t'I(\overline{\mu}) + \frac{|\lambda|}{2}t'(1-t')\varepsilon^2$$

Bounding $I(\widetilde{\mu}(T))$ by $I(\overline{\mu}) + \Delta - \delta$, we obtain,

$$\begin{aligned} I(\overline{\mu}) + \Delta &\leq I(\overline{\mu}) + (1-t')\Delta - (1-t')\delta + \frac{|\lambda|}{2}t'(1-t')\varepsilon^2 \\ &\leq I(\overline{\mu}) + \Delta - (1-t')\delta + \frac{|\lambda|}{2}(1-t')\varepsilon^2 \,. \end{aligned}$$

From this it follows that

$$\delta \leq \frac{|\lambda|}{2}\varepsilon^2 \,.$$

Thus we obtain

$$\inf_{\mu\in F} S(\mu) \geq \inf_{\mu\in F_{\mathrm{epi}}} \max_{t\in[0,T]} \big(\mathcal{G}_I(\mu(t),\xi(t)) - \mathcal{G}_I(\mu_0, I(\mu_0))\big) = \Delta - 2\delta \geq \Delta - |\lambda|\varepsilon^2$$

Finally, we can apply the result of Theorem 5.8 (b), to obtain that

$$\limsup_{N\to\infty} N^{-1}\log \mathbb{P}\Big(\mu^{(N)}(\cdot)\in F, \mu^{(N)}(0) = \mu_0^{(N)}\Big) \leq -\inf_{\mu(\cdot)\in F, \mu(0)=\mu_0} S(\mu(\cdot)) \leq -\Delta + |\lambda|\varepsilon^2 \,.$$

The result then follows from the above estimate, where we use that $W_2(\mu_0^{(N)}, \mu_0) \to 0$ as $N\to\infty$ is implied by the strong law of large numbers. □

**Remark 5.10.** *The $O(\varepsilon^2)$ appearing in the exponent $\exp\big(-N(\Delta - O(\varepsilon^2))\big)$ can be removed if one can show that the minimiser $\overline{\mu}$ is a local basin of attraction for the McKean–Vlasov dynamics, i.e., there exists some $\varepsilon > 0$ such that all measures in $\overline{B}_\varepsilon^{W_2}(\overline{\mu})$ converge to $\overline{\mu}$ under the flow of the McKean–Vlasov PDE as $t\to\infty$. In this case we can choose the continuous curve between $\widetilde{\mu}(T)$ and $\overline{\mu}$ (in the proof of Theorem 5.9) to be the solution of the McKean–Vlasov PDE starting $\widetilde{\mu}(T)$. This solution does not increase the energy and thus the $O(\varepsilon^2)$ error from the $\lambda$-convexity argument will not appear in the exponent. Such a characterisation of $\overline{\mu}$ is expected under more specific assumptions on the potential $W$.*

**Acknowledgements.** The authors would like to thanks José A. Carrillo and Greg Pavliotis for useful discussions during the course of this work.

## References

[ADPZ11] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Comm. Math. Phys.*, 307(3):791–815, 2011.

[ADPZ13] S. Adams, N. Dirr, M. Peletier, and J. Zimmer. Large deviations and gradient flows. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 371(2005):20120341, 17, 2013.

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[BB18] K. Bashiri and A. Bovier. Gradient flow approach to local mean-field spin systems. *arXiv e-prints*, page arXiv:1806.07121, Jun 2018.

[CEMS01] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.

[CGPS18] J. A. Carrillo, R. S. Gvalani, G. A. Pavliotis, and A. Schlichting. Long-time behaviour and phase transitions for the McKean–Vlasov equation on the torus. *arXiv e-prints*, page arXiv:1806.01719, Jun 2018.

[CP10] L. Chayes and V. Panferov. The McKean-Vlasov equation in finite volume. *J. Stat. Phys.*, 138(1-3):351–380, 2010.

[Daw83] D. A. Dawson. Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Stat. Phys.*, 31(1):29–85, apr 1983.

[DG87] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[DGMT80] E. De Giorgi, A. Marino, and M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3):180–187, 1980.

[DM94] M. Degiovanni and M. Marzocchi. A critical point theory for nonsmooth functionals. *Ann. Mat. Pura Appl. (4)*, 167:73–100, 1994.

[DPZ13] M. H. Duong, M. A. Peletier, and J. Zimmer. GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to large-deviation principles. *Nonlinearity*, 26(11):2951–2971, 2013.

[DS10] S. Daneri and G. Savaré. Lecture Notes on Gradient Flows and Optimal Transport. *arXiv e-prints*, page arXiv:1009.3737, Sep 2010.

[EFLS16] M. Erbar, M. Fathi, V. Laschos, and A. Schlichting. Gradient flow structure for McKean-Vlasov equations on discrete spaces. *Discrete Contin. Dyn. Syst.*, 36(12):6799–6833, 2016.

[Fat16] M. Fathi. A gradient flow approach to large deviations for diffusion processes. *J. Math. Pures Appl. (9)*, 106(5):957–993, 2016.

[FS16] M. Fathi and M. Simon. The gradient flow approach to hydrodynamic limits for the simple exclusion process. In *From particle systems to partial differential equations. III*, volume 162 of *Springer Proc. Math. Stat.*, pages 167–184. Springer, [Cham], 2016.

[FV18] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems*. Cambridge University Press, Cambridge, 2018. A concrete mathematical introduction.

[GP18] S. N. Gomes and G. A. Pavliotis. Mean Field Limits for Interacting Diffusions in a Two-Scale Potential. *J. Nonlinear Sci.*, 28(3):905–941, jun 2018.

[GPY13] J. Garnier, G. Papanicolaou, and T.-W. Yang. Large deviations for a mean field model of systemic risk. *SIAM J. Financial Math.*, 4(1):151–184, 2013.

[IS96] A. Ioffe and E. Schwartzman. Metric critical point theory. I. Morse regularity and homotopic stability of a minimum. *J. Math. Pures Appl. (9)*, 75(2):125–153, 1996.

[Kat94] G. Katriel. Mountain pass theorems and global homeomorphism theorems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 11(2):189–209, 1994.

[KJZ18] M. Kaiser, R. L. Jack, and J. Zimmer. A Variational Structure for Interacting Particle Systems and their Hydrodynamic Scaling Limits. *arXiv e-prints*, page arXiv:1805.01411, May 2018.

[Loe06] G. Loeper. Uniqueness of the solution to the Vlasov-Poisson system with bounded density. *J. Math. Pures Appl. (9)*, 86(1):68–79, 2006.

[LP66] J. L. Lebowitz and O. Penrose. Rigorous treatment of the van der Waals-Maxwell theory of the liquid-vapor transition. *J. Mathematical Phys.*, 7:98–113, 1966.

[McC97] R. J. McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.

[McC01] R. J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[Rey18] J. Reygner. Equilibrium large deviations for mean-field systems with translation invariance. *Ann. Appl. Probab.*, 28(5):2922–2965, 2018.

[San15] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

[Shi87] M. Shiino. Dynamical behavior of stochastic systems of infinitely many coupled nonlinear oscillators exhibiting phase transitions of mean-field type: H theorem on asymptotic approach to equilibrium and critical slowing down of order-parameter fluctuations. *Phys. Rev. A*, 36(5):2393–2412, sep 1987.

[Sin82] Y. G. Sinaĭ. *Theory of phase transitions: rigorous results*, volume 108 of *International Series in Natural Philosophy*. Pergamon Press, Oxford-Elmsford, N.Y., 1982. Translated from the Russian by J. Fritz, A. Krámli, P. Major and D. Szász.

[Szn91] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[Vil08] C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

Department of Mathematics, Imperial College London, London SW7 2AZ
*Email address*: rishabh.gvalani14@imperial.ac.uk

Institut für Angewandte Mathematik, Universität Bonn
*Email address*: schlichting@iam.uni-bonn.de