

# QUANTITATIVE STABILITY FOR THE BRASCAMP-LIEB INEQUALITY AND MOMENT MEASURES

JOÃO MIGUEL MACHADO AND JOÃO P. G. RAMOS

**ABSTRACT.** By employing the recently obtained sharp stability versions of the Prékopa–Leindler inequality, we are able to obtain a sharp quantitative stability version for the Brascamp–Lieb inequality, as well as several different results on the stability of moment measures.

As main features of our results, we highlight the *independence* of the stability constant for the Brascamp–Lieb inequality on the convex function considered, a completely novel feature. In the realm of moment measures, we highlight in the same vein that the stability results obtained are *uniform*, which we expect to be particularly valuable not only from a purely mathematical point of view, but also for applications.

**Keywords.** Brascamp–Lieb inequality, moment measures, optimal transport, stability, log-concave measures

**2020 Mathematics Subject Classification.** 39B82, 26D10, 49Q22

## CONTENTS

1. Introduction	1
1.1. Stability of the Brascamp–Lieb variance inequality	2
1.2. Stability for moment measures	3
Acknowledgments	6
2. Some tools on optimal transport and functional inequalities	6
2.1. A primer on optimal transport and Wasserstein distances	6
2.2. General facts on moment measures	9
2.3. State of the art on the Prékopa–Leindler inequality	13
3. Sharp Quantitative Stability for Brascamp–Lieb inequality	14
4. Quantitative Stability of Moment Measures	19
4.1. Stability in a compact domain	21
4.2. Convergence rates w.r.t. regularization parameter	22
4.3. Stability over $\mathbb{R}^d$	23
References	28

## 1. INTRODUCTION

In recent years, the fields of calculus of variations and functional inequalities have become deeply intertwined. If from one hand, many of the inequalities used in analysis and probability have deep variational interpretations, functional and geometric inequalities are fundamental in the study of many variational problems. To name a few examples we can cite the use of Sobolev type inequalities in the stability of matter [LS10,Lew22], the use of isoperimetric type inequalities in the study of shape optimization problems [LLR25], and more recently the use of Brascamp–Lieb [BL76] and Prékopa–Leindler inequalities [Pré71, Pre73, Lei73] has become fundamental on the quantitative stability theory of optimal transport [DM23, LM24, Let25].

In the present work, we study the quantitative stability of the *Brascamp–Lieb inequality* and of *moment measures* – as in, for instance, [BL76, CEK15, San16] – and their deep connections. Indeed, that connection has not only been made rigorous several times over (see for instance [Kla14]

where a regularity theory for moment measures is combined with a suitable Brascamp-Lieb inequality to obtain bounds on a Poincaré constant [KL25]), but they also have a common denominator in *optimal transport theory*; see [San15, Vil09, ABS21].

More specifically, the Brascamp-Lieb inequality is not only ubiquitous in the modern stability theory of optimal transport, but it can also be derived with linearization arguments of transportation type inequalities, as done for example in [CE17]. On the other hand, in [San16] it has been showcased how to obtain moment measures with the minimization of functionals coming from optimal transport.

By relying on the recent advances in the quantitative stability of the Prékopa-Leindler inequality in a series of recent works [BFR23, FR24, BFR25, FvHT25], we are able to obtain both a *sharp* stability version of the Brascamp-Lieb variance inequality and novel explicit stability estimates for moment measures. In spite of the fact that both objects can be studied from many different perspectives, a common theme between them is that both can be approached from the viewpoint of the Prékopa-Leindler inequality.

Indeed, we mention, for instance, the classical work [BL00] where a proof of Brascamp-Lieb's variance inequality with a linearization argument using Prékopa-Leindler is provided, as well as [CEK15] (or Section 2.2 below), where a variational approach for moment measures is demonstrated, consisting on the maximization of a functional which is proven to be concave with the same functional inequality.

**1.1. Stability of the Brascamp-Lieb variance inequality.** In order to describe our main contributions, let us first recall the Brascamp-Lieb variance inequality [BL76, BL00]: given a convex function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ , let  $\varrho \propto e^{-\varphi}$  be its associated Gibbs measure. Then for every sufficiently smooth function  $f$ , the following quantity

$$(1.1) \quad \delta_{\text{BL}}(f) \stackrel{\text{def.}}{=} \int_{\mathbb{R}^d} \langle (D^2\varphi)^{-1} \nabla f, \nabla f \rangle d\varrho - \text{Var}_{\varrho}(f) \geq 0$$

is non-negative. It is well known that the optimal functions for such inequality – which make  $\delta_{\text{BL}}(f)$  vanish – are the affine functions on the geometry induced by the log-concave measure  $\varrho$ , i.e.  $f(x) = a \cdot \nabla \varphi(x) + b$ ; see for instance [CE17] or the derivation at the start of Section 3.

Given the characterization of the optimizers of a functional inequality, the question of stability consists of: under which topology can we estimate the distance of  $f$  to the manifold of optimizers?

Results of this type have increasingly gained importance over the past two decades. Indeed, the question of stability has been extensively studied in the case of the isoperimetric inequality [FMP08, FMP10], as well as in other geometric inequalities such as the Faber-Krahn inequality [BDPV15], and, more recently, in [FMP10, FJ17, FvHT23, FvHT24, vHST21, vHST23] the Brunn-Minkowski inequality, in [DT16, DEF<sup>+</sup>25, BK25] for Sobolev and log-Sobolev inequalities and relations with non-linear evolution equations, in [BFR23, FvHT25, FR24] for the Prékopa-Leindler and Borell-Brascamp-Lieb inequalities.

In that vein, our first main contribution is a *sharp  $L^1$ -based stability estimate* for the Brascamp-Lieb variance inequality. In order to state it, we define the finite dimensional manifold of optimal functions

$$(1.2) \quad \mathcal{O}_{\text{BL}} \stackrel{\text{def.}}{=} \{f : \mathbb{R}^d \rightarrow \mathbb{R} : f(x) = a \cdot \nabla \varphi(x) + b, \text{ for some } a \in \mathbb{R}^d, b \in \mathbb{R}\}.$$

Clearly, if  $f \in \mathcal{O}_{\text{BL}}$ , then the parameter  $b$  must be given by  $\mathbb{E}_{\varrho_\varphi} f$  since

$$\mathbb{E}_{\varrho_\varphi} f = \int_{\mathbb{R}^d} f d\varrho_\varphi = a \cdot \int_{\mathbb{R}^d} \nabla \varphi(x) e^{-\varphi} dx + b = a \cdot \int_{\mathbb{R}^d} \nabla(e^{-\varphi}) dx + b = b.$$

Our first main result must then be an estimate of the distance of a function  $f$  to the manifold of optimizers

$$(1.3) \quad \text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{\text{BL}}) \stackrel{\text{def.}}{=} \inf_{g \in \mathcal{O}_{\text{BL}}} \|f - g\|_{L^1(\varrho_\varphi)} = \inf_{a \in \mathbb{R}^d} \|f - (a \cdot \nabla \varphi + \mathbb{E}_{\varrho_\varphi} f)\|_{L^1(\varrho_\varphi)},$$

in terms of the deficit of the Brascamp-Lieb inequality:

**Theorem 1.1.** *Let  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  be an essentially continuous and convex function, and  $\int_{\mathbb{R}^d} e^{-\varphi} < +\infty$  so that the Gibbs measure  $\varrho_\varphi$  is well defined as in (1.4).*

*Then, there exists a universal constant  $C_d$  depending only on the ambient dimension such that*

$$\text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{BL}) \leq C_d \delta_{BL}(f)^{1/2}$$

*for all locally Lipschitz functions  $f \in L^2(\varrho_\varphi)$ .*

As the classical Brascamp-Lieb is independent of the dimension, the first efforts in the literature into improving it were focused on next order terms that would exploit the ambient dimension [Har08, BGG18]. These results are equivalent with bounding the deficit from below with a dimensional term which vanishes in the manifold of optimizers for this inequality. A similar improvement was also obtained in [CE17], by measuring the distance to the manifold of optimal functions with a  $L^2$  term in a geometry that compensates for the curvature induced by the potential  $V$ . This is very natural to expect for the Brascamp-Lieb inequality as it is closely related to the spectral gap of the diffusion operator induced by  $\nabla V$  that is central to the *carré du champ method* [BÉ06, BGL13], see also Section 3. We refer the reader to [LLR25] for a particular version for the Gaussian Poincaré inequality with relations to uncertainty principles.

It is worth pointing out that not only the present work is the first that proposes a quantitative stability result for the Brascamp-Lieb variance inequality in full generality, but that the result is *sharp* in terms of the dependence on  $\delta_{BL}$  on the right-hand side, with constant *independent* of  $\varphi$ . In particular, the independence of our stability results on the convex function provides us with the perfect platform to apply Theorem 1.10 to deduce several novel results about the stability of moment measures, as we shall see below.

Our method of proof will be based on the approach of Bobkov and Ledoux [BL00]. A main new twist in our proof, however, is the usage of the newly developed *sharp stability estimates* for the Prékopa-Leindler inequality stemming from the work of A. Figalli, P. van Hintum and M. Tiba [FvHT25], together with a careful use of the geometry of the functionals involved in order to conclude. This explains our result being stated in terms of the  $L^1$ -topology, which is natural viewing the Brascamp-Lieb inequality as a geometric inequality stemming from Brunn-Minkowski's.

**1.2. Stability for moment measures.** Our next few results use the previous machinery to obtain sharp quantitative stability results for moment measures. Using Prékopa-Leindler, we obtain stability of the Gibbs measures in the moment measures representation, while the quantitative version of Brascamp-Lieb yields stability of the potentials.

Given a convex function  $\varphi$ , we recall the notation of its associated Gibbs probability measure

$$(1.4) \quad \varrho_\varphi \stackrel{\text{def.}}{=} \frac{e^{-\varphi}}{\int_{\mathbb{R}^d} e^{-\varphi}}.$$

On the other hand, given  $\psi$  also convex, which we will view informally as in the dual space of variables of  $\varphi$  by frequently considering  $\psi = \varphi^*$ , we define the *moment measure* associated with  $\psi$  as

$$(1.5) \quad \mu_\psi \stackrel{\text{def.}}{=} (\nabla \psi)_\sharp \varrho_\varphi.$$

In [CEK15, San16], it was proven that for any measure  $\mu \in \mathcal{P}_1(\mathbb{R})$ , that is a Radon probability measure over  $\mathbb{R}^d$  with finite first moments, admits a moment measure representation, there exists a convex and essentially continuous (see the definition in Section 2.2) potential  $\psi_\mu$  such that (1.5) holds. In addition, it is also unique up to translations. Using the Prékopa-Leindler inequality, it was shown in [CEK15] that the moment measure of  $\psi$  is  $\mu$  if and only if its Legendre transform  $\varphi \stackrel{\text{def.}}{=} \psi^*$  is a maximizer of

$$(1.6) \quad \mathcal{J}_\mu(\varphi) \stackrel{\text{def.}}{=} \log \int_{\mathbb{R}^d} e^{-\varphi^*} dx - \int_{\mathbb{R}^d} \varphi d\mu.$$

The same result was recovered in [San16] using optimal transport, but this time for a dual perspective: a measure  $\varrho$  is the Gibbs measure in (1.5) if and only if it minimizes

$$(1.7) \quad \mathcal{E}_\mu(\varrho) \stackrel{\text{def.}}{=} \int_{\mathbb{R}^d} \log \varrho(x) d\varrho(x) + \mathcal{T}(\varrho, \mu),$$

where  $\mathcal{T}$  is the maximal correlation formulation of the quadratic optimal transport problem. The reader is referred to Sections 2.1 and 2.2 for more details on optimal transport theory and for general results on moment measures.

The quantitative stability question that we wish to answer then becomes: given two measures  $\mu, \nu$  can we quantify the distance between either  $\psi_\mu, \psi_\nu$  or between  $\varrho_\mu, \varrho_\nu$ . Since the variational characterization described above is necessary and sufficient, one can only expect to answer this question as an estimate on the distance to the optimal sets defined as

$$(1.8) \quad \mathcal{M}_\mu \stackrel{\text{def.}}{=} \operatorname{argmin} \mathcal{E}_\mu, \quad \mathcal{N}_\mu \stackrel{\text{def.}}{=} \operatorname{argmax} \mathcal{J}_\mu,$$

which can be characterized as follows: fixed some reference  $\varrho_0 \in \mathcal{M}_\mu$  and  $\varphi_0 \in \mathcal{N}_\mu$ , the manifolds of optimizers can be expressed as

$$(1.9) \quad \begin{aligned} \mathcal{M}_\mu &= \{\varrho_0(\cdot - x_0) : x_0 \in \mathbb{R}^d\}, \\ \mathcal{N}_\mu &= \{\varphi_0 + (a \cdot x + b) : a \in \mathbb{R}^d, b \in \mathbb{R}\}. \end{aligned}$$

We start with the following Theorem, which is a backbone to the stability results that will follow. We hope it can be used to other purposes, such as deriving explicit convergence rates for numerical algorithms for computing moment measures.

**Theorem 1.2.** *Given a Radon measure  $\mu \in \mathscr{P}_1(\mathbb{R}^d)$  whose barycenter lies that the origin and  $\dim \operatorname{supp} \mu = d$ , let  $\bar{\varphi}$  be a maximizer of  $\mathcal{J}_\mu$ . Then for any convex function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  it holds that*

$$(1.10) \quad \operatorname{dist}_{L^1(\mathbb{R}^d)}(\varrho_{\varphi^*}, \mathcal{M}_\mu) \leq C_d (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi))^{1/2}.$$

In addition, there exists  $\lambda \in (0, 1/2)$  such that, setting  $v \stackrel{\text{def.}}{=} \bar{\varphi} - \varphi$  and  $\varphi_\lambda \stackrel{\text{def.}}{=} \bar{\varphi} + \lambda v$ , it holds that

$$(1.11) \quad \operatorname{dist}_{L^1(\mu_\lambda)}(\varphi, \mathcal{N}_\mu) \leq C_d (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi))^{1/2},$$

where  $\mu_\lambda = \mu_{\varphi_\lambda^*}$  is the moment measure associated with the interpolation  $(\varphi_\lambda)^*$ .

We note that this result may be regarded as a Polyak-Lojasiewicz type inequality for the functional  $\mathcal{J}_\mu$ . More precisely, estimate (1.11) gives a way of using this functional to measure the distance of any  $\varphi$  to the class of maximizers of  $\mathcal{J}_\mu$ . This is done by establishing a connection between the equality case of the Brascamp-Lieb inequality and the invariance of the functional  $\mathcal{J}_\mu$  with respect to the addition of affine functions. Estimate (4.1) relates the invariance of the moment measure representation with respect to translations to the equality cases in the Prékopa-Leindler inequality, estimating the distance of  $\varrho_{\varphi^*}$  to the set of minimizers of  $\mathcal{E}_\mu$ .

In spite of the fact that Theorem 1.2 provides us, by itself, with a novel quantitative way to show that almost maximizers of the functionals  $\mathcal{J}_\mu$  and  $\mathcal{E}_\mu$  must indeed converge to the original potentials, we highlight that its true scope lies in the *deeper applications* to the theory of moment measures it unveils, as pointed below.

*Stability in a compact domain.* Using Theorem 1.2, we are able to obtain several quantitative stability results for moment measures in different contexts. When the moment measure and the associated Gibbs measure in (1.5) are restricted to a compact domain we obtain a strong stability result controlling the  $L^1$  distances between  $\varrho_\mu, \varrho_\nu$  and  $\varphi_\mu, \varphi_\nu$  with the 1-Wasserstein distance of  $\mu, \nu$ ; this is the content of Theorem 4.2.

In spite of the fact that compactness allows for a fairly simpler mathematical structure, diverging from the objects described by the existence theorem by [CEK15, San16], we remark that such a result means a remarkable control *especially for numerical applications*, as we are able

to control a strong norm with a distance which metrizes a fairly weak convergence of Radon measures, see Section 2.1 for more details on these topologies.

*Rates of convergence for quadratic regularization.* The classical definition of moment measures from (1.5) has strong motivations from convex and toric geometry [BB13], but variations thereof have recently found several applications to sampling and generative models [VBC25]. For such applications, the essential property of moment measures is their capacity of representing a given measure via the pushforward of a log-concave distribution through the gradient of the same convex convex potential.

This can also be achieved with the *quadratically regularized moment measures* recently proposed in [DF25]. That is, given  $\mu \in \mathcal{P}_1(\mathbb{R}^d)$  there is a convex potential  $\psi_\alpha$  such that

$$(1.12) \quad \mu = (\nabla \psi_\alpha)_\sharp \varrho_\alpha, \text{ where } \varrho_\alpha \propto e^{-(\varphi_\alpha + \frac{\alpha}{2}|\cdot|^2)};$$

see also Section 2.2.

By introducing the quadratic regularization, the quantitative stability investigated in [DF25] turns out to become a bit simplified for a few reasons: first, we mention that the uniqueness modulo translations is replaced by simply uniqueness. This way, the quantitative stability can be understood without the distance to a certain manifold of optimizers. Secondly, the regularization makes the functional corresponding to (1.7) strongly geodesically convex in Wasserstein space. On the other hand, the stability is formulated with the 2-Wasserstein distance.

For these reasons, both of theoretical and of applied nature, obtaining explicit rates of convergence with respect to the regularization parameter is of great interest. This is achieved once again by using Theorem 1.2.

**Theorem 1.3.** *Fix a Radon  $\mu \in \mathcal{P}_1(\mathbb{R}^d)$  whose barycenter lies at the origin and such that  $\dim \text{supp } \mu = d$ . Then there exists a constant  $C_{d,\mu}$  depending on the ambient dimension and on  $\mu$  such that*

$$(1.13) \quad \text{dist}_{L^1(\mathbb{R}^d)}(\varrho_\alpha, \mathcal{M}_\mu) \leq C_{d,\mu} \alpha^{1/2},$$

for  $\varrho_\alpha$  given in (1.12).

In addition, set  $\varphi_\alpha \stackrel{\text{def.}}{=} \psi_\alpha$ , where  $\psi_\alpha$  is given in (1.12). Then, for each  $\alpha$ , there exists  $\lambda \in (0, 1/2)$  such that, setting  $v \stackrel{\text{def.}}{=} \varphi - \varphi_\alpha$  and  $\varphi_t \stackrel{\text{def.}}{=} \varphi + \lambda v$ , it holds that

$$(1.14) \quad \text{dist}_{L^1(\mu_\lambda)}(\varphi_\alpha, \mathcal{N}_\mu) \leq C_{d,\mu} \alpha^{1/2},$$

where  $\mu_\lambda = \mu_{\varphi_\lambda^*}$  is the moment measure associated with the interpolation  $(\varphi_\lambda)^*$ .

*Stability in  $\mathbb{R}^d$ .* To tackle the stability question in  $\mathbb{R}^d$  we can in principle combine the quantitative stability of regularized moment measures arguments based on strong geodesic convexity from [DF25] with our explicit rates of convergence 1.3. But, in order to achieve this, we have to make quantitative the assumption that  $\mu$  is not supported on a hyperplane, otherwise such a result would allow the construction of a  $\psi$  yielding (1.5) for such a singular  $\mu$ , contradicting the existence theory of [CEK15, San16]. See Proposition 4.8 for more details on this construction.

Our approach is then to either impose a *uniform lower bound* on the geometric quantity

$$\Theta(\mu) \stackrel{\text{def.}}{=} \inf_{\theta \in \mathbb{S}^{d-1}} \int_{\mathbb{R}^d} |\theta \cdot y| d\mu(y) \geq \vartheta,$$

or to control the Hessian of the log-density of  $\mu$ . This is done by defining the two classes

$$\begin{aligned} \mathcal{K}_\vartheta &\stackrel{\text{def.}}{=} \left\{ \mu \in \mathcal{P}(\Omega) : \begin{array}{l} \mu \text{ is centered,} \\ \Theta(\mu) \geq \vartheta \end{array} \right\} \text{ for a compact } \Omega, \\ \mathcal{K}_\Lambda &\stackrel{\text{def.}}{=} \left\{ \mu \in \mathcal{P}_{\text{ac}}(\mathbb{R}^d) : \begin{array}{l} \mu \propto e^{-V} \text{ is centered,} \\ D^2 V \leq \Lambda \text{id} \end{array} \right\}. \end{aligned}$$

The results we obtain can be summarized as follows, where the reader is referred to Section 4.3 for more information.

**Theorem 1.4.** *Take  $\mathcal{K}$  to be either  $\mathcal{K}_\vartheta$  or  $\mathcal{K}_\Lambda$ . Then for each  $p > 2$  there exists a constant  $C$  such that, for all  $\mu, \nu \in \mathcal{K}$  it holds that*

$$(1.15) \quad \text{dist}_{W_2}(\mathcal{M}_\mu, \mathcal{M}_\nu) \leq CW_2(\mu, \nu)^{\frac{p-2}{6p-4}},$$

where  $C$  grows linearly in  $p$ .

In both cases, our approach is to control the moments of order  $p$  of the Gibbs measures  $\varrho_\mu$ , uniformly for all elements  $\mu$  of the classes  $\mathcal{K}_\vartheta$  and  $\mathcal{K}_\Lambda$ . With this finer information we can come back to the strong geodesic convexity arguments for regularized moment measures from [DF25] and obtain an optimal dependence on  $\alpha$  in the quantitative stability results in this case. For the first class  $\mathcal{K}_\vartheta$ , this can be done with a Lemma 4.6 due to Klartag. For the class  $\mathcal{K}_\Lambda$ , similar controls on the  $p$ -moments can be obtained with a regularity result of moment measures, namely

$$(1.16) \quad \mu \propto e^{-V} = (\nabla \psi)_\sharp e^{-\psi}, \text{ such that } D^2V \leq \Lambda \text{id}, \text{ then } \psi \text{ is } \Lambda^{-1/3}\text{-strongly convex.}$$

The proof of (1.16) is done in Theorem 4.11, and is based on a bootstrap argument using Caffarelli's contraction theorem [Caf92, Caf00]. It states that if  $\mu$  is as in (1.16) and  $\nu \propto e^{-W}$  with  $D^2W \geq \lambda \text{id}$ , then the optimal transportation map  $T$  from  $\mu$  to  $\nu$  is globally Lipschitz with  $\text{Lip } T \leq \sqrt{\Lambda/\lambda}$ . Since we cannot know in principle that  $\psi$  is strongly convex, what we do instead is consider the quadratic regularized moment measure presentation of  $\mu$ . In this case, the Gibbs measure has indeed a strongly convex potential, with a modulus of convexity  $\alpha$ . This modulus of convexity can be iteratively increased with successive applications of the contraction theorem, which yields  $\Lambda^{-1/3}$  at the end. Besides the application to the stability of moment measures, we believe that this result might have important applications in sampling, see the discussion after the proof of Theorem 4.11.

#### ACKNOWLEDGMENTS

During the final preparation steps of the present work we have discovered that other colleagues were working on similar questions on the stability of moment measures [BJR], but with a particular focus on semi-discrete approximations and applications in Kähler—Ricci solitons. We wish to thank them, specially Guillaume Bonnet, for clarifying discussions. The first author wishes to thank Quentin Mérigot and Filippo Santambrogio for stimulating discussions about moment measures. He also warmly thanks the support of the Lagrange Mathematical and Computational Research Center.

## 2. SOME TOOLS ON OPTIMAL TRANSPORT AND FUNCTIONAL INEQUALITIES

**2.1. A primer on optimal transport and Wasserstein distances.** The Wasserstein distances are defined via the value function of the optimal transport problem as

$$(2.1) \quad W_p^p(\mu, \nu) \stackrel{\text{def.}}{=} \min_{\gamma \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma(x, y),$$

where  $\Pi(\mu, \nu)$  the set of probability measures over  $\mathbb{R}^d \times \mathbb{R}^d$  whose marginals are respectively  $\mu$  and  $\nu$ , the so called set of transportation plans. This quantity is finite if and only if  $\mu$  and  $\nu$  have finite  $p$ -moments,  $M_p(\mu), M_p(\nu) < +\infty$ , where

$$(2.2) \quad M_p(\mu) \stackrel{\text{def.}}{=} \int_{\mathbb{R}^d} |x|^p d\mu.$$

The reader is referred to [San15, Vil09] for a complete introduction to optimal transport and Wasserstein distances, in what follows we introduce its properties that will be useful in the sequel.

We are particularly interested in the case  $p = 2$ , where the optimal transport problem 2.1 has a close relation with convex analysis. This can be easily seen by introducing the *maximal correlation formulation*

$$(2.3) \quad \mathcal{T}(\mu, \nu) = \sup_{\gamma \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \langle x, y \rangle d\gamma = \inf_{\varphi \text{ convex}} \int_{\mathbb{R}^d} \varphi d\mu + \int_{\mathbb{R}^d} \varphi^* d\nu,$$

in such a way that

$$\frac{1}{2} W_2^2(\mu, \nu) = \frac{1}{2} M_2(\mu) + \frac{1}{2} M_2(\nu) - \mathcal{T}(\mu, \nu).$$

The infimum on the RHS is attained by a convex function  $\varphi$ , and Brenier's Theorem states that the optimal transportation plan is unique and given by  $\gamma = (\text{id}, \nabla \varphi)_\sharp \mu$ , whenever  $\mu$  is absolutely continuous w.r.t. the Lebesgue measure.

*Topological properties of Wasserstein distances.* A very important property of Wasserstein distances is that they metrize the narrow topology of probability measures. A sequence of Radon probability measures  $(\mu_n)_{n \in \mathbb{N}}$  converges narrowly to  $\mu$  if for all  $f \in \mathscr{C}_b(\mathbb{R}^d)$ , continuous and bounded function, it holds that

$$\int_{\mathbb{R}^d} f d\mu_n \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} f d\mu,$$

and we write  $\mu_n \xrightarrow{n \rightarrow \infty} \mu$ . A very important property is that

$$(2.4) \quad W_p(\mu_n, \mu) \xrightarrow{n \rightarrow \infty} 0 \text{ if and only if } \begin{cases} \mu_n \xrightarrow{n \rightarrow \infty} \mu, \\ M_p(\mu_n) \xrightarrow{n \rightarrow \infty} M_p(\mu). \end{cases}$$

In addition, the space of Radon probability measures with finite  $p$ -moments  $\mathscr{P}_p(\mathbb{R}^d)$  becomes itself complete and separable when endowed with the topology of  $W_p$ .

This is conceptually very important for the contributions of this work since many of the quantitative stability results regarding moment measure representations from Section 4 are stated with respect to the Wasserstein distance  $W_2$ . This allows to compare very singular objects with a weak topology. On the other hand, if we know that two probability measures are also in  $L^1(\mathbb{R}^d)$  and have finite  $p$ -moments, then we can interpolate the Wasserstein distance with the  $L^1$  distance via the following inequality.

**Lemma 2.1.** *Let  $\mu, \nu$  be two probability densities in  $L^1(\mathbb{R}^d)$  with finite  $p$ -moments, for some  $p > 1$ . Then, for all  $1 \leq q < p$  there exists a constant  $C_{p,q}$  such that*

$$(2.5) \quad W_q(\mu, \nu) \leq C_{p,q} (M_p(\mu) + M_p(\nu))^{1/p} \|\mu - \nu\|_{L^1}^{1/q-1/p}.$$

*Proof.* Suppose that  $\mu, \nu$  have densities  $f, g$ , set  $\varepsilon \stackrel{\text{def.}}{=} \|f - g\|_{L^1(\mathbb{R}^d)}$ , and define the common density and the residuals  $f_r, g_r$  as

$$h(x) \stackrel{\text{def.}}{=} \min\{f(x), g(x)\}, \quad f_r \stackrel{\text{def.}}{=} f - h, \quad g_r \stackrel{\text{def.}}{=} g - h.$$

As a result, we observe that

$$\int_{\mathbb{R}^d} f_r(x) dx + \int_{\mathbb{R}^d} g_r(x) dx = \varepsilon, \quad \int_{\mathbb{R}^d} f_r(x) dx - \int_{\mathbb{R}^d} g_r(x) dx = 0.$$

so that  $\int_{\mathbb{R}^d} f_r(x) dx = \int_{\mathbb{R}^d} g_r(x) dx = \varepsilon/2$ .

Now, we can construct a transportation plan  $\gamma \in \Pi(\mu, \nu)$  as follows: we first transport the common part of both densities via the identity map, and to transport the residuals we use the product measure. More precisely, we set  $\alpha \stackrel{\text{def.}}{=} \varepsilon/2$  and

$$\gamma \stackrel{\text{def.}}{=} (\text{id}, \text{id})_\sharp h + \frac{1}{\alpha} f_r(x) g_r(y) dx dy.$$

We then have

$$W_q^q(\mu, \nu) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^q d\gamma(x, y) = \frac{1}{\alpha} \underbrace{\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^q f_r(x) g_r(y) dx dy}_{\stackrel{\text{def.}}{=} I}.$$

To estimate the RHS, we split the integral into two parts as

$$\begin{aligned} I &= I_1 + I_2, \\ &= \int_{B_R \times B_R} |x - y|^q f_r(x) g_r(y) dx dy + \int_{(B_R \times B_R)^c} |x - y|^q f_r(x) g_r(y) dx dy. \end{aligned}$$

The first term can be easily estimated as

$$I_1 \leq (2R)^q \alpha^2,$$

while the second can be estimated as with the  $p$ -moments, for any  $p > q$  by noticing that for  $|x| > R$  we have  $|x|^q \leq R^{q-p}|x|^p$ , so that

$$\begin{aligned} I_2 &\leq 2^{q-1} \alpha \int_{|x| \geq R} |x|^q f_r(x) dx + 2^{q-1} \alpha \int_{|y| \geq R} |y|^q g_r(y) dy \\ &\leq 2^{q-1} R^{q-p} \alpha (M_p(\mu) + M_p(\nu)). \end{aligned}$$

Combining both estimates, we obtain

$$W_q^q(\mu, \nu) \leq (2R)^q \alpha + 2^{q-1} R^{q-p} (M_p(\mu) + M_p(\nu)).$$

Optimizing in  $R > 0$  gives the desired result.  $\square$

*Geodesic convexity in Wasserstein spaces.* Given two probability measures  $\mu_0, \mu_1 \in \mathcal{P}_p(\mathbb{R}^d)$ , any optimal transportation plan  $\gamma$  yields a natural interpolation between them given by

$$\mu_t \stackrel{\text{def.}}{=} \pi_t \sharp \gamma, \quad \text{where } \pi_t(x, y) = (1-t)x + ty, \quad t \in [0, 1].$$

It turns out that this interpolation is a constant speed geodesic in the metric space  $(\mathcal{P}_p(\mathbb{R}^d), W_p)$ , see [ABS21, San15, Vil09] for details.

This allows to define the notion of geodesic convexity for functionals defined over Wasserstein spaces. A functional  $\mathcal{F} : \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  is said to be geodesically convex if for every  $\mu_0, \mu_1 \in \mathcal{P}_p(\mathbb{R}^d)$  and any geodesic interpolation  $\mu_t$  between them the function  $t \mapsto \mathcal{F}(\mu_t)$  is convex in the classical sense. It is strictly geodesically convex if the same function is strictly convex.

The 2-Wasserstein distance is not itself geodesically convex, this is a feature of the geometry of the Wasserstein space, see [AGS08, Chapter 9]. In fact, it satisfies the opposite inequality. In [San16, Proposition 3.3], Santambrogio exploited this fact to show the geodesic convexity of the maximal correlation functional  $\varrho \mapsto \mathcal{T}(\varrho, \mu)$ . This fact is at the heart of the study of moment measures via optimal transport, as well as the geodesic convexity of the entropy functional defined as

$$(2.6) \quad H(\varrho) \stackrel{\text{def.}}{=} \begin{cases} \int_{\mathbb{R}^d} \log \varrho d\varrho, & \text{if } \varrho \ll \mathcal{L}^d, \\ +\infty, & \text{otherwise,} \end{cases}$$

which follows from McCann's criterion [McC97]<sup>1</sup>.

Another important example of geodesically convex functionals are the  $p$ -moments,  $\mu \mapsto M_p(\mu)$ , this follows directly from the convexity of  $x \mapsto |x|^p$ , see [San15, Chapter 7]. In fact, it is easy

---

<sup>1</sup>And seems to be equivalent ot the Prékopa-Leindler inequality [Pré71, Lei73].

to see that the second moment  $M_2$  is even *strongly* geodesically convex, that is, for any geodesic interpolation  $\mu_t$  between  $\mu_0$  and  $\mu_1$  it holds that

$$(2.7) \quad M_2(\mu_t) \leq (1-t)M_2(\mu_0) + tM_2(\mu_1) - \frac{1}{2}t(1-t)W_2^2(\mu_0, \mu_1).$$

This property is particularly useful to derive quantitative stability results for minimizers, as done in [DF25] and will be exploited here.

**2.2. General facts on moment measures.** In [CEK15], Cordero-Erasquin and Klartag gave a complete characterization of the measures  $\mu$  which admit a meaningful moment measure representation. Their analysis goes through the functionals defined as in (1.6)

They identified that if the potential  $\psi$  in (1.5) does not satisfy certain regularity properties, its moment measure can loose a lot of information. They shed some light on this issue with the following example: take  $y \in \mathbb{R}^d$ , a convex body  $C \subset \mathbb{R}^d$  and consider

$$\psi(x) \stackrel{\text{def.}}{=} \begin{cases} x \cdot y, & x \in C \\ +\infty, & \text{otherwise.} \end{cases}$$

Then one can readily check that  $\mu_\psi = \delta_y$ , which does not give much information on the measure  $\delta_y$ , and certainly is not amenable to a quantitative stability result since one can drastically change the convex set  $C$  and obtain the same moment measure. For this reason they focus on characterizing measures that admit a moment measure representation with an *essentially continuous* and convex function.

**Definition 2.2.** A convex function  $\psi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  is said to be *essentially continuous* if it is lower semi-continuous and the set of points where it is discontinuous has zero  $\mathcal{H}^{d-1}$  measure.

**Lemma 2.3** ([CEK15]). *Let  $\psi$  be a convex and essentially continuous function such that  $0 < \int_{\mathbb{R}^d} e^{-\psi} < +\infty$ , and such that  $0 < \int_{\mathbb{R}^d} e^{-\psi} < +\infty$ . Then, its moment measure  $\mu_\psi$  is centered and its support spans  $\mathbb{R}^d$ . In particular,*

$$\inf_{\theta \in \mathbb{S}^{d-1}} \int_{\mathbb{R}^d} |x \cdot \theta| d\mu_\psi(x) > 0.$$

Alternatively, Santambrogio proposed an optimal transport proof of the same result by exploring the *geodesic convexity* of the following functional

$$(2.8) \quad \min_{\varrho \in \mathcal{P}_1(\mathbb{R}^d)} \mathcal{E}_\mu(\varrho) \stackrel{\text{def.}}{=} H(\varrho) + \mathcal{T}(\varrho, \mu),$$

where  $H$  corresponds to the negative entropy functional (2.6) and  $\mathcal{T}$  denotes Brenier's formulation of the quadratic optimal transport problem, or maximimal correlation problem (2.3). Santambrogio relates both problems and recovers the characterization of moment measures with minimizers of  $\mathcal{E}_\mu$ , which are log-concave probability densities whose potential is essentially continuous, thus recovering the results of Cordero-Erasquin and Klartag.

More recently, Delalande and Farineli [DF25] studied the quantitative stability of *regularized moment measures*, which they define via the unique minimizer of the following functional, which consists of a second moment regularized version of  $\mathcal{E}_\mu$ :

$$(2.9) \quad \mathcal{E}_{\mu, \alpha}(\varrho) \stackrel{\text{def.}}{=} H(\varrho) + \mathcal{T}(\varrho, \mu) + \frac{\alpha}{2} M_2(\varrho).$$

Minimizers of the regularized version are now unique and of the following form: there exists  $\psi_\alpha$  convex such that

$$\varrho_\alpha = e^{-(\psi_\alpha + \frac{\alpha}{2}|x|^2)} \text{ and } (\nabla \psi_\alpha)_\sharp e^{-(\psi_\alpha + \frac{\alpha}{2}|x|^2)} = \mu.$$

A major interest of introducing regularized moment measures is computational; indeed regularizing the functional  $\mathcal{E}_\mu$  with the second moment term  $M_2(\varrho)$  makes it  $\alpha$ -strongly geodesically convex in the Wasserstein geometry while preserving the most important feature of the representation via moment measures, that is being the push-forward of a Gibbs measure through (a

perturbation of) its potential's gradient. For these reasons, it's not surprising that the arguments in [DF25] are based on strong geodesic convexity and actually do not require the dual formulation of  $\mathcal{E}_{\mu,\alpha}$ .

To fully exploit the sharp stability version of Prékopa-Leindler's inequality, we shall require the connection of both primal and dual formulations. For this reason we introduce the regularized dual functional defined for any  $\alpha \geq 0$  as

$$(2.10) \quad \mathcal{J}_{\mu,\alpha}(\varphi) \stackrel{\text{def.}}{=} \log \int_{\mathbb{R}^d} e^{-(\varphi^* + \frac{\alpha}{2}|x|^2)} dx - \int_{\mathbb{R}^d} \varphi d\mu.$$

In the following theorem, we summarize the contributions from the literature and show a strong-duality type result with a small adaptation of a remark from Santambrogio in [San16], but that will be useful in the sequel. More importantly, we identify a way to map minimizers of  $\mathcal{E}_{\mu,\alpha}$  into maximizers of  $\mathcal{J}_{\mu,\alpha}$  and vice-versa.

**Theorem 2.4** ([CEK15, San16, DF25]). *Let  $\mu$  be a Radon probability measure in  $\mathcal{P}_1(\mathbb{R}^d)$  whose barycenter lies at the origin and such that  $\dim \text{supp } \mu = d$ . The following assertions hold:*

- For every  $\alpha \geq 0$  there is strong duality:

$$(2.11) \quad \sup \mathcal{J}_{\mu,\alpha} = -\inf \mathcal{E}_{\mu,\alpha}.$$

- $\mathcal{E}_{\mu,\alpha}$  admits a minimizer of the form

$$(2.12) \quad \varrho_\alpha = e^{-(\psi_\alpha + \frac{\alpha}{2}|x|^2)} \text{ and } (\nabla \psi_\alpha)_\sharp \varrho_\alpha = \mu,$$

where  $\psi_\alpha$  is an essentially continuous and convex function, so that  $\nabla \psi_\alpha$  is the unique Brenier map from  $\varrho_\alpha$  to  $\mu$ .

If  $\alpha > 0$  this minimizer is unique; and if  $\alpha = 0$  it is unique up to translations. In any case, if  $\varrho_\alpha$  is a minimizer, defining

$$(2.13) \quad \varphi_\alpha \stackrel{\text{def.}}{=} \psi_\alpha^* \text{ for } \psi_\alpha \stackrel{\text{def.}}{=} -\log \varrho_\alpha - \frac{\alpha}{2}|x|^2,$$

it follows that  $\varphi_\alpha$  is a maximizer of  $\mathcal{J}_{\mu,\alpha}$ .

- $\mathcal{J}_{\mu,\alpha}$  admits a maximizer  $\varphi_\alpha$ , whose Legendre transform  $\psi_\alpha \stackrel{\text{def.}}{=} \varphi_\alpha^*$  is an essentially continuous and convex function, and gives the regularized moment measure representation for  $\mu$

$$(2.14) \quad (\nabla \psi_\alpha)_\sharp e^{-(\psi_\alpha + \frac{\alpha}{2}|x|^2)} = \mu$$

If  $\alpha > 0$ , the maximizer is unique; whereas if  $\alpha = 0$  it is unique up to the addition of an affine function.

In any case, if  $\varphi_\alpha$  is a maximizer, defining  $\psi_\alpha = \varphi_\alpha^*$  it follows that

$$(2.15) \quad \varrho_\alpha \stackrel{\text{def.}}{=} e^{-(\psi_\alpha + \frac{\alpha}{2}|x|^2)}$$

is a minimizer of  $\mathcal{E}_{\mu,\alpha}$ .

*Proof.* First let us prove the strong duality type-result. It was already proven in the case  $\alpha = 0$  in [San16], we prove it here for general  $\alpha$  for completeness and for the reader's convenience.

First, we compute the Legendre transform of  $\mathcal{E}_{\mu,\alpha}$ . It is defined as a functional over the space of convex functions as

$$\begin{aligned} \mathcal{E}_{\mu,\alpha}^*(f) &\stackrel{\text{def.}}{=} \sup_{\rho \in \mathcal{P}(\mathbb{R}^d)} \langle f, \rho \rangle - H(\rho) - \frac{\alpha}{2} \int_{\mathbb{R}^d} |x|^2 d\rho - \inf_{\varphi \text{ convex}} \int_{\mathbb{R}^d} \varphi^* d\rho + \int_{\mathbb{R}^d} \varphi d\mu \\ &= \sup_{\varphi \text{ convex}} - \int_{\mathbb{R}^d} \varphi d\mu + \sup_{\rho \in \mathcal{P}_{\text{ac}}(\mathbb{R}^d)} \left\langle f - \varphi^* - \frac{\alpha}{2}|x|^2, \rho \right\rangle - \int_{\mathbb{R}^d} \log \rho d\rho \end{aligned}$$

The supremum on  $\rho$  is attained, and the maximizer is characterized by the following Euler-Lagrange equations; see [San15]:  $\rho > 0$  a.e. over  $\mathbb{R}^d$  and

$$\log \rho + \varphi^* + \frac{\alpha}{2}|x|^2 - f \equiv \text{cte},$$

so that  $\rho \propto e^{-(\varphi^* + \frac{\alpha}{2}|x|^2 - f)}$ . A straightforward computation then gives

$$\mathcal{E}_{\mu,\alpha}^*(f) = \sup_{\varphi \text{ convex}} \log \int_{\mathbb{R}^d} e^{-(\varphi^* + \frac{\alpha}{2}|x|^2 - f)} dx - \int_{\mathbb{R}^d} \varphi d\mu.$$

As a result, taking  $f \equiv 0$  we get that

$$\sup_{\varphi \text{ convex}} \mathcal{J}_{\mu,\alpha}(\varphi) = - \inf_{\varrho \in \mathcal{P}_1(\mathbb{R}^d)} \mathcal{E}_{\mu,\alpha}(\varrho).$$

The existence of minimizers for  $\mathcal{E}_{\mu,\alpha}$  and maximizers for  $\mathcal{J}_{\mu,\alpha}$  is completely equivalent, as we will see in the sequel. However, the proof of existence for  $\mathcal{E}_{\mu,\alpha}$  and general  $\alpha \geq 0$  has been done in [DF25], along with the claimed characterization of minimizers and the relation with moment measures. Therefore, to be economic the reader referred to this work for a proof of the second assertion, we only comment that the uniqueness in the case  $\alpha > 0$  stems from the strong geodesic convexity of the energy, while the functional  $\mathcal{E}_{\mu}$  is invariant with respect to translations, [San16, Prop. 3.1].

Now given a minimizer  $\varrho_\alpha$ , for any  $\alpha \geq 0$ , consider  $\varrho_\alpha$  as defined above and let us compute  $\mathcal{J}_{\mu,\alpha}(\varrho_\alpha)$  as follows

$$\begin{aligned} \mathcal{J}_{\mu,\alpha}(\varphi_\alpha) &= \log \int_{\mathbb{R}^d} e^{-(\varphi_\alpha^* + \frac{\alpha}{2}|x|^2)} dx - \int_{\mathbb{R}^d} \varphi_\alpha d\mu \\ &= \underbrace{\log \int_{\mathbb{R}^d} e^{\log \varrho_\alpha} dx}_{=0} - \int_{\mathbb{R}^d} \varphi_\alpha d\mu. \end{aligned}$$

Now using the regularized moment measure representation for  $\mu$  in terms of  $\psi_\alpha = \varphi_\alpha^*$  and Fenchel's identity  $\varphi_\alpha(\nabla \varphi_\alpha^*(x)) = \langle x, \nabla \varphi_\alpha^*(x) \rangle - \varphi_\alpha^*(x)$ , it holds that

$$\begin{aligned} \mathcal{J}_{\mu,\alpha}(\varphi_\alpha) &= - \int_{\mathbb{R}^d} \varphi_\alpha(\nabla \varphi_\alpha^*(x)) e^{-(\varphi_\alpha^*(x) + \frac{\alpha}{2}|x|^2)} dx \\ &= - \underbrace{\int_{\mathbb{R}^d} \langle x, \nabla \varphi_\alpha^*(x) \rangle d\varrho_\alpha}_{=\mathcal{T}(\varrho_\alpha, \mu)} + \underbrace{\int_{\mathbb{R}^d} \left( \varphi_\alpha^*(x) + \frac{\alpha}{2}|x|^2 \right) d\varrho_\alpha}_{=-H(\varrho_\alpha)} - \frac{\alpha}{2} \int_{\mathbb{R}^d} |x|^2 d\varrho_\alpha \\ &= -\mathcal{E}_{\mu,\alpha}(\varrho_\alpha). \end{aligned}$$

As a result, we finally obtain that

$$\sup \mathcal{J}_{\mu,\alpha} \geq \mathcal{J}_{\mu,\alpha}(\psi_\alpha) = -\mathcal{E}_{\mu,\alpha}(\varrho_\alpha) = -\inf \mathcal{E}_{\mu,\alpha} = \sup \mathcal{J}_{\mu,\alpha},$$

so that  $\varphi_\alpha$  must be a maximizer of  $\mathcal{J}_{\mu,\alpha}$ .

Doing the same reasoning, but on the inverse, starting an arbitrary maximizer of  $\mathcal{J}_{\mu,\alpha}$ , we can construct a log-concave probability measure which minimizes  $\mathcal{E}_{\mu,\alpha}$ , but since those are unique thanks to the strong geodesic convexity, or at least unique up to a translation in the case  $\alpha = 0$ , we obtain that the maximizers of  $\mathcal{J}_{\mu,\alpha}$  must also be unique, or unique up to adding an affine function in the case  $\alpha = 0$ .  $\square$

It can also be shown that the Gibbs measure associated with the moment measure (resp. regularized moment measure) representation of a given  $\mu$  is a minimizer for  $\mathcal{E}_{\mu}$  (resp.  $\mathcal{E}_{\mu,\alpha}$ ), see [San16, Prop. 5.1] for classical moment measures, [DF25, Prop. 4.2] and [CEK15, Thm. 8] for the corresponding statement for the dual functional  $\mathcal{J}_{\mu}$ .

The moment measure representation can be seen as a stationarity condition for  $\mathcal{J}_{\mu}$  and its maximization can be studied through its concavity, which is equivalent to the Brascamp-Lieb inequality. Indeed, its first and second variations are computed as follows:

**Lemma 2.5.** *The first and second variations of  $\mathcal{J}_\mu$  are computed as follows: given two convex functions  $\varphi_0, \varphi_1$ , setting  $v \stackrel{\text{def.}}{=} \varphi_1 - \varphi_0$ ,  $\varphi_t \stackrel{\text{def.}}{=} \varphi_0 + tv$ , and the Gibbs probability measure  $\varrho_{\varphi_t^*, \alpha} \propto e^{-(\varphi_t^* + \frac{\alpha}{2}|x|^2)}$  and its regularized moment measure  $\mu_{\varphi_t^*, \alpha} = (\nabla \varphi_t^*)_\sharp \varrho_{\varphi_t^*, \alpha}$  it holds that*

$$\frac{d}{dt} \mathcal{J}_{\mu, \alpha}(\varphi_t) = \int_{\mathbb{R}^d} v d(\mu_{\varphi_t^*, \alpha} - \mu).$$

In addition, defining  $w(x) \stackrel{\text{def.}}{=} v(\nabla \varphi^*(x))$ , it holds that

$$\frac{d^2}{dt^2} \mathcal{J}_{\mu, \alpha}(\varphi_t) = \text{Var}_{\varrho_{\varphi_t^*, \alpha}}(w) - \int_{\mathbb{R}^d} \left\langle (D^2 \varphi_t^*)^{-1} \nabla w, \nabla w \right\rangle d\varrho_{\varphi_t^*, \alpha}.$$

*Proof.* For fixed  $x \in \mathbb{R}^d$ , we start by computing the first and second time derivatives of  $\varphi_t^*(x)$  with an application of the envelope theorem. This quantity can be shown to be  $\mathscr{C}^2$  with an implicit function argument as done rigorously in [DM23].

Let  $y_t$  be optimal for the supremum defining  $\varphi_t^*(x)$ , that is

$$(2.16) \quad y_t = \underset{y \in \mathbb{R}^d}{\operatorname{argmax}} \langle y, x \rangle - \varphi_t(y), \text{ so that } y_t = \nabla \varphi_t^{-1}(x) = \nabla \varphi_t^*(x).$$

From the envelope theorem we have that

$$\frac{d}{dt} \varphi_t^*(x) = \partial_t (\langle y, x \rangle - \varphi_t(y))|_{y=y_t} = -v(\nabla \varphi_t^*(x)).$$

To compute the second time derivative, we then have

$$\frac{d^2}{dt^2} \varphi_t^*(x) = -\frac{d}{dt} v(\nabla \varphi_t^*(x)) = -\nabla v(\nabla \varphi_t^*(x)) \cdot y'_t.$$

Taking derivatives w.r.t.  $t$  on the optimality condition defining  $y_t$ , that is  $x = \nabla \varphi_t(y_t)$ , we get

$$y'_t = -D^2 \varphi_t(y_t)^{-1} \nabla v(y_t).$$

In addition, for the gradient of the quantity  $w$  we have that

$$\nabla w(x) = \nabla(v(y_t)) = D^2 \varphi_t^*(x)(\nabla v) \circ y_t.$$

Therefore, using that  $D^2 \varphi_t(y_t)^{-1} = D^2 \varphi_t^*(x)$ , the previous computation gives

$$\frac{d^2}{dt^2} \varphi_t^*(x) = \left\langle (D^2 \varphi_t^*)^{-1} \nabla w, \nabla w \right\rangle.$$

We can then compute the first and second time derivatives of  $\mathcal{J}_{\mu, \alpha}(\varphi + tv)$  via derivation under the integral sign as

$$\begin{aligned} \frac{d}{dt} \mathcal{J}_{\mu, \alpha}(\varphi_t) &= \frac{1}{Z_t} \int_{\mathbb{R}^d} e^{-(\varphi_t^* + \frac{\alpha}{2}|x|^2)} \left( -\frac{d}{dt} \varphi_t^* \right) - \int_{\mathbb{R}^d} v d\mu \\ &= \int_{\mathbb{R}^d} v d(\mu_{\varphi_t^*, \alpha} - \mu), \end{aligned}$$

for  $Z_t = \int_{\mathbb{R}^d} e^{-(\varphi_t^* + \frac{\alpha}{2}|x|^2)}$ .

For the second time derivative, we get that

$$\begin{aligned} \frac{d^2}{dt^2} \mathcal{J}_{\mu, \alpha}(\varphi_t) &= -\frac{Z'_t}{Z_t} \int_{\mathbb{R}^d} v(\nabla \varphi_t^*) d\varrho_{\varphi_t^*, \alpha} + \int_{\mathbb{R}^d} \left( v(\nabla \varphi_t^*)^2 - \frac{d^2}{dt^2} \varphi_t^* \right) d\varrho_{\varphi_t^*, \alpha} \\ &= \text{Var}_{\varrho_{\varphi_t^*, \alpha}}(w) - \int_{\mathbb{R}^d} \left\langle (D^2 \varphi_t^*)^{-1} \nabla w, \nabla w \right\rangle d\varrho_{\varphi_t^*, \alpha}. \end{aligned}$$

The result follows.  $\square$

For  $\alpha = 0$ , the above computation for the second variation of  $\mathcal{J}_\mu$  gives precisely minus the deficit of the Brascamp-Lieb inequality. In the case that  $\alpha > 0$  we need to be a bit careful to perform a Taylor expansion on the Hessian and get the same deficit plus lower order terms in  $\alpha$ . Either way, these computations give a clear intuition on why Cordero-Erasquin and Klartag's characterization should hold, but it does not provide the essential continuity of the potential, which is at the heart of their statement.

**Remark 2.6.** More importantly to our purposes, it becomes clear that having access to a sharp quantitative stability version of Brascamp-Lieb's inequality shall give stability results for the potentials of moment measure representations. As discussed in the introduction, there are a few next order expansions of the Brascamp-Lieb inequality in the literature [CE17, BGG18, Har08]. However, not only our Theorem 4.13 is more adapted to our purposes as we show that the deficit  $\delta_{BL}$  controls a distance, but also we do it with a multiplicative constant  $C_d$  that is independent of the measure  $\varrho_\varphi$ . This last feature is crucial for the applications to moment measures, as we shall see below.

**2.3. State of the art on the Prékopa-Leindler inequality.** The original version of the Prékopa-Leindler inequality was initially proposed independently by Prékopa in [Pré71, Pré73] and Leindler [Lei73] and can be stated as follows: given a triplet of measurable functions  $f, g, h : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that for a given parameter  $s \in (0, 1)$  the so-called *Prékopa condition* holds

$$(2.17) \quad f(x)^s g(y)^{1-s} \leq h(sx + (1-s)y) \text{ for all } x, y \in \mathbb{R}^d,$$

then we have the following bound

$$(2.18) \quad \left( \int_{\mathbb{R}^d} f(x) dx \right)^s \left( \int_{\mathbb{R}^d} g(x) dx \right)^{1-s} \leq \int_{\mathbb{R}^d} h(x) dx.$$

This inequality is a functional generalization of the Brunn-Minkowski inequality, for which many stability results have been derived, see for instance the seminal work [FMP10], where stability version of such geometric inequalities have been derived.

A stability version for the Prékopa-Leindler inequality quantifies how functions nearly satisfying the equality case must be close to an optimal log-concave profile. The result takes the following form: if  $h, f, g : \mathbb{R}^d \rightarrow \mathbb{R}_+$  satisfy Prékopa condition (2.17) and (2.18) is almost satisfied in the sense that

$$(2.19) \quad \int_{\mathbb{R}^d} h(x) dx \leq (1 + \varepsilon) \left( \int_{\mathbb{R}^d} f(x) dx \right)^s \left( \int_{\mathbb{R}^d} g(x) dx \right)^{1-s}.$$

for some  $\varepsilon > 0$ , there exists a log-concave function  $\tilde{h} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  and parameters  $a > 0$ ,  $x_0 \in \mathbb{R}^d$  such that the following proximity estimates hold:

$$(2.20) \quad \begin{aligned} \int_{\mathbb{R}^d} |f(x) - a^{-s}\tilde{h}(x - sx_0)| dx &\leq C(\tau)\varepsilon^{\alpha_n(\tau)} \int_{\mathbb{R}^d} f(x) dx, \\ \int_{\mathbb{R}^d} |g(x) - a^{1-s}\tilde{h}(x + (1-s)x_0)| dx &\leq C(\tau)\varepsilon^{\alpha_n(\tau)} \int_{\mathbb{R}^d} g(x) dx, \\ \int_{\mathbb{R}^d} |h(x) - \tilde{h}(x)| dx &\leq C(\tau)\varepsilon^{\alpha_n(\tau)} \int_{\mathbb{R}^d} h(x) dx, \end{aligned}$$

where

$$\tau \stackrel{\text{def.}}{=} \min(s, 1-s), \quad a \stackrel{\text{def.}}{=} \frac{\int_{\mathbb{R}^d} f}{\int_{\mathbb{R}^d} g}$$

$C(\tau)$  is a constant that depends on  $\tau$ , and  $\alpha_n(\tau) > 0$  is a computable exponent that typically depends on both the dimension  $n$  and the parameter  $\tau$ .

Obtaining sharper versions of this inequality is currently a very active area of research. On one side, the sharp exponent  $\alpha$  is expected to be  $1/2$ , as is the case, for instance, for the isoperimetric and Brunn-Minkowski inequalities [FMP10]. Indeed, this exponent has been obtained

in [FvHT25], but at the cost of a sub-optimal constant  $C(\tau)$  that does not behave properly as  $\tau \rightarrow 0^+$ . On the other hand, in [FR24], a quantitative stability version has been shown with  $\alpha = 1/2$  and  $C \propto \tau^{-1/2}$  in the 1D or radial in  $\mathbb{R}^d$ , when the involved functions are log-concave.

In the present work, we will only need a stability version of Prékopa-Leindler for the case  $\tau = 1/2$ . Therefore, the  $\tau$ -dependence of the constant  $C(\tau)$  is not particularly important for us, and we may hence use some of the results from [FvHT25] directly, which we collect in the following Lemma.

**Lemma 2.7.** *Let  $f, g, h : \mathbb{R}^d \rightarrow \mathbb{R}_+$  be measurable functions satisfying Prékopa's condition (2.17) with  $s = 1/2$ . If there is  $\varepsilon > 0$  such that (2.19) holds, there exists  $x_0 \in \mathbb{R}^d$  and a dimensional constant  $C_d$  such that*

$$\int_{\mathbb{R}^d} \left| \frac{f(x)}{\int_{\mathbb{R}^d} f} - \frac{g(x + x_0)}{\int_{\mathbb{R}^d} g} \right| dx \leq C_d \varepsilon^{1/2}.$$

*Proof.* Recall  $a \stackrel{\text{def.}}{=} \frac{\int_{\mathbb{R}^d} f}{\int_{\mathbb{R}^d} g}$ , so from [FvHT25], estimates (2.20) hold with  $\alpha = 1/2$  and a constant  $C(\tau) = C(1/2) = C_d$ , that depends only on the dimension. So there is a log-concave function  $\tilde{h}$  such that

$$\begin{aligned} \int_{\mathbb{R}^d} \left| \frac{f(x)}{\int_{\mathbb{R}^d} f} - \frac{\tilde{h}(x - \frac{1}{2}x_0)}{(\int_{\mathbb{R}^d} f \int_{\mathbb{R}^d} g)^{1/2}} \right| dx &\leq C_d \varepsilon^{1/2}, \\ \int_{\mathbb{R}^d} \left| \frac{g(x)}{\int_{\mathbb{R}^d} g} - \frac{\tilde{h}(x + \frac{1}{2}x_0)}{(\int_{\mathbb{R}^d} f \int_{\mathbb{R}^d} g)^{1/2}} \right| dx &\leq C_d \varepsilon^{1/2}. \end{aligned}$$

A change of variables on the inequality concerning  $g$  and an easy triangular inequality give the desired bound.  $\square$

### 3. SHARP QUANTITATIVE STABILITY FOR BRASCAMP-LIEB INEQUALITY

In this section we are focused on the quantitative stability of the Brascamp-Lieb inequality, as stated in Theorem 1.1. Due to its importance to derive quantity stability results, and for the reader's convenience, we recall how this can be obtained with the *carré du champ* method, which is based on the geometry of the following operator

$$Lu \stackrel{\text{def.}}{=} \Delta u - \nabla \varphi \cdot \nabla u.$$

This method is based on the so called  $\Gamma$ -calculus, which revolves around the following commutator operators

$$\begin{aligned} \Gamma(f, g) &\stackrel{\text{def.}}{=} \frac{1}{2} (L(fg) - fLg - gLf) = \langle \nabla f, \nabla g \rangle, \\ \Gamma_2(f, g) &\stackrel{\text{def.}}{=} \frac{1}{2} (L\Gamma(f, g) - \Gamma(f, Lg) - \Gamma(g, Lf)) = D^2 f : D^2 g + \langle D^2 \varphi \nabla f, \nabla g \rangle. \end{aligned}$$

It is then a straight-forward computation to check the integration by parts formula

$$(3.1) \quad \int_{\mathbb{R}^d} \Gamma(f, g) d\varrho_\varphi = \int_{\mathbb{R}^d} \langle f, g \rangle d\varrho_\varphi = - \int_{\mathbb{R}^d} f L g d\varrho_\varphi = - \int_{\mathbb{R}^d} g L f d\varrho_\varphi.$$

**Lemma 3.1.** *Inequality (1.1) holds, and we have equality if, and only if, there exists  $a \in \mathbb{R}^d$  such that*

$$f(x) = a \cdot \nabla \varphi(x) + \mathbb{E}_{\varrho_\varphi} f.$$

*Proof.* Given  $f \in \mathscr{C}_c^\infty(\mathbb{R}^d)$ , set  $\bar{f} = f - \mathbb{E}_{\varrho_\varphi} f$  and let  $u$  be the unique weak solution of

$$Lu = \bar{f}.$$

From the classical regularity theory of uniformly elliptic equations,  $u \in \mathcal{C}^\infty(\mathbb{R}^d)$ . So we can use the  $\Gamma$  calculus tools to write the variance of  $f$  as

$$\begin{aligned}\text{Var}_{\varrho_\varphi}(f) &= \int_{\mathbb{R}^d} \bar{f}^2 d\varrho_\varphi = \int_{\mathbb{R}^d} \bar{f} L u d\varrho_\varphi = - \int_{\mathbb{R}^d} \langle \nabla f, \nabla u \rangle d\varrho_\varphi \\ &= - \int_{\mathbb{R}^d} \left\langle (D^2 \varphi)^{-1/2} \nabla f, (D^2 \varphi)^{1/2} \nabla u \right\rangle d\varrho_\varphi \\ &\leq \left( \int_{\mathbb{R}^d} \left\| (D^2 \varphi)^{-1/2} \nabla f \right\|^2 d\varrho_\varphi \right)^{1/2} \left( \int_{\mathbb{R}^d} \left\| (D^2 \varphi)^{1/2} \nabla u \right\|^2 d\varrho_\varphi \right)^{1/2} \\ &= \left( \int_{\mathbb{R}^d} \left\langle (D^2 \varphi)^{-1} \nabla f, \nabla f \right\rangle d\varrho_\varphi \right)^{1/2} \left( \int_{\mathbb{R}^d} \langle D^2 \varphi \nabla u, \nabla u \rangle d\varrho_\varphi \right)^{1/2}.\end{aligned}$$

Now we notice that

$$\int_{\mathbb{R}^d} \Gamma_2(u) d\varrho_\varphi = \frac{1}{2} \int_{\mathbb{R}^d} L |\nabla u|^2 d\varrho_\varphi - \int_{\mathbb{R}^d} \langle \nabla u, \nabla L u \rangle d\varrho_\varphi = - \int_{\mathbb{R}^d} \langle \nabla u, \nabla L u \rangle d\varrho_\varphi$$

since the integrand of the first integral is a divergence. In addition, by definition of  $\Gamma_2$  we have that

$$\begin{aligned}\int_{\mathbb{R}^d} \langle D^2 \varphi \nabla u, \nabla u \rangle d\varrho_\varphi &= \int_{\mathbb{R}^d} \Gamma_2(u) d\varrho_\varphi - \int_{\mathbb{R}^d} \|D^2 u\|_F^2 d\varrho_\varphi \\ &= - \int_{\mathbb{R}^d} \langle \nabla u, \nabla L u \rangle d\varrho_\varphi - \int_{\mathbb{R}^d} \|D^2 u\|_F^2 d\varrho_\varphi \\ &= \text{Var}_{\varrho_\varphi}(f) - \int_{\mathbb{R}^d} \|D^2 u\|_F^2 d\varrho_\varphi.\end{aligned}$$

Plugging it back into the original estimations of  $\text{Var}_{\varrho_\varphi}(f)$ , we notice that

$$\text{Var}_{\varrho_\varphi}(f) \leq \left( \int_{\mathbb{R}^d} \left\langle (D^2 \varphi)^{-1} \nabla f, \nabla f \right\rangle d\varrho_\varphi \right)^{1/2} \left( \text{Var}_{\varrho_\varphi}(f) - \int_{\mathbb{R}^d} \|D^2 u\|_F^2 d\varrho_\varphi \right)^{1/2}$$

which gives inequality (1.1). This proof is also enlightening on the equality cases; indeed notice that the chain of inequalities are all equalities, except for the application of a Cauchy-Scharwz and the non-positive term on  $-\|D^2 u\|_F^2$ .

As a result, we have equality if, and only if,

$$D^2 u \equiv 0 \text{ and } (D^2 \varphi)^{-1/2} \nabla f \parallel (D^2 \varphi)^{1/2} \nabla u.$$

The first condition implies that  $u = a \cdot x + b$  for constant  $a \in \mathbb{R}^d$  and  $b \in \mathbb{R}$ . Since  $u$  is a classical solution of the equation  $L u = \bar{f}$ , it follows that, to have equality  $f$  must satisfy

$$f = a \cdot \nabla \varphi(x) + \mathbb{E}_{\varrho_\varphi} f.$$

The result follows.  $\square$

With the knowledge of the equality case for (1.1), we now aim at a quantitative stability version of this inequality in terms of the  $L^1$  distance to the optimal set (see Remark 2.6), as highlighted in Theorem 1.1. We state that result below once more for the reader's convenience.

**Theorem 3.2.** *Let  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  be an essentially continuous and convex function, and  $\int_{\mathbb{R}^d} e^{-\varphi} < +\infty$  so that the Gibbs mesuare  $\varrho_\varphi$  is well defined as in (1.4).*

*Then, there exists a universal constant  $C_d$  depending only on the ambient dimension such that*

$$\text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{BL}) \leq C_d \delta_{BL}(f)^{1/2}$$

*for all locally Lipschitz functions  $f \in L^2(\varrho_\varphi)$ .*

Our proof of Theorem 3.2 will be based, as mentioned in the introduction, on the linealization argument from Bobkov-Ledoux [BL00], while employing the main results from [FvHT25] in order to control the deficit, and carefully using the geometry induced by the potential in order to control the several stability parameters involved.

*Proof.* We first prove the result of the case that  $\varphi$  is strongly convex and essentially continuous function. That is, we assume that there exists  $\alpha > 0$  such that  $D^2\varphi \geq \alpha \text{id}$ , where the gradients and hessians  $\nabla\varphi, D^2\varphi$  are understood in the sense of Alexandrov (see for instance [EG15, Chapter 6]). Therefore  $\mathcal{L}^d$  and  $\varrho_\varphi$  are a.e. well defined.

Given  $f \in L^2(\varrho_\varphi)$ , it can be arbitrarily approximated in  $L^2(\varrho_\varphi)$  by a sequence of smooth functions, so for the moment we assume  $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ . As a result, from the  $\alpha$ -strong convexity of  $\varphi$  there exists  $\delta_0 > 0$  such that for all  $\delta < \delta_0$  the function  $2\delta f - \varphi$  is concave.

Our goal is to apply the quantitative stability version of the Prékopa-Leindler inequality with convexity parameter  $s = 1/2$ . In order to do so, we consider the functions

$$u_\delta \stackrel{\text{def.}}{=} e^{2\delta f - \varphi}, \quad v_\delta \stackrel{\text{def.}}{=} e^{-\varphi}, \quad w_\delta \stackrel{\text{def.}}{=} e^{f_\delta - \varphi},$$

where  $f_\delta(z)$  is pointwise defined as

$$\begin{aligned} f_\delta(z) &= \sup_{z=\frac{x+y}{2}} \delta f(x) - \left[ \frac{1}{2}\varphi(x) + \frac{1}{2}\varphi(y) - \varphi\left(\frac{x+y}{2}\right) \right] \\ &= \sup_{h \in \mathbb{R}^d} \delta f(z+h) - \left[ \frac{1}{2}\varphi(z+h) + \frac{1}{2}\varphi(z-h) - \varphi(z) \right]. \end{aligned}$$

It is not too hard to check that the triple of functions  $u_\delta, v_\delta, w_\delta$  satisfies the condition in Prékopa–Leindler's inequality. Hence, as the total mass of  $v_\delta$  is 1 by assumption, the standard Prékopa–Leindler inequality holds and we can define the non-negative quantity  $\varepsilon_\delta$  such that

$$(3.2) \quad \int_{\mathbb{R}^d} w_\delta dx = (1 + \varepsilon_\delta) \left( \int_{\mathbb{R}^d} u_\delta dx \right)^{1/2}, \quad \varepsilon_\delta \stackrel{\text{def.}}{=} \frac{\int_{\mathbb{R}^d} w_\delta dx - (\int_{\mathbb{R}^d} u_\delta dx)^{1/2}}{(\int_{\mathbb{R}^d} u_\delta dx)^{1/2}}.$$

Analogously as in Bobkov-Ledoux [BL00], our argument to obtain the sharp quantity version of the Brascamp–Lieb inequality will consist of studying the dependence on  $\delta$  of the above quantities. Starting with the integral of  $u_\delta$ , we get the following Taylor expansion on  $\delta$ :

$$(3.3) \quad \int_{\mathbb{R}^d} u_\delta dx = 1 + 2\delta \int_{\mathbb{R}^d} f d\varrho_\varphi + 2\delta^2 \int_{\mathbb{R}^d} f^2 d\varrho_\varphi + o(\delta^2).$$

For the expansion on the integral of  $w_\delta$ , recall the classical result from log-concave measures that

$$\text{if } \int_{\mathbb{R}^d} e^{-\varphi} < +\infty, \text{ then } \varphi \text{ is coercive.}$$

Therefore, the quantity being maximized in the definition of  $f_\delta$  is strongly concave and goes to  $-\infty$  as  $h \rightarrow +\infty$ . Hence for  $\varrho_\varphi$ -a.e.  $z$  let  $h_\delta = h_\delta(z)$  be optimal for the supremum in  $f_\delta(z)$ . The optimality conditions for  $h_\delta$  give, for  $\varrho_\varphi$ -a.e. point where the gradients and hessian of  $\varphi$  are uniquely defined, that

$$\delta \nabla f(z + h_\delta) = \frac{1}{2} [\nabla \varphi(z + h_\delta) - \nabla \varphi(z - h_\delta)] = D^2 \varphi(z) h_\delta + o(\delta).$$

Since  $D^2\varphi \geq \alpha \text{id}$ , the hessian is invertible and  $(D^2\varphi)^{-1} \leq \alpha^{-1} \text{id}$ , and developing the Taylor expansion for  $\nabla f$  in the identity above we obtain that

$$h_\delta = \delta (D^2 \varphi)^{-1} \nabla f(z) + o(\delta).$$

In particular  $h_\delta = O(\delta)$ , uniformly in  $z$ . As a result, we can expand the value of  $f_\delta(z)$  as

$$f_\delta(z) = \delta f(z) + \frac{\delta^2}{2} \langle (D^2 \varphi)^{-1} \nabla f, \nabla f \rangle + o(\delta^2).$$

Therefore, with a similar argument to the expansion of the integral of  $u_\delta$ , we get

$$\begin{aligned} \int_{\mathbb{R}^d} w_\delta dx &= \int_{\mathbb{R}^d} e^{f_\delta} d\varrho_\varphi \\ &= 1 + \delta \int_{\mathbb{R}^d} f d\varrho_\varphi + \frac{\delta^2}{2} \int_{\mathbb{R}^d} f^2 d\varrho_\varphi + \frac{\delta^2}{2} \int_{\mathbb{R}^d} \langle (D^2\varphi)^{-1} \nabla f, \nabla f \rangle d\varrho_\varphi + o(\delta^2). \end{aligned}$$

In addition, using the Taylor expansion  $(1+s)^{1/2} = 1 + \frac{1}{2}s - \frac{1}{8}s^2 + o(s^2)$  and the expansion for the integral of  $u_\delta$  we obtain that

$$\left( \int_{\mathbb{R}^d} u_\delta dx \right)^{1/2} = 1 + \delta \int_{\mathbb{R}^d} f d\varrho_\varphi + \delta^2 \int_{\mathbb{R}^d} f^2 d\varrho_\varphi - \frac{\delta^2}{2} \left( \int_{\mathbb{R}^d} f d\varrho_\varphi \right)^2 + o(\delta^2).$$

Joining these estimates, we obtain the following expansion for  $\varepsilon_\delta$ :

$$\begin{aligned} \varepsilon_\delta &= \left( \int_{\mathbb{R}^d} u_\delta dx \right)^{-1/2} \left( \int_{\mathbb{R}^d} w_\delta dx - \left( \int_{\mathbb{R}^d} u_\delta dx \right)^{1/2} \right) \\ &= (1 + O(\delta)) \frac{\delta^2}{2} \left[ \int_{\mathbb{R}^d} \langle (D^2\varphi)^{-1} \nabla f, \nabla f \rangle d\varrho_\varphi - \text{Var}_{\varrho_\varphi}(f) + \frac{o(\delta^2)}{\delta^2} \right] \\ &= (1 + O(\delta)) \frac{\delta^2}{2} \left[ \delta_{\text{BL}}(f) + \frac{o(\delta^2)}{\delta^2} \right] \end{aligned}$$

From the quantitative stability of the Prékopa-Leindler inequality we obtain that there exists a universal constant  $C_d$ , depending only on the dimension and on the convexity parameter which is fixed  $s = 1/2$ , and some vector  $x_\delta \in \mathbb{R}^d$  such that

$$(3.4) \quad \int_{\mathbb{R}^d} \left| \frac{u_\delta(x)}{\int_{\mathbb{R}^d} u_\delta} - e^{-\varphi(x+x_\delta)} \right| dx \leq C_d \varepsilon_\delta^{1/2}.$$

Taking into account the scaling of  $\varepsilon_\delta$ , which behaves as  $\delta^2$ , we see that the RHS of the above inequality is of the order  $\delta$ . Developing the LHS depends strongly on the behavior of  $x_\delta$  when  $\delta \rightarrow 0$ . Clearly it must hold that  $x_\delta \rightarrow 0$  as  $\delta \rightarrow 0$ , as a result we obtain that

$$\begin{aligned} \int_{\mathbb{R}^d} \left| \frac{u_\delta(x)}{\int_{\mathbb{R}^d} u_\delta} - e^{-\varphi(x+x_\delta)} \right| dx &= \left( \int_{\mathbb{R}^d} u_\delta \right)^{-1} \int_{\mathbb{R}^d} \left| e^{2\delta f} - e^{-(\varphi(x+x_\delta)-\varphi(x))} \int_{\mathbb{R}^d} u_\delta \right| d\varrho_\varphi \\ &= (1 + O(\delta)) \int_{\mathbb{R}^d} \left| 1 + 2\delta f + o(\delta) - (1 + \nabla\varphi(x) \cdot x_\delta + o(x_\delta)) (1 + 2\delta \mathbb{E}_{\varrho_\varphi} f + o(\delta)) \right| d\varrho_\varphi \\ &= (1 + O(\delta)) 2\delta \int_{\mathbb{R}^d} \left| f - \nabla\varphi(x) \cdot \frac{x_\delta}{2\delta} - \mathbb{E}_{\varrho_\varphi} f \right| d\varrho_\varphi + o(\delta + x_\delta). \end{aligned}$$

Therefore, in order to obtain a meaningful estimate in terms of  $\text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{\text{BL}})$ , we must show that  $x_\delta/2\delta$  has a cluster point as  $\delta \rightarrow 0^+$ . For this, define the sets

$$\Omega_n \stackrel{\text{def.}}{=} \{x \in \text{dom } \varphi : D^2\varphi(x) \leq n \text{id}\}.$$

From the convexity of  $\varphi$ , for every  $n \in \mathbb{N}$  we can choose  $\delta$  sufficiently small so that  $\varphi(x + x_\delta) - \varphi(x) = O(x_\delta)$ . More precisely, there is  $\delta_n$  such that for  $\delta \leq \delta_n$  we have

$$\begin{aligned} \|v_\delta - v_\delta(\cdot + x_\delta)\|_{L^1(\mathbb{R}^d)} &\geq \|v_\delta - v_\delta(\cdot + x_\delta)\|_{L^1(\Omega_n)} \\ &= \int_{\Omega_n} \left| x_\delta \cdot \nabla\varphi(x) + \frac{1}{2} \langle D^2\varphi(x)x_\delta, x_\delta \rangle + o(x_\delta^2) \right| d\varrho_\varphi \\ &\geq \frac{1}{2} \int_{\Omega_n} |x_\delta \cdot \nabla\varphi(x)| d\varrho_\varphi. \end{aligned}$$

Using a suitable triangle inequality, we get for  $\delta \leq \delta_n$

$$\begin{aligned} \frac{1}{2} \int_{\Omega_n} |x_\delta \cdot \nabla \varphi(x)| d\varrho_\varphi &\leq \|v_\delta - v_\delta(\cdot + x_\delta)\|_{L^1(\mathbb{R}^d)} \\ &\leq \|v_\delta - u_\delta\|_{L^1(\mathbb{R}^d)} + \|u_\delta - v_\delta(\cdot + x_\delta)\|_{L^1(\mathbb{R}^d)} \leq C\delta, \end{aligned}$$

where the bound  $\|v_\delta - u_\delta\|_{L^1(\mathbb{R}^d)} = O(\delta)$  follows directly from the definition of  $u_\delta$  as a perturbation of  $v_\delta$  and  $\|u_\delta - v_\delta(\cdot + x_\delta)\|_{L^1(\mathbb{R}^d)} = O(\delta)$  is a consequence of the previous estimations on the deficit of the Prékopa-Leindler inequality. But then we conclude that the quantity

$$\int_{\Omega_n} \left| \frac{x_\delta}{2\delta} \cdot \nabla \varphi(x) \right| d\varrho_\varphi$$

must remain uniformly bounded for  $\delta \leq \delta_n$ . Now, we wish to exploit the fact that, since  $\varphi$  is essentially continuous, its moment measure  $\mu_\varphi$  has full-dimensional support, as highlighted in Lemma 2.3. Since  $x_\delta / \|x_\delta\| \in \mathbb{S}^{d-1}$ , we can find a subsequence of  $\delta_n \xrightarrow{k \rightarrow \infty} 0$  such that  $x_{\delta_n} / \|x_{\delta_n}\| \xrightarrow{k \rightarrow \infty} \bar{\theta}$ , for some  $\bar{\theta} \in \mathbb{S}^{d-1}$ . But then using Lebesgue's dominated convergence theorem and the fact that  $\mathbf{1}_{\Omega_n}$  converges strongly in  $L^1(\varrho_\varphi)$  to 1, we get that

$$\begin{aligned} C &\geq \limsup_{n \rightarrow \infty} \int_{\Omega_n} \left| \frac{x_{\delta_n}}{2\delta_n} \cdot \nabla \varphi(x) \right| d\varrho_\varphi = \limsup_{n \rightarrow \infty} \frac{\|x_{\delta_n}\|}{2\delta_n} \int_{\Omega_n} \left| \frac{x_{\delta_n}}{\|x_{\delta_n}\|} \cdot \nabla \varphi(x) \right| d\varrho_\varphi \\ &= \left( \int_{\mathbb{R}^d} |\bar{\theta} \cdot y| d\mu_\varphi(y) \right) \limsup_{n \rightarrow \infty} \frac{\|x_{\delta_n}\|}{2\delta_n} \end{aligned}$$

Since the moment measure  $\mu_\varphi$  has full-dimensional support,  $x_{\delta_n} / \delta_n$  remains bounded as  $n \rightarrow \infty$ .

Therefore, we can extract a subsequence  $\delta_n \xrightarrow{k \rightarrow \infty} 0$  such that  $\frac{x_{\delta_n}}{2\delta_n} \xrightarrow{k \rightarrow \infty} a$ , for some vector  $a \in \mathbb{R}^d$ . Coming back to the estimate on  $\varepsilon_\delta$  from (3.4), dividing both sides by  $\delta$  and passing to the limit of these estimates as  $\delta_n \rightarrow 0$  we obtain that

$$\begin{aligned} \text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{\text{BL}}) &\leq \int_{\mathbb{R}^d} |(f - \nabla \varphi(x) \cdot a - \mathbb{E}_{\varrho_\varphi} f)| d\varrho_\varphi \\ &\leq \liminf_{k \rightarrow \infty} (1 + O(\delta_n)) \int_{\mathbb{R}^d} \left| \left( f - \nabla \varphi(x) \cdot \frac{x_{\delta_n}}{2\delta_n} - \mathbb{E}_{\varrho_\varphi} f \right) \right| d\varrho_\varphi + o(\delta_n) \\ &\leq \limsup_{k \rightarrow \infty} C_d \left( \delta_{\text{BL}}(f) + \frac{o(\delta_n^2)}{\delta_n^2} \right)^{1/2} \\ &= C_d \delta_{\text{BL}}(f)^{1/2}, \end{aligned}$$

and the result follows when  $\varphi$  is  $\alpha$ -strongly convex and  $f \in L^2(\varrho_\varphi) \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$ . By approximation, it follows also for all  $f \in L^2(\varrho_\varphi)$ .

For an arbitrary  $\varphi$  convex and essentially continuous, we finish the proof with an approximation argument. In order to make it clearer, we modify the notation used previously in order for it to emphasize the dependence on the potential  $\varphi$ . Hence, we shall write

$$\delta_{\text{BL}}(f; \varphi) \text{ and } \text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{\text{BL}, \varphi}),$$

in order to denote the deficit of the Brascamp-Lieb inequality and the distance to optimal functions with respect to the potential  $\varphi$ , respectively. Our strategy is to add a little quadratic regularization by defining  $\varphi_\alpha \stackrel{\text{def.}}{=} \varphi + \frac{\alpha}{2} |\cdot|^2$ . The result will follow if we prove that the deficit is upper semi-continuous

$$(3.5) \quad \limsup_{\alpha \rightarrow 0^+} \delta_{\text{BL}}(f; \varphi_\alpha) \leq \delta_{\text{BL}}(f; \varphi),$$

and that the distance to optimal functions is lower semi-continuous

$$(3.6) \quad \text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{\text{BL}, \varphi}) \leq \liminf_{\alpha \rightarrow 0^+} \text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{\text{BL}, \varphi_\alpha}).$$

For the first identity (3.5), first notice that the Gibbs measures  $\varrho_{\varphi_\alpha}$  converge pointwise, even uniformly, to  $\varrho_\varphi$  as  $\alpha \rightarrow 0^+$ . Then, for any  $f \in L^2(\varrho_\varphi)$ , convergence of the variance follows from Lebesgue's dominated convergence theorem:

$$\text{Var}_{\varrho_{\varphi_\alpha}}(f) \xrightarrow[\alpha \rightarrow 0^+]{\longrightarrow} \text{Var}_{\varrho_\varphi}(f).$$

For the positive term in the deficit, notice that  $D^2\varphi_\alpha = D^2\varphi + \alpha \text{id} \geq D^2\varphi$ , hence  $D^2\varphi_\alpha$  is invertible and it follows that  $(D^2\varphi_\alpha)^{-1} \leq (D^2\varphi)^{-1}$  in the partial ordering of positive semi-definite matrices, even if  $D^2\varphi$  becomes singular. As a result, assuming that  $\delta_{\text{BL}}(f; \varphi) < +\infty$ , we have for all  $\alpha > 0$  and  $f \in L^2(\varrho_\varphi)$  that

$$\int_{\mathbb{R}^d} \left\langle (D^2\varphi_\alpha)^{-1} \nabla f, \nabla f \right\rangle d\varrho_{\varphi_\alpha} \leq \int_{\mathbb{R}^d} \left\langle (D^2\varphi)^{-1} \nabla f, \nabla f \right\rangle d\varrho_{\varphi_\alpha}$$

and passing to the limit as  $\alpha \rightarrow 0^+$  yields (3.5).

Moving on the second identity (3.6), notice that the finite dimensional function

$$(a, b) \mapsto \|f - (a \cdot \nabla \varphi_\alpha + b)\|_{L^1(\varrho_{\varphi_\alpha})}$$

is convex, due to the triangle inequality, and coercive. So the infimum is attained at some  $(a_\alpha, b_\alpha)$  and, in particular, we know that it must be the case that  $b_\alpha = \mathbb{E}_{\varrho_{\varphi_\alpha}} f \xrightarrow[\alpha \rightarrow 0^+]{\longrightarrow} \mathbb{E}_{\varrho_\varphi} f \stackrel{\text{def.}}{=} \bar{b}$ . In addition, the optimal  $a_\alpha$  must be all contained in the same compact ball since  $(\varrho_{\varphi_\alpha})_{\alpha > 0}$  have equibounded moments. So we can assume w.l.o.g. that  $a_\alpha \xrightarrow[\alpha \rightarrow 0^+]{\longrightarrow} \bar{a}$ . It then follows from Fatou's Lemma that

$$\text{dist}_{L^1(\varrho_\varphi)}(f, \mathcal{O}_{\text{BL}, \varphi}) \leq \|f - (\bar{a} \cdot \nabla \varphi + \bar{b})\|_{L^1(\varrho_\varphi)} \leq \liminf_{\alpha \rightarrow 0^+} \text{dist}_{L^1(\varrho_{\varphi_\alpha})}(f, \mathcal{O}_{\text{BL}, \varphi_\alpha}).$$

The result follows.  $\square$

#### 4. QUANTITATIVE STABILITY OF MOMENT MEASURES

In this section we use the previous machinery to obtain sharp quantitative stability results for moment measures. We start by proving Theorem 1.2. As usual, we shall state it again below for the reader's convenience:

**Theorem 4.1.** *Given a Radon measure  $\mu \in \mathscr{P}_1(\mathbb{R}^d)$  whose barycenter lies that the origin and  $\dim \text{supp } \mu = d$ , let  $\bar{\varphi}$  be a maximizer of  $\mathcal{J}_\mu$ . Then for any convex function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  it holds that*

$$(4.1) \quad \text{dist}_{L^1(\mathbb{R}^d)}(\varrho_{\varphi^*}, \mathcal{M}_\mu) \leq C_d (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi))^{1/2}.$$

*In addition, there exists  $\lambda \in (0, 1/2)$  such that, setting  $v \stackrel{\text{def.}}{=} \bar{\varphi} - \varphi$  and  $\varphi_\lambda \stackrel{\text{def.}}{=} \bar{\varphi} + \lambda v$ , it holds that*

$$(4.2) \quad \text{dist}_{L^1(\mu_\lambda)}(\varphi, \mathcal{N}_\mu) \leq C_d (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi))^{1/2},$$

*where  $\mu_\lambda = \mu_{\varphi_\lambda^*}$  is the moment measure associated with the interpolation  $(\varphi_\lambda)^*$ .*

*Proof.* To prove the first estimate (4.1), we let  $\bar{\varphi}$  be a maximizer of  $\mathcal{J}_\mu$ , while  $\varphi$  is a general convex function. We then employ the quantitative version of the Prékopa-Leindler inequality with the functions

$$f = e^{-\bar{\varphi}^*}, \quad g = e^{-\varphi^*} \text{ and } h = e^{-\varphi_{1/2}^*} \text{ with } \varphi_{1/2} = \frac{\bar{\varphi} + \varphi}{2}.$$

Therefore  $f, g, h$  satisfy the Prékopa condition with convexity parameter  $s = 1/2$ . In addition, due to concavity of  $\mathcal{J}_\mu$  and the optimality of  $\bar{\varphi}$  we have that

$$\begin{aligned} 0 \leq \delta &\stackrel{\text{def.}}{=} \mathcal{J}_\mu(\varphi_{1/2}) - \frac{1}{2} (\mathcal{J}_\mu(\bar{\varphi}) + \mathcal{J}_\mu(\varphi)) \\ &\leq \mathcal{J}_\mu(\bar{\varphi}) - \frac{1}{2} (\mathcal{J}_\mu(\bar{\varphi}) + \mathcal{J}_\mu(\varphi)) = \frac{1}{2} (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi)) \stackrel{\text{def.}}{=} \varepsilon. \end{aligned}$$

The quantity  $\delta$  can be precisely related to the deficit of the Prékopa-Leindler inequality. Indeed:

$$\begin{aligned} \delta &= \log \left( \int_{\mathbb{R}^d} e^{-\varphi_{1/2}^*} dx \right) - \log \left( \int_{\mathbb{R}^d} e^{-\bar{\varphi}^*} dx \cdot \int_{\mathbb{R}^d} e^{-\varphi^*} dx \right)^{1/2} \\ &\quad - \underbrace{\int_{\mathbb{R}^d} (\varphi_{1/2} - \frac{1}{2}(\bar{\varphi} + \varphi)) d\mu}_{=0} \\ &= \log \left( \frac{\int_{\mathbb{R}^d} e^{-\varphi_{1/2}^*} dx}{\left( \int_{\mathbb{R}^d} e^{-\bar{\varphi}^*} dx \cdot \int_{\mathbb{R}^d} e^{-\varphi^*} dx \right)^{1/2}} \right) \leq \varepsilon. \end{aligned}$$

Now, notice that since  $\varrho_{\bar{\varphi}^*}$  and all elements in  $\mathcal{M}_\mu$  are probability measures, their distance in  $L^1(\mathbb{R}^d)$  is at most 2. As a result, we can assume that  $\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi) < 1$  otherwise the inequality is trivial. From the previous estimates we have that

$$\int_{\mathbb{R}^d} e^{-\varphi_{1/2}^*} dx < (1 + \varepsilon) \left( \int_{\mathbb{R}^d} e^{-\bar{\varphi}^*} dx \cdot \int_{\mathbb{R}^d} e^{-\varphi^*} dx \right)^{1/2}.$$

Then, the quantitative version of Prékopa-Leindler inequality in the form of Lemma 2.7 implies that there exists a universal constant  $C_d$  depending only on the dimension and  $x_0 \in \mathbb{R}^d$  such that

$$\int_{\mathbb{R}^d} |\varrho_{\bar{\varphi}^*}(x) - \varrho_{\varphi^*}(x + x_0)| dx \leq C_d (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi))^{1/2}.$$

Estimate (4.1) follows.

For the second estimate recall the interpolation  $\varphi_t = \bar{\varphi} + tv$ , so that  $\varphi_0 = \bar{\varphi}$ , and define the one-dimensional function

$$t \mapsto J(t) \stackrel{\text{def.}}{=} \mathcal{J}_\mu(\varphi_t).$$

Hence a second order Taylor expansion and the characterization of the second variation of  $\mathcal{J}_\mu$  from Lemma 2.5 we obtain that

$$\begin{aligned} \mathcal{J}_\mu(\varphi) - \mathcal{J}_\mu(\bar{\varphi}) &= J(1) - J(0) = J'(0) + \int_0^1 (1-t) J''(t) dt \\ &= \underbrace{\langle \nabla \mathcal{J}_\mu(\bar{\varphi}), v \rangle}_{=0} + \int_0^1 (1-t) \frac{d^2}{dt^2} \mathcal{J}_\mu(\varphi_t) dt \\ &= - \int_0^1 (1-t) \delta_{BL} (v(\nabla \varphi_t^*), \varrho_{\varphi_t^*}) dt, \end{aligned}$$

where we have used the fact that if  $\bar{\varphi} \in \operatorname{argmax} \mathcal{J}_\mu$  then the moment measure  $\mu_{\varphi^*}$  coincides with  $\mu$  so that the first variation vanishes. In the sequel, using the quantitative stability version of the Brascamp-Lieb inequality provided by Theorem 3.2, there exists a dimensional constant  $C_d$  such that

$$\int_0^1 (1-t) \inf_{a \in \mathbb{R}^d, b \in \mathbb{R}} \|(\varphi - \bar{\varphi}) - (a \cdot x + b)\|_{L^1(\mu_{\varphi_t^*})}^2 dt \leq C_d^2 (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi)).$$

In particular, there must exist some  $\lambda \in (0, 1/2)$  such that the integrand above evaluated at the interpolation  $\mu_{\varphi_\lambda^*}$  yields

$$\inf_{a \in \mathbb{R}^d, b \in \mathbb{R}} \|(\varphi - \bar{\varphi}) - (a \cdot x + b)\|_{L^1(\mu_{\varphi_\lambda^*})}^2 \leq 2C_d (\mathcal{J}_\mu(\bar{\varphi}) - \mathcal{J}_\mu(\varphi))^{1/2}.$$

This finishes our proof.  $\square$

**4.1. Stability in a compact domain.** The simplest case of stability is when both  $\varrho$  and  $\mu$  are restricted to be supported on a convex and compact domain  $\Omega$ , that is  $\varrho, \mu \in \mathcal{P}(\Omega)$ . The attentive reader might have hesitated if this is even possible without contradicting the uniqueness result of Cordero-Erasquin and Klartag for moment measure representations. Here the question is to find  $\psi \in \mathcal{C}(\Omega)$  such that

$$(4.3) \quad \mu = (\nabla \psi)_\sharp e^{-\psi} \mathcal{L}^d \llcorner \Omega.$$

The existence question follows with simplified versions of the arguments from [CEK15, San16] by considering the functionals

$$\mathcal{E}_{\mu, \Omega}(\varrho) \stackrel{\text{def.}}{=} \int_{\Omega} \log \varrho d\varrho + \mathcal{T}(\varrho, \mu), \quad \mathcal{J}_{\mu, \Omega}(\varphi) \stackrel{\text{def.}}{=} \log \int_{\Omega} e^{-\varphi^*} dx - \int_{\Omega} \varphi d\mu.$$

Uniqueness then follows, now without modulo translations, due to the strict geodesic convexity of the entropy since  $\Omega$  is convex.

One could embed  $\mu$  in  $\mathcal{P}(\Omega')$  for any  $\Omega'$  containing its support and obtain another moment measure representation. What is happening here? For any representation in a bounded domain the potential  $\psi_\Omega$  is necessarily Lipschitz continuous, for instance due to  $\nabla \psi_\Omega$  being the Brenier map with the compactly supported target  $\mu$ , and hence bounded in  $\Omega$ . Therefore, if one wishes to use  $\psi_\Omega$  for a moment measure representation in  $\mathbb{R}^d$  we must extend it with  $+\infty$  outside of  $\Omega$  so that  $e^{-\varphi_\Omega^*}$  remains a probability measure. But this means that this extension is not continuous in all of  $\partial\Omega$ , and hence it is not essentially continuous. Therefore have no contradiction with the results of [CEK15], since their result states uniqueness up to translations of a moment measure representation with an essentially continuous potential.

From one hand, this case is much less interesting from a strictly mathematical point of view, since the compactness of the domain makes many arguments easier. However, it is particularly relevant for numerical applications where one often restricts the problem to a bounded domain. Furthermore, in this case we obtain stronger stability results as a direct corollary of Theorem 4.1.

**Theorem 4.2.** *Let  $\Omega$  be a convex and compact subset of  $\mathbb{R}^d$ . Given two Radon measures  $\mu, \nu \in \mathcal{P}(\Omega)$  whose barycenters lie at the origin and  $\dim \text{supp } \mu = \dim \text{supp } \nu = d$ , let  $\varphi_\mu, \varphi_\nu$  be maximizers of  $\mathcal{J}_{\mu, \Omega}, \mathcal{J}_{\nu, \Omega}$  respectively. Then, there exists a universal constant  $C_{d, \Omega}$  depending only on the dimension and on the diameter of  $\Omega$  such that*

$$(4.4) \quad \inf_{x_0 \in \mathbb{R}^d} \left\| \varrho_{\varphi_\mu^*}(\cdot) - \varrho_{\varphi_\nu^*}(\cdot + x_0) \right\|_{L^1(\mathbb{R}^d)} \leq C_{d, \Omega} W_1(\mu, \nu)^{1/2}$$

and

$$(4.5) \quad \inf_{a \in \mathbb{R}^d, b \in \mathbb{R}} \|(\varphi_\mu - \varphi_\nu) - (a \cdot x + b)\|_{L^1(\mu_\lambda)} \leq C_{d, \Omega} W_1(\mu, \nu)^{1/2},$$

where  $\mu_\lambda = \mu_{\psi_\lambda}$  is the moment measure of  $\psi_\lambda = ((1 - \lambda)\varphi_\mu + \lambda\varphi_\nu)^*$  for some  $\lambda \in (0, 1)$ .

*Proof.* Consider  $\mu, \nu \in \mathcal{P}(\Omega)$ , and consider two convex functions  $\varphi_\mu, \varphi_\nu$  such that

$$\varphi_\mu \in \operatorname{argmax} \mathcal{J}_\mu, \quad \varphi_\nu \in \operatorname{argmax} \mathcal{J}_\nu.$$

Since  $\varphi_\mu, \varphi_\nu$  are optimal Kantorovitch potentials for the  $L^2$ -optimal transportation problem in a compact domain, they are both Lipschitz with constant  $\text{diam } \Omega$ . Then by optimality of  $\varphi_\mu$ , for

the maximization of  $\mathcal{J}_\mu$ , and concavity of  $\mathcal{J}_\mu$  we have that

$$\begin{aligned} 0 \leq \mathcal{J}_\mu(\varphi_\mu) - \mathcal{J}_\mu(\varphi_\nu) &\leq \langle \nabla \mathcal{J}_\mu(\varphi_\nu), \varphi_\mu - \varphi_\nu \rangle = \langle \nu - \mu, \varphi_\mu - \varphi_\nu \rangle \\ &\leq \text{Lip}(\varphi_\mu - \varphi_\nu) W_1(\mu, \nu), \end{aligned}$$

which can then be easily combined with Lemma 4.1 to obtain bounds on the distance between the sets  $\text{argmin } \mathcal{E}_\mu$ ,  $\text{argmin } \mathcal{E}_\nu$ , and between  $\text{argmax } \mathcal{J}_\mu$ ,  $\text{argmax } \mathcal{J}_\nu$ .  $\square$

**Remark 4.3.** A similar argument can be employed to estimate the distance between the sets  $\text{argmin } \mathcal{E}_{\mu,\alpha}$ ,  $\text{argmin } \mathcal{E}_{\nu,\alpha}$  in the  $L^1(\mathbb{R}^d)$  topology. Indeed, for all  $\alpha \geq 0$  the functional  $\mathcal{J}_{\mu,\alpha}$  remains concave, with an application of the classical Prékopa-Leindler inequality. Hence taking  $\varphi_{\mu,\alpha}, \varphi_{\nu,\alpha}$  such that

$$\varphi_{\mu,\alpha} \in \text{argmax } \mathcal{J}_{\mu,\alpha}, \quad \varphi_{\nu,\alpha} \in \text{argmax } \mathcal{J}_{\nu,\alpha},$$

these are also Lipschitz continuous, and we have the same bound

$$\begin{aligned} 0 \leq \mathcal{J}_{\mu,\alpha}(\varphi_{\mu,\alpha}) - \mathcal{J}_{\nu,\alpha}(\varphi_{\nu,\alpha}) &\leq \langle \nabla \mathcal{J}_{\mu,\alpha}(\varphi_{\nu,\alpha}), \varphi_\mu - \varphi_\nu \rangle = \langle \nu - \mu, \varphi_{\mu,\alpha} - \varphi_{\nu,\alpha} \rangle \\ &\leq \text{diam } \Omega W_1(\mu, \nu). \end{aligned}$$

Which can also be used as in Lemma 4.1 to bound the distance  $\text{argmin } \mathcal{E}_{\mu,\alpha}$ ,  $\text{argmin } \mathcal{E}_{\nu,\alpha}$ , but not of  $\text{argmax } \mathcal{J}_{\mu,\alpha}$ ,  $\text{argmax } \mathcal{J}_{\nu,\alpha}$ , since for  $\alpha > 0$  the argument via quantitative stability of Brascamp-Lieb does not apply.

**4.2. Convergence rates w.r.t. regularization parameter.** We are also able to prove results on the sharp rates of convergence of regularized moment measures introduced in [DF25] to the classical ones from [CEK15].

Recall that for each  $\alpha > 0$ , the regularization present in the energies  $\mathcal{E}_{\mu,\alpha}$  and  $\mathcal{J}_{\mu,\alpha}$  induces a unique minimizer/maximizer pair  $(\varrho_\alpha, \varphi_\alpha)$  associated with the moment regularized measure representation of  $\mu$ . Therefore, it makes sense to study the convergence of these pairs as  $\alpha \rightarrow 0$  by estimating the distance of  $\varrho_\alpha$  to  $\mathcal{M}_\mu$  and of  $\varphi_\alpha$  to  $\mathcal{N}_\mu$ .

**Theorem 4.4.** *Fix a Radon measure  $\mu \in \mathcal{P}_1(\mathbb{R}^d)$  whose barycenter lies at the origin and such that  $\dim \text{supp } \mu = d$ . Then there exists a constant  $C_{d,\mu}$  depending on the ambient dimension and on  $\mu$  such that*

$$(4.6) \quad \text{dist}_{L^1(\mathbb{R}^d)}(\varrho_\alpha, \mathcal{M}_\mu) \leq C_{d,\mu} \alpha^{1/2},$$

where the Gibbs probability measures  $\varrho_\alpha \propto e^{-(\varphi_\alpha^* + \frac{\alpha}{2}|x|^2)}$ , give the regularized moment measure representation for  $\mu$ .

In addition, let  $(\varphi_\alpha)_{\alpha>0}$  be the family of unique maximizers of  $(\mathcal{J}_{\mu,\alpha})_{\alpha>0}$ . Then, for each  $\alpha$ , there exists  $\lambda \in (0, 1/2)$  such that, setting  $v \stackrel{\text{def.}}{=} \varphi - \varphi_\alpha$  and  $\varphi_t \stackrel{\text{def.}}{=} \varphi + \lambda v$ , it holds that

$$(4.7) \quad \text{dist}_{L^1(\mu_\lambda)}(\varphi_\alpha, \mathcal{N}_\mu) \leq C_{d,\mu} \alpha^{1/2},$$

where  $\mu_\lambda = \mu_{\varphi_\lambda^*}$  is the moment measure associated with the interpolation  $(\varphi_\lambda)^*$ .

**Example 4.5.** Set  $S_\lambda(x) = \lambda x$ , and consider the scalings

$$\mu_\lambda = (S_\lambda)_\sharp \mu, \text{ and } \varrho_\lambda = e^{\psi_\lambda} \text{ with } (\varphi_\lambda)_\sharp \varrho_\lambda = \mu.$$

Then one can check that  $\varphi_\lambda(\cdot) = \lambda^d \varphi(\lambda \cdot)$ , where  $\varphi$  is the moment map. As a result we have that

$$M_2(\mu_\lambda) = \lambda^2 M_2(\mu), \quad M_2(\varrho_\lambda) = \lambda^{-2} M_2(\varrho),$$

so it is perfectly possible that  $M_2(\varrho)$  explodes while  $M_2(\mu)$  remains bounded.

For the purposes of Theorem 4.4, the only important fact is that  $M_2(\varrho_\mu)$  is finite, which is always the case for log-concave measures, but in the next section we identify conditions on  $\mu$  which ensure that  $M_2(\varrho_\mu)$  can be controlled uniformly.

As we will see in the proof below, the constant  $C_{d,\mu} = C_d M_2(\varrho)$ , where  $\varrho = \varrho_\mu$ , the Gibbs measure associated with the moment measure representation of  $\mu$ . Therefore, the constant

depends on  $\mu$  only via the second moment of  $\varrho_\mu$ , which is finite since  $\varrho_\mu$  is log-concave, but cannot be controlled in general only by bounding the second moment of  $\mu$  as the following example shows.

*Proof of Theorem 4.4.* The goal is to compare the  $\mathcal{J}_\mu$  energies of  $\varphi_\alpha$ , the unique maximizer of  $\mathcal{J}_{\mu,\alpha}$  for  $\alpha > 0$ , and the maximizer  $\varphi$  of  $\mathcal{J}_\mu$  such that  $\varrho = e^{-\varphi^*}$  is a centered probability measure.

From Theorem 2.4, we know that the unique minimizer of  $\mathcal{E}_{\mu,\alpha}$  can be obtained from  $\varphi_\alpha$  as

$$\varrho_\alpha \propto e^{-(\varphi_\alpha^* + \frac{\alpha}{2}|x|^2)},$$

and similarly for the case  $\alpha = 0$ , we can map  $\varphi$  into a minimizer of  $\mathcal{E}_\mu$  defined as the Gibbs probability measure

$$\varrho \propto e^{-\varphi^*}.$$

To compare the  $\mathcal{J}_\mu$  energies, we start by recalling the regularized energy  $\mathcal{J}_{\mu,\alpha}$  and noticing the following trivial bound

$$\begin{aligned} \mathcal{J}_{\mu,\alpha}(\varphi_\alpha) &= \log \int_{\mathbb{R}^d} e^{-(\varphi_\alpha^* + \frac{\alpha}{2}|x|^2)} dx - \int_{\mathbb{R}^d} \varphi_\alpha d\mu \\ &\leq \log \int_{\mathbb{R}^d} e^{-\varphi_\alpha^*} dx - \int_{\mathbb{R}^d} \varphi_\alpha d\mu = \mathcal{J}_\mu(\varphi_\alpha). \end{aligned}$$

Therefore we have that

$$\begin{aligned} \mathcal{J}_\mu(\varphi) - \mathcal{J}_\mu(\varphi_\alpha) &\leq \mathcal{J}_\mu(\varphi) - \mathcal{J}_{\mu,\alpha}(\varphi_\alpha) = \mathcal{E}_{\mu,\alpha}(\varrho_\alpha) - \mathcal{E}_\mu(\varrho) \\ &\leq \mathcal{E}_{\mu,\alpha}(\varrho) - \mathcal{E}_\mu(\varrho) = \alpha M_2(\varrho) = C_\mu \alpha, \end{aligned}$$

where the first equality comes from the strong duality formula between the minimization of  $\mathcal{E}_{\mu,\alpha}$  and the maximization of  $\mathcal{J}_{\mu,\alpha}$  proved in Theorem 2.4. Since  $\varrho$  is a log-concave measure, its second moment is finite and hence a constant  $C_\mu$  depending only on  $\mu$ .

Combining this estimate with Lemma 4.1, the result follows.  $\square$

**4.3. Stability over  $\mathbb{R}^d$ .** The stability result in a compact setting proven in Section 4.1 covers most cases of applications, since for computational purposes one usually needs to truncate the measures to a compact domain. Although the restriction of  $\mu$  to a compact domain is not problematic due to many interesting cases already being in this setting, the Gibbs measure  $\varrho_\mu$  is in general supported over the whole  $\mathbb{R}^d$ . Therefore, it is interesting to obtain stability results also for moment measure representations over the whole  $\mathbb{R}^d$ .

As mentionned above, this was achieved for regularized moment measures in [DF25], using the strong geodesic convexity of the functional  $\mathcal{E}_{\mu,\alpha}$ . The convergence results from Section 4.2 depending on  $M_2(\varrho_\mu)$  give a hint on why the stability for regularized moment measures appears to be a more direct issue, as in this case one can expect to control the second moment of the Gibbs measure  $\varrho_{\mu,\alpha}$  independently of  $\mu$ , thanks to the regularization term. Without such a uniform control, we cannot know if the constant multiplying any quantitative stability inequality will remain bounded. Example 4.5 above shows that we cannot in general bound  $M_2(\varrho_\mu)$  in terms of  $M_2(\mu)$ .

In this section, we identify two very natural classes of measures  $\mu$ , for which  $M_2(\varrho_\mu)$  remains uniformly bounded. We can then leverage the stability proof via geodesic convexity from [DM23] and the convergence rates in Theorem 4.4 to obtain quantitative stability results for moment measure representations over  $\mathbb{R}^d$  in terms of the  $W_2$ -distance.

The first approach is to impose a lower bound on the quantity

$$(4.8) \quad \Theta(\mu) \stackrel{\text{def.}}{=} \inf_{\theta \in \mathbb{S}^{d-1}} \int_{\mathbb{R}^d} |y \cdot \theta| d\mu(y),$$

that measures the minimal spread of the measure  $\mu$  in all directions, being therefore very natural for the study of moment measures. We shall also require that  $\mu$  is supported on a convex and

compact set  $\Omega$ , so that we can control the Lipschitz constant of the potential  $\varphi_\mu^*$ . Therefore, we consider the class of measures defined as

$$(4.9) \quad \mathcal{K}_\vartheta \stackrel{\text{def.}}{=} \left\{ \mu \in \mathscr{P}(\Omega) : \begin{array}{l} \mu \text{ is centered,} \\ \Theta(\mu) \geq \vartheta \end{array} \right\}.$$

The fact that  $\mu \in \mathscr{P}(\Omega)$  implies that the potential  $\varphi_\mu^*$  is Lipschitz continuous with constant  $L = \text{diam}(\Omega)$ . This is not sufficient to control  $M_2(\varrho_\mu)$ , but enforcing the lower bound on  $\Theta(\mu)$  we have

$$(4.10) \quad \Theta(\mu) = \inf_{\theta \in \mathbb{S}^{d-1}} \int_{\mathbb{R}^d} |\nabla \varphi_\mu^*(x) \cdot \theta| e^{-\varphi_\mu^*(x)} dx \geq \vartheta.$$

Combining this with the Lipschitz continuity of  $\varphi_\mu^*$ , we can employ the following growth estimate for convex functions due to Klartag [Kla14].

**Lemma 4.6.** [Kla14, Lemma 2] *Suppose that  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $L$ -Lipschitz, convex function such that  $\varrho \propto e^{-\psi}$  has barycenter at the origin. If there exists  $c > 0$  such that*

$$\inf_{\theta \in \mathbb{S}^{d-1}} \int_{\mathbb{R}^d} |\nabla \psi(x) \cdot \theta| e^{-\psi(x)} dx \geq c,$$

then there exists  $\ell, \beta_1, \beta_2$  depending only on  $d, L, c$  such that

$$\ell|x| - \beta_1 \leq \psi(x) \leq L|x| + \beta_2, \quad \text{for all } x \in \mathbb{R}^d.$$

In particular, the second moment of  $\varrho$  is bounded by a constant depending only on  $d, L, c$ .

We can now state our quantitative stability result for measures in  $\mathcal{K}_\vartheta$ .

**Theorem 4.7.** *Given  $\Omega$  a convex and compact subset of  $\mathbb{R}^d$  and  $\vartheta > 0$ , then for every  $p > 2$  there exists a constant  $C$  such that for all  $\mu, \nu \in \mathcal{K}_\vartheta$*

$$(4.11) \quad \text{dist}_{W_2}(\mathcal{M}_\mu, \mathcal{M}_\nu) \leq CW_2(\mu, \nu)^{\frac{p-2}{6p-4}},$$

where  $C$  depends only on the  $d$ ,  $\text{diam} \Omega$ ,  $\vartheta$ , and grows linearly in  $p$ .

*Proof of Theorem 4.7.* The proof consists of combining the convergence rates of regularized moment measures from Theorem 4.4 with strong geodesic convexity argument from [DM23] and the uniform bounds on the second moment that we obtain for  $\varrho_\mu$  whenever  $\mu \in \mathcal{K}_\vartheta$  through Lemma 4.6.

First, notice that for all  $\mu \in \mathcal{K}_\vartheta$  and  $\alpha \geq 0$ , letting  $\varrho_{\mu,\alpha}$  denote the Gibbs measure associated with the (regularized if  $\alpha > 0$ ) moment measure representation of  $\mu$ , we have that  $M_2(\varrho_{\mu,\alpha}) \leq C_{\Omega,\vartheta,d}$ , where the constant depends only on  $\Omega, c$  and  $d$ . This is a direct consequence of Lemma 4.6, since the potentials  $\varphi_{\mu,\alpha}^*$  are all  $\text{diam}(\Omega)$ -Lipschitz continuous.

For the reader's convenience, we detail the part of [DF25]'s strong geodesic convexity argument that is relevant to us. Given  $\mu, \nu \in \mathcal{K}_\vartheta$  and some  $\alpha > 0$ , let  $(\varrho_{\alpha,t})_{t \in [0,1]}$  be the constant speed geodesic with endpoints

$$\varrho_{\alpha,0} = \varrho_{\mu,\alpha}, \quad \varrho_{\alpha,1} = \varrho_{\nu,\alpha}.$$

Since both  $\mathcal{E}_{\mu,\alpha}$  and  $\mathcal{E}_{\nu,\alpha}$  are  $\alpha$ -strongly geodesically convex functionals over  $\mathscr{P}_2(\mathbb{R}^d)$ , we have that

$$\mathcal{E}_{\mu,\alpha}(\varrho_{\mu,\alpha}) \leq \mathcal{E}_{\mu,\alpha}(\varrho_{\alpha,1/2}) \leq \frac{1}{2} \mathcal{E}_{\mu,\alpha}(\varrho_{\mu,\alpha}) + \frac{1}{2} \mathcal{E}_{\mu,\alpha}(\varrho_{\nu,\alpha}) - \frac{\alpha}{8} W_2^2(\varrho_{\mu,\alpha}, \varrho_{\nu,\alpha}),$$

and similarly for  $\mathcal{E}_{\nu,\alpha}$ . Rearranging we get

$$\frac{\alpha}{4} W_2^2(\varrho_{\mu,\alpha}, \varrho_{\nu,\alpha}) \leq \mathcal{E}_{\mu,\alpha}(\varrho_{\nu,\alpha}) - \mathcal{E}_{\mu,\alpha}(\varrho_{\mu,\alpha}),$$

$$\frac{\alpha}{4} W_2^2(\varrho_{\mu,\alpha}, \varrho_{\nu,\alpha}) \leq \mathcal{E}_{\nu,\alpha}(\varrho_{\mu,\alpha}) - \mathcal{E}_{\nu,\alpha}(\varrho_{\nu,\alpha}).$$

Adding these inequalities, and recalling that  $\mathcal{T}(\mu_0, \mu_1) = -W_2^2(\mu_0, \mu_1) + M_2(\mu_0) + M_2(\mu_1)$  and  $W_2^2(\mu_0, \mu_1) \leq 2M_2(\mu_0) + 2M_2(\mu_1)$ , we obtain

$$\begin{aligned} \frac{\alpha}{4} W_2^2(\varrho_{\mu,\alpha}, \varrho_{\nu,\alpha}) &\leq \mathcal{T}(\varrho_{\nu,\alpha}, \mu) - \mathcal{T}(\varrho_{\nu,\alpha}, \nu) + \mathcal{T}(\varrho_{\mu,\alpha}, \nu) - \mathcal{T}(\varrho_{\mu,\alpha}, \mu) \\ &= (W_2(\varrho_{\nu,\alpha}, \mu) + W_2(\varrho_{\nu,\alpha}, \nu))(W_2(\varrho_{\nu,\alpha}, \mu) - W_2(\varrho_{\nu,\alpha}, \nu)) \\ &\quad + (W_2(\varrho_{\mu,\alpha}, \nu) + W_2(\varrho_{\mu,\alpha}, \mu))(W_2(\varrho_{\mu,\alpha}, \nu) - W_2(\varrho_{\mu,\alpha}, \mu)) \\ &\leq (W_2(\varrho_{\nu,\alpha}, \mu) + W_2(\varrho_{\nu,\alpha}, \nu) + W_2(\varrho_{\mu,\alpha}, \nu) + W_2(\varrho_{\mu,\alpha}, \mu))W_2(\mu, \nu) \\ &\leq C_{\Omega, \vartheta, d} W_2(\mu, \nu), \end{aligned}$$

where we have used that for any  $\mu, \nu \in \mathcal{K}_\vartheta$ , all second moments above are uniformly bounded by a constant depending only on  $\Omega, \vartheta$  and  $d$ . In particular, this gives an explicit behaviour on  $\alpha$  for the bound proven in [DF25, Theorem 3.1] for the particular case of  $\mu, \nu \in \mathcal{K}_\vartheta$ , i.e.

$$(4.12) \quad W_2(\varrho_{\mu,\alpha}, \varrho_{\nu,\alpha}) \leq \alpha^{-1/2} C_{\Omega, \vartheta, d} W_2(\mu, \nu)^{1/2}.$$

To finish the proof we need to exploit the explicit rates of convergence of regularized moment measures from Theorem 4.4. The issue is that these convergence rates concern the  $L^1$  distance between  $\varrho_{\mu,\alpha}$  and  $\mathcal{M}_\mu$ . But for absolutely continuous measures with bounded  $p$ -moments, the  $L^1$  norm can be used to control Wasserstein distances. There exists a constant  $C_{p,q}$  such that

$$(4.13) \quad W_q(\varrho_0, \varrho_1) \leq C_{p,q} (M_p(\varrho_0) + M_p(\varrho_1))^{1/p} \|\varrho_0 - \varrho_1\|_{L^1}^{1/q-1/p},$$

see Lemma 2.1 in Section 2.

As a result, combining (4.13), with the stability of regularized moment measures from (4.12), and the explicit rates of convergence from 4.4, we let  $\varrho_\mu = \varrho_\mu(\cdot + x_\alpha) \in \mathcal{M}_\mu$  be a Gibbs measure associated with the moment measure representation of  $\mu$  such that

$$\|\varrho_{\mu,\alpha} - \varrho_\mu\|_{L^1(\mathbb{R}^d)} \leq 2\text{dist}_{L^1}(\varrho_{\mu,\alpha}, \mathcal{M}_\mu) \leq C_{\Omega, d, \vartheta} \alpha^{1/2},$$

and similarly for  $\varrho_\nu$ . Then, for every  $\alpha > 0$ , we have

$$\begin{aligned} \text{dist}_{W_2}(\mathcal{M}_\mu, \mathcal{M}_\nu) &\leq W_2(\varrho_\mu, \varrho_\nu) \leq W_2(\varrho_\mu, \varrho_{\mu,\alpha}) + W_2(\varrho_{\mu,\alpha}, \varrho_{\nu,\alpha}) + W_2(\varrho_{\nu,\alpha}, \varrho_\nu) \\ &\leq C \left( \alpha^{1/4-1/2p} + \alpha^{-1/2} W_2(\mu, \nu)^{1/2} \right), \end{aligned}$$

where the constant  $C$  above depends on  $\Omega, \vartheta, d, p$ , more specifically it depends on  $M_p(\varrho_\mu)^{1/p}$ , which behaves linearly in  $p$ , so that the constant explodes as  $p \rightarrow \infty$ .

But for every finite  $p > 2$ , we can optimize the above estimate in  $\alpha$  to obtain the best possible bound, which is achieved for

$$\alpha = W_2(\mu, \nu)^{\frac{2p}{3p-2}},$$

which gives

$$\text{dist}_{W_2}(\mathcal{M}_\mu, \mathcal{M}_\nu) \leq C W_2(\mu, \nu)^{\frac{p-2}{6p-4}},$$

for all  $p > 2$ , with a constant  $C$  depending on  $\Omega, \vartheta, d, p$ .  $\square$

**Proposition 4.8.** *Inequality (4.11) never holds with a constant independent of  $\vartheta$ .*

*Proof.* Suppose that there was a constant  $C$ , independent of  $\vartheta$ , such that (4.11) holds for all moment measures  $\mu, \nu$ . Then consider the one-dimensional measure over the segment  $L = [0, 1]e_1 \subset \mathbb{R}^d$

$$\mu = \mathcal{H}^1 \llcorner L,$$

and a sequence of normalized mollifications  $(\mu_\varepsilon)_{\varepsilon>0}$ , so that  $W_2(\mu, \mu_\varepsilon) \rightarrow 0$ .

If a inequality such as (4.11) held with a constant independent of  $\vartheta$ , then we would have that  $\varrho_\varepsilon = e^{-\psi_\varepsilon}$  would be Cauchy in  $W_2$  as  $\varepsilon \rightarrow 0$ , and hence would converge to some limit measure  $\varrho$ . Since all  $\psi_\varepsilon$  as  $L$ -Lipschitz, we can extract a locally uniformly converging subsequence, so that  $\psi_\varepsilon \rightarrow \psi$  locally uniformly, and hence  $\varrho = e^{-\psi}$ .

In addition,  $\gamma_\varepsilon \stackrel{\text{def.}}{=} (\text{id}, \nabla \psi_\varepsilon)_\# \varrho_\varepsilon$  is the optimal transport plan between  $\varrho_\varepsilon$  and  $\mu_\varepsilon$ , and since  $W_2(\mu_\varepsilon, \mu) \rightarrow 0$  we have that  $\gamma_\varepsilon$  converges weakly to some limit plan  $\gamma$ . By stability of optimal transport plans,  $\gamma$  is an optimal transport plan between  $\varrho$  and  $\mu$  and must be given by  $(\text{id}, \nabla \psi)_\# \varrho$ . But then  $\psi$  is  $L$ -Lipschitz and a moment measure representation for  $\mu$ , which is impossible since  $\mu$  is supported on a hyperplane.  $\square$

**Remark 4.9.** The exponent  $\frac{p-2}{6p-4}$  in Theorem 4.7 can be made arbitrarily close to  $1/6$  by taking  $p$  large enough, at the cost of getting a worse constant  $C$ , which behaves linearly at  $p$ . It is somewhat natural to expect a worse exponent than  $1/2$  obtained in [DF25] for regularized moment measures, since the regularization term improves the second moment of solutions. However, it is not clear whether the limiting exponent  $1/6$  is optimal or not.

The second class of measures for which we can derive quantitative stability results is obtained by bounding the curvature they induce. Given  $\Lambda > 0$  we define

$$(4.14) \quad \mathcal{K}_\Lambda \stackrel{\text{def.}}{=} \left\{ \mu \in \mathscr{P}_{\text{ac}}(\mathbb{R}^d) : \mu \propto e^{-V}, \text{ where } D^2V \leq \Lambda \text{id} \right\}.$$

In (4.14), the condition  $D^2V \leq \Lambda \text{id}$  is understood only on  $\text{dom } V$  and we do not make any convexity assumption. In this case that  $\mu$  has bounded curvature, we can expect that the Gibbs measure  $\varrho_\mu$  has a stronger geometry than the log-concavity of the orginal measure. Indeed we shall show that the optimal potential  $\varphi_\mu^*$  that yields the moment measure representation of  $\mu$  is  $\Lambda^{-1/3}$ -strongly convex.

This improved strong-convexity result will be obtained by bootstrapping the supplementary regularity enjoyed by the regularized moment measure representation in conjunction with *Caffarelli's contraction theorem*, a global Lipschitz regularity result for the optimal transport between log-concave measures. See for instance the original works of Caffarelli [Caf92, Caf00], or [GS25] for a modern approach via entropic regularization and the references therein. For the reader's convenience, we recall that it is stated as follows:

**Theorem 4.10** ([Caf92, Caf00]Caffarelli's contraction theorem). *Let  $\mu, \nu \in \mathscr{P}_{\text{ac}}(\mathbb{R}^d)$  be such that  $\text{dom } \mu = \mathbb{R}^d$ ,  $\text{dom } \nu$  is convex, and satisfying*

$$(4.15) \quad \mu \propto e^{-V}, \quad \nu \propto e^{-W}, \quad \text{where } D^2V \leq \Lambda \text{id} \text{ and } D^2W \geq \lambda \text{id},$$

*where  $\Lambda, \lambda > 0$ . Then the  $L^2$ -optimal transport map  $T_\# \mu = \nu$  is Lipschitz continuous with constant  $\sqrt{\Lambda/\lambda}$ .*

With this result we prove the following regularity result for moment measures associated with elements of  $\mathcal{K}_\Lambda$ .

**Theorem 4.11.** *For any  $\mu \in \mathcal{K}_\Lambda$ , let  $\varphi_\mu^*$  be a potential inducing the moment measure representation  $\mu = (\nabla \varphi_\mu^*)_\# e^{-\varphi_\mu^*}$ . Then  $\varphi_\mu^*$  is  $\Lambda^{-1/3}$ -convex.*

*Proof.* Fix some  $\alpha > 0$  and consider the regularized moment measure representation of  $\mu$  as

$$\mu = (\nabla \varphi_\alpha^*)_\# e^{-(\varphi_\alpha^* + \frac{\alpha}{2}|x|^2)} = (\nabla \varphi_\alpha^*)_\# \varrho_\alpha,$$

so that  $\varrho_\alpha \propto e^{-V_\alpha}$  where  $D^2V_\alpha \geq \alpha \text{id}$ . In addition, from Brenier's Theorem  $\nabla \varphi_\alpha$  is the  $L^2$  optimal transportation map from  $\mu$  to  $\varrho_\alpha$ .

From Caffarelli's contraction theorem we have that

$$\text{Lip}(\nabla \varphi_\alpha) \leq \sqrt{\frac{\Lambda}{\alpha}},$$

which in turns implies further regularity for the Legendre transform  $\varphi_\alpha^*$ , being therefore  $\sqrt{\alpha/\Lambda}$ -strongly convex. This implies a better bound for strong convexity of  $V_\alpha$  and can be exploited

once again in Caffarelli's contraction theorem, implying that

$$\text{Lip}(\nabla \varphi_\alpha) \leq \sqrt{\frac{\Lambda}{\sqrt{\alpha/\Lambda}}} = \frac{\Lambda^{1/4}}{\alpha^{1/4}}.$$

This argument can be bootstrapped indefinitely, and one can check by induction that at the  $k$ -th time it yields

$$\text{Lip}(\nabla \varphi_\alpha) \leq \frac{\Lambda^{S_k}}{\alpha^{1/2^k}} \text{ and } \varphi_\alpha^* \text{ is } \frac{\alpha^{1/2^k}}{\Lambda^{S_k}}\text{-strongly convex,}$$

with  $S_k = \sum_{i=1}^k \frac{(-1)^{i+1}}{2^i}$ . Passing to the limit as  $k \rightarrow \infty$ , we obtain that for all  $\alpha > 0$

$$\text{Lip}(\nabla \varphi_\alpha) \leq \Lambda^{1/3} \text{ and } \varphi_\alpha^* \text{ is } \Lambda^{-1/3}\text{-strongly convex.}$$

In particular, this implies that  $\nabla \varphi_\alpha$  converges locally uniformly to  $\nabla \varphi_\mu$  a maximizer of  $\mathcal{J}_\mu$ , so that  $\varphi$  is also  $\Lambda^{1/3}$ -Lipschitz, so that the optimal potential  $\varphi^*$  is  $\Lambda^{-1/3}$ -strongly convex, as we wanted to prove.  $\square$

Besides the relevance to obtaining another quantitative stability result for moment measures in the class  $\mathcal{K}_\Lambda$ , see Theorem 4.13 below, this result is particularly interesting for applications in sampling. Whenever  $\mu \propto e^{-V}$  is log-concave there are efficient sampling algorithms based on the Langevin dynamics associated with the potential  $V$  [Che23, CNWR25, DCWY19]. In the case that  $V$  is strongly convex, its associate Gibbs measure satisfies a log-Sobolev inequality as we've learned from the Bakry-Émery technique [BGL13] and as a result implies exponential convergence rates for the associated Langevin dynamics [MV00]. This can then be exploited in sampling algorithms, see [CKSV] for a recent application or the recent monographs [Che23, CNWR25]. With only the upper bound  $D^2V \leq \Lambda$ , the rate of convergence of the Langevin dynamics, and hence of sampling algorithms, can be only polynomial if any convergence guarantee is even available. However, if the moment measure representation for  $\mu$  has been computed, we know from Theorem 4.11 that its potential  $\varphi_\mu^*$  is strongly convex, so we can use a Langevin-based sampling algorithm with exponential convergence rates to obtain a sample of the (strongly) log-concave measure  $\varrho_\mu$  and map this sample into a sample of  $\mu$  directly.

In Lemma 4.6 we have exploited the lower bound on  $\Theta(\mu)$  aligned with the Lipschitz continuity of the moment map induced by  $\mu$  to bound the moments  $M_p(\varrho_\mu)$ . This can even more easily be achieved for  $\mu \in \mathcal{K}_\Lambda$  by exploiting the strong convexity result from Theorem 4.11 and the following bound on the  $p$ -moments.

**Lemma 4.12.** *Let  $\psi$  be a centered,  $\lambda$ -strongly convex function. Then for every  $p > 2$  the  $p$ -moment of its associated Gibbs measure can be bounded as*

$$M_p(e^{-\psi})^{1/p} = \left( \int_{\mathbb{R}^d} |x|^p e^{-\psi(x)} dx \right)^{1/p} \leq Cp,$$

where  $C$  is a constant depending only on  $\lambda$  and the dimension  $d$ .

*Proof.* Since the  $\lambda$ -strongly convex function  $\psi$  is centered, it is a known fact in convex analysis that:

$$(4.16) \quad \psi(0) \leq d + \inf \psi = d + \psi(x_*)$$

where  $x^*$  denotes its minimizer. This is proven for instance in [Fra97]. As a result of the  $\lambda$ -strong convexity of  $\psi$ , and the fact that  $\nabla \psi(x_*) = 0$  that

$$(4.17) \quad \langle \nabla \psi(x), x - x_* \rangle \geq \lambda |x - x_*|^2.$$

On the other hand, the divergence theorem gives that

$$(4.18) \quad \int_{\mathbb{R}^d} \text{div}(|x - x_*|^{p-2}(x - x_*) e^{-\psi}) dx = \int_{\mathbb{R}^d} |x - x_*|^{p-2}(x - x_*) \cdot \nabla \psi(x) e^{-\psi} dx.$$

Now define the following quantity

$$V(p) \stackrel{\text{def.}}{=} \int_{\mathbb{R}^d} |x - x_*|^p e^{-\psi(x)} dx.$$

Combining identities (4.17), (4.17), and developing the divergence, we obtain that

$$V(p) \leq \frac{(d+p-2)}{\lambda} V(p-2).$$

Noticing that  $V(0) = 1$ , by induction we obtain that for all  $k \in \mathbb{N}$

$$V(2k) \leq \lambda^{-k} \prod_{i=0}^{k-1} (d+2i) = \lambda^{-k} 2^k \frac{\Gamma(\frac{d}{2} + k)}{\Gamma(\frac{d}{2})},$$

where  $\Gamma$  corresponds to the gamma function.

Using Hölder's inequality, in the form  $\|f\|_p^p \leq \|f\|_{p_0}^{(1-t)p_0} \|f\|_{p_1}^{tp_1}$ , for a convex combination  $p = (1-t)p_0 + tp_1$ , we can interpolate the above inequality for any  $p > 2$ , by choosing  $p_0 = 2k$  and  $p_1 = 2(k+1)$  for a convenient  $k$  such that  $p \in (2k, 2(k+1))$ . Due to Stirling's formula, this yields an estimate such that

$$\left( \int_{\mathbb{R}^d} |x - x_*|^p e^{-\psi(x)} dx \right)^{1/p} \leq Cp,$$

for a constant  $C$  depending on  $d$  and  $\lambda$ .

To conclude, we need only to bound  $|x_*|$ , this can be done with (4.16) and the strongl convexity. Indeed, using once again that  $\nabla\psi(x_*) = 0$ , we have

$$\lambda|x - x_*|^2 \leq \psi(0) - \psi(x_*) \leq d.$$

Hence, combing all these estimates we get that

$$M_p(e^{-\psi})^{1/p} \leq \left( \int_{\mathbb{R}^d} |x - x_*|^p e^{-\psi(x)} dx \right)^{1/p} + |x_*| \leq Cp + \sqrt{\frac{d}{\lambda}},$$

and the result follows.  $\square$

Combining this and the reasoning of Theorem 4.7 we can obtain a stability result for elements of  $\mathcal{K}_\Lambda$  with a completely analogous proof.

**Theorem 4.13.** *Given  $\Lambda > 0$ , then for each  $p > 2$  there exists a constant  $C$  such that, for all  $\mu, \nu \in \mathcal{K}_\Lambda$  it holds thaths*

$$(4.19) \quad \text{dist}_{W_2}(\mathcal{M}_\mu, \mathcal{M}_\nu) \leq CW_2(\mu, \nu)^{\frac{p-2}{6p-4}},$$

where  $C$  depends only on the  $d$ ,  $\Lambda$ , and grows linearly in  $p$ .

## REFERENCES

- [ABS21] Luigi Ambrosio, Elia Brué, and Daniele Semola. *Lectures on optimal transport*, volume 130. Springer, 1 edition, 2021.
- [AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [BB13] Robert J Berman and Bo Berndtsson. Real monge-ampère equations and kähler-ricci solitons on toric log fano varieties. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 22, pages 649–711, 2013.
- [BDPV15] Lorenzo Brasco, Guido De Philippis, and Bozhidar Velichkov. Faber–krahn inequalities in sharp quantitative form. 2015.
- [BÉ06] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84: Proceedings*, pages 177–206. Springer, 2006.
- [BFR23] Károly J Böröczky, Alessio Figalli, and Joao PG Ramos. A quantitative stability result for the prékopa–leindler inequality for arbitrary measurable functions. *Annales de l’Institut Henri Poincaré C*, 41(3):565–614, 2023.

- [BFR25] Károly J Böröczky, Alessio Figalli, and Joao PG Ramos. Isoperimetric inequalities, brunn-minkowski theory and minkowski type monge-ampère equations on the sphere. 2025.
- [BGG18] François Bolley, Ivan Gentil, and Arnaud Guillin. Dimensional improvements of the logarithmic sobolev, talagrand and brascamp–lieb inequalities. *The Annals of Probability*, 46(1):261–301, 2018.
- [BGL13] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348. Springer Science & Business Media, 2013.
- [BJR] Guillaume Bonnet, Yuxiang Ji, and Yanir A. Rubinstein. Semidiscrete approximation of moment measure potentials and kähler–ricci solitons. *In preparation*.
- [BK25] Zoltán M Balogh and Alexandru Kristály. Sharp stability in hypercontractivity estimates and logarithmic sobolev inequalities. *arXiv preprint arXiv:2508.21552*, 2025.
- [BL76] Herm Jan Brascamp and Elliott H Lieb. On extensions of the brunn-minkowski and prékopa-leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of functional analysis*, 22(4):366–389, 1976.
- [BL00] Sergey G Bobkov and Michel Ledoux. From brunn-minkowski to brascamp-lieb and to logarithmic sobolev inequalities. *Geometric and Functional Analysis*, 10(5):1028–1052, 2000.
- [Caf92] Luis A Caffarelli. The regularity of mappings with a convex potential. *Journal of the American Mathematical Society*, 5(1):99–104, 1992.
- [Caf00] Luis A Caffarelli. Monotonicity properties of optimal transportation and the fkg and related inequalities. *Communications in Mathematical Physics*, 214(3):547–563, 2000.
- [CE17] Dario Cordero-Erausquin. Transport inequalities for log-concave measures, quantitative forms, and applications. *Canadian Journal of Mathematics*, 69(3):481–501, 2017.
- [CEK15] Dario Cordero-Erausquin and Bo’az Klartag. Moment measures. *Journal of Functional Analysis*, 268(12):3834–3866, 2015.
- [Che23] Sinho Chewi. *An optimization perspective on log-concave sampling and beyond*. Massachusetts Institute of Technology, 2023.
- [CKSV] Omar Chehab, Anna Korba, Austin J Stromme, and Adrien Vacher. Provable convergence and limitations of geometric tempering for langevin dynamics. In *The Thirteenth International Conference on Learning Representations*.
- [CNWR25] Sinho Chewi, Jonathan Niles-Weed, and Philippe Rigollet. Statistical optimal transport: École d’Été de probabilités de saint-flour xlix – 2019. 2025.
- [DCWY19] Raaz Dwivedi, Yuansi Chen, Martin J Wainwright, and Bin Yu. Log-concave sampling: Metropolis-hastings algorithms are fast. *Journal of Machine Learning Research*, 20(183):1–42, 2019.
- [DEF<sup>+</sup>25] Jean Dolbeault, Maria J Esteban, Alessio Figalli, Rupert L Frank, and Michael Loss. Sharp stability for sobolev and log-sobolev inequalities, with optimal dimensional dependence. *Camb. J. Math.*, 13(2):359–430, 2025.
- [DF25] Alex Delalande and Sara Farinelli. Regularized moment measures. *arXiv preprint arXiv:2506.13218*, 2025.
- [DM23] Alex Delalande and Quentin Merigot. Quantitative stability of optimal transport maps under variations of the target measure. *Duke Mathematical Journal*, 172(17):3321–3357, 2023.
- [DT16] Jean Dolbeault and Giuseppe Toscani. Stability results for logarithmic sobolev and gagliardo-nirenberg inequalities. *International Mathematics Research Notices*, 2016(2):473–498, 2016.
- [EG15] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. CRC press, 2015.
- [FJ17] Alessio Figalli and David Jerison. Quantitative stability for the brunn–minkowski inequality. *Advances in Mathematics*, 314:1–47, 2017.
- [FMP08] Nicola Fusco, Francesco Maggi, and Aldo Pratelli. The sharp quantitative isoperimetric inequality. *Annals of mathematics*, pages 941–980, 2008.
- [FMP10] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. *Inventiones Mathematicae*, 182(1):167–211, 2010.
- [FR24] Alessio Figalli and Joao PG Ramos. Improved stability versions of the pr\’ekopa-leindler inequality. *arXiv preprint arXiv:2410.01122*, 2024.
- [Fra97] Matthieu Fradelizi. Sections of convex bodies through their centroid. *Archiv der Mathematik*, 69(6):515–522, 1997.
- [FvHT23] Alessio Figalli, Peter van Hintum, and Marius Tiba. Sharp quantitative stability of the brunn-minkowski inequality. *arXiv preprint arXiv:2310.20643*, 2023.
- [FvHT24] Alessio Figalli, Peter van Hintum, and Marius Tiba. Sharp stability of the brunn-minkowski inequality via optimal mass transportation. *arXiv preprint arXiv:2407.10932*, 2024.
- [FvHT25] Alessio Figalli, Peter van Hintum, and Marius Tiba. Sharp quantitative stability for the pr\’ekopa-leindler and borell-brascamp-lieb inequalities. *arXiv preprint arXiv:2501.04656*, 2025.
- [GS25] Nathael Gozlan and Maxime Sylvestre. Global regularity estimates for optimal transport via entropic regularisation. *arXiv preprint arXiv:2501.11382*, 2025.

- [Har08] Gilles Hargé. Reinforcement of an inequality due to brascamp and lieb. *Journal of Functional Analysis*, 254(2):267–300, 2008.
- [KL25] Boaz Klartag and Joseph Lehec. Isoperimetric inequalities in high-dimensional convex sets. *Bulletin of the American Mathematical Society*, 62(4):575–642, 2025.
- [Kla14] Bo'az Klartag. Logarithmically-concave moment measures i. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2011–2013*, pages 231–260. Springer, 2014.
- [Lei73] L. Leindler. On a certain converse of hölder's inequality ii. *Acta Scientiarum Mathematicarum*, 34:–, 1973.
- [Let25] Cyril Letrouit. Lectures on quantitative stability of optimal transport. 2025.
- [Lew22] Mathieu Lewin. *Théorie spectrale et mécanique quantique*, volume 87. Springer Nature, 2022.
- [LLR25] Nguyen Lam, Guozhen Lu, and Andrey Russanov. Stability of gaussian poincar\'{e} inequalities and heisenberg uncertainty principle with monomial weights. *arXiv preprint arXiv:2503.17567*, 2025.
- [LM24] Cyril Letrouit and Quentin Mérigot. Gluing methods for quantitative stability of optimal transport maps. *arXiv preprint arXiv:2411.04908*, 2024.
- [LS10] Elliott H Lieb and Robert Seiringer. *The stability of matter in quantum mechanics*. Cambridge university press, 2010.
- [McC97] Robert J McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.
- [MV00] Peter A Markowich and Cédric Villani. On the trend to equilibrium for the fokker-planck equation: an interplay between physics and functional analysis. *Mat. Contemp*, 19:1–29, 2000.
- [Pré71] András Prékopa. Logarithmic concave measures with applications to stochastic programming. *Acta Scientiarum Mathematicarum*, 1971.
- [Pre73] Andras Prékopa. On logarithmic concave measures and functions. *Acta Scientiarum Mathematicarum*, 34:335–343, 1973.
- [San15] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkhäuser, NY*, 55(58-63):94, 2015.
- [San16] Filippo Santambrogio. Dealing with moment measures via entropy and optimal transport. *Journal of Functional Analysis*, 271(2):418–436, 2016.
- [VBC25] Nina Vesseron, Louis Béthune, and Marco Cuturi. Sample and map from a single convex potential: Generation using conjugate moment measures. *arXiv preprint arXiv:2503.10576*, 2025.
- [vHST21] Peter van Hintum, Hunter Spink, and Marius Tiba. Sharp stability of brunn–minkowski for homothetic regions. *Journal of the European Mathematical Society*, 24(12):4207–4223, 2021.
- [vHST23] Peter van Hintum, Hunter Spink, and Marius Tiba. Sharp quantitative stability of the planar brunn–minkowski inequality. *Journal of the European Mathematical Society*, 26(2):695–730, 2023.
- [Vil09] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

LAGRANGE MATHEMATICAL AND COMPUTATIONAL CENTER, 103 RUE DE GRENELLE, PARIS, 75007, FRANCE  
*Email address:* joao-miguel.machado@ceremade.dauphine.fr

INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA, ESTRADA DONA CASTORINA 110, HORTO - RIO DE JANEIRO, RJ - BRAZIL  
*Email address:* joao.ramos@impa.br