OPTIMAL TRANSPORT MAPS, MAJORIZATION, AND LOG-SUBHARMONIC MEASURES

GUIDO DE PHILIPPIS AND YAIR SHENFELD

Abstract. Caffarelli's contraction theorem bounds the derivative of the optimal transport map between a log-convex measure and a strongly log-concave measure. We show that an analogous phenomenon holds on the level of the trace: The trace of the derivative of the optimal transport map between a log-subharmonic measure and a strongly log-concave measure is bounded. We show that this trace bound has a number of consequences pertaining to volume-contracting transport maps, majorization and its monotonicity along Wasserstein geodesics, growth estimates of log-subharmonic functions, the Wehrl conjecture for Glauber states, and two-dimensional Coulomb gases. We also discuss volume-contraction properties for the Kim-Milman transport map.

1. Introduction

1.1. **The divergence of the Brenier map.** The study of regularity properties of transport maps between probability measures is an important topic in analysis, probability, and geometry. The most classical result in the field is due to Caffarelli [5] who established Lipschitz regularity of the Brenier map of optimal transport under log-convexity/concavity assumptions on the source and target measures. In this work we consider a low regularity notion of convexity:

Definition 1.1. A probability measure ρ on $\mathbb{R}^n$, with density $\mathrm{d}\rho = e^{-U}\,\mathrm{d}x$ with respect to the Lebesgue measure, is *c-log-convex* (res. *c-log-concave*), for $c \in \mathbb{R}$, if the distributional derivative $\nabla^2 U$ satisfies $\nabla^2 U \preceq c\,\mathrm{Id}_n$ (res. $\nabla^2 U \succeq c\,\mathrm{Id}_n$), where $\preceq$ (res. $\succeq$) are in the sense of positive semidefinite order.

It follows from Caffarelli's work that if the source measure is α-log-convex, and the target measure is κ-log-concave, then the Lipschitz constant of the Brenier map between the source and target can be bounded in terms of α and κ.

Theorem 1.2 (Caffarelli [5], Kolesnikov [30])**.** *Let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ and $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ be probability measures on $\mathbb{R}^n$, with μ supported on all of $\mathbb{R}^n$, such that there exist $\alpha > 0, \kappa > 0$ with*

$$\nabla^2 V \preceq \alpha\,\mathrm{Id}_n \quad \text{and} \quad \nabla^2 W \succeq \kappa\,\mathrm{Id}_n\,.$$

Let $\nabla\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be the Brenier map transporting μ to ν. Then,

$$\|\nabla^2\Phi\|_{L^\infty(\mathrm{d}x)} \leq \sqrt{\frac{\alpha}{\kappa}}, \tag{1.1}$$

where $\|\cdot\|_{L^\infty(\mathrm{d}x)}$ is the L^∞ operator norm.

The bound (1.1) is sharp as can be seen by taking Gaussians for the source and target, $\mu = \mathcal{N}(0, \alpha\,\mathrm{Id}_n)$ and $\nu = \mathcal{N}(0, \kappa\,\mathrm{Id}_n)$. One special feature of (1.1) is that the right-hand side is dimension-independent, which leads to numerous dimension-independent functional inequalities [8, 10, 25, 27]. For example, suppose μ satisfies logarithmic Sobolev inequality with constant c_μ, that is, for every test function f,

$$\int (\log f^2) f^2\,\mathrm{d}\mu - \int f^2\,\mathrm{d}\mu \,\log \int f^2\,\mathrm{d}\mu \leq c_\mu \int |\nabla f|^2\,\mathrm{d}\mu. \tag{1.2}$$

Then, the bound (1.1) can be used to show that ν satisfies a logarithmic Sobolev inequality with constant $\frac{\alpha}{\kappa}c_\mu$:

(1.3)
$$\int (\log f^2) f^2 \,\mathrm{d}\nu - \int f^2 \,\mathrm{d}\nu \log \int f^2 \,\mathrm{d}\nu \overset{(\nabla\Phi)_\sharp\mu=\nu}{=} \int (\log(f\circ\nabla\Phi)^2)(f\circ\nabla\Phi)^2 \,\mathrm{d}\mu$$
$$- \int (f\circ\nabla\Phi)^2 \,\mathrm{d}\mu \log \int (f\circ\nabla\Phi)^2 \,\mathrm{d}\mu \overset{(1.2)}{\leq} c_\mu \int |\nabla(f\circ\nabla\Phi)|^2 \,\mathrm{d}\mu \leq c_\mu \int \|\nabla^2\Phi\|_{L^\infty}^2 |(\nabla f)\circ\nabla\Phi|^2 \,\mathrm{d}\mu$$
$$\overset{(1.1)}{\leq} \frac{\alpha}{\kappa} c_\mu \int |(\nabla f)\circ\nabla\Phi|^2 \,\mathrm{d}\mu = \frac{\alpha}{\kappa} c_\mu \int |\nabla f|^2 \,\mathrm{d}\nu.$$

While Theorem 1.2 is highly useful there are numerous classes of probability measures which fall outside its scope as they do not satisfy the required convexity. In this work we will focus on a particular such class of probability measures.

Definition 1.3. A probability measure ρ on $\mathbb{R}^n$, with density $\mathrm{d}\rho = e^{-U}\,\mathrm{d}x$ with respect to the Lebesgue measure where $U \in L^1_{\mathrm{loc}}(\mathrm{d}x)$, is *c-log-subharmonic* (res. *c-log-superharmonic*), for $c \in \mathbb{R}$, if the distributional derivative ΔU satisfies $\Delta U \leq c$ (res. $\Delta U \geq c$).

Log-subharmonic measures arise naturally in complex-analytic settings since the modulus of a holomorphic function on $\mathbb{C}^n$ is log-subharmonic (cf. Theorem 1.15). In particular, we will see below (cf. Section 1.4) that c-log-subharmonic measures are the right objects to investigate in the context of the generalized Wehrl conjecture. In addition, probability measures in two-dimensional one-component plasma models are log-subharmonic (cf. Section 1.5). We refer the reader to [23, 20, 21, 3] for various functional inequalities enjoyed by c-log-subharmonic measures.

Our first main result establishes an analogue of Theorem 1.2 on the level of the trace under log-subharmonicity assumptions.

Theorem 1.4. *Let* $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ *and* $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ *be probability measures on* $\mathbb{R}^n$*, with* μ *supported on all of* $\mathbb{R}^n$*, such that there exist* $\alpha > 0, \kappa > 0$ *with*

$$\Delta V \leq \alpha n \quad \text{and} \quad \nabla^2 W \succeq \kappa \,\mathrm{Id}_n\,.$$

Let $\nabla\Phi : \mathbb{R}^n \to \mathbb{R}^n$ *be the Brenier map transporting* μ *to* ν*. Then, the Laplacian* $\Delta\Phi$ *satisfies*

$$\|\Delta\Phi\|_{L^\infty(\mathrm{d}x)} \leq n\sqrt{\frac{\alpha}{\kappa}}. \tag{1.4}$$

The bound (1.4) is sharp as can be seen by taking

$$\mu = \mathcal{N}(0, \sigma^2 \,\mathrm{Id}_n), \quad \nu = \mathcal{N}(0, \mathrm{Id}_n); \quad \kappa = 1, \quad \alpha = \frac{1}{\sigma^2},$$

since in this case $\nabla\Phi(x) = \frac{x}{\sigma}$ so

$$\Delta\Phi(x) = \frac{n}{\sigma} = n\sqrt{\frac{\alpha}{\kappa}} \quad \text{for all } x.$$

In particular the right-hand side of (1.4) must be dimension-dependent. Let us also remark that following the first version of this work, Gozlan and Sylvestre showed [19, §6] that our assumption on W can be weakened.

Remark 1.5 (Lipschitz bounds)**.** Since the Brenier map $\nabla\Phi$ between μ and ν is a gradient of a convex function, our bound (1.4) implies that under the assumptions of Theorem 1.4 we get the Lipschitz bound

$$\|\nabla^2\Phi\|_{L^\infty(\mathrm{d}x)} \leq n\sqrt{\frac{\alpha}{\kappa}}. \tag{1.5}$$

The bound (1.5) is sharp as can be seen by taking

$$\mu = \mathcal{N}(0, \Sigma_\epsilon), \quad \nu = \mathcal{N}(0, \mathrm{Id}_n); \quad \kappa = 1, \quad \alpha_\epsilon = 1 + \frac{n-1}{n}\epsilon,$$

where Σ_ϵ is a diagonal matrix with one entry equal to $1/n$ and the rest of the entries equal to $1/\epsilon$ for some $\epsilon \in (0, 1]$. In this case $\nabla\Phi(x) = \Sigma_\epsilon^{-1} x$ so that

$$\|\nabla^2\Phi\|_{L^\infty(\mathrm{d}x)} \leq n = \lim_{\epsilon\to 0} n\sqrt{\frac{\alpha_\epsilon}{\kappa}}.$$

The bound (1.5) can be seen as the dimensional price that one has to pay if the log-convexity assumption in Theorem 1.2 is relaxed to log-subharmonicity assumption.

While Theorem 1.4 is the trace analogue of Theorem 1.2, its applications to functional inequalities is different in nature. The argument in (1.3) can no longer be used if we only have the bound (1.4) to transfer functional inequalities (though one can always of course use the dimension-dependent bound (1.5)). Instead, the natural playground for trace bounds such as (1.4) is majorization.

1.2. **Volume contracting maps and majorization.** The starting point of this section is the result of Melbourne and Roberto [34] on the relation between volume-contracting transport maps and majorization. To introduce their ideas we recall the basic definitions from the theory of majorization between probability measures [33].

Definition 1.6. A probability measure $h\,\mathrm{d}x$ on $\mathbb{R}^n$ *majorizes* a probability measure $g\,\mathrm{d}x$ on $\mathbb{R}^n$ if

$$\int_{\mathbb{R}^n} \varphi(g)\,\mathrm{d}x \leq \int_{\mathbb{R}^n} \varphi(h)\,\mathrm{d}x \quad \text{for every convex function } \varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}.$$

For example, we may take $\varphi(x) = x\log x$ to get that the differential entropy of $h\,\mathrm{d}x$ must be bigger than the differential entropy of $g\,\mathrm{d}x$. The observation of Melbourne and Roberto is that the type of regularity of transport maps relevant to majorzation is volume-contraction.

Definition 1.7. A differentiable transport map $\mathrm{T} : \mathbb{R}^n \to \mathbb{R}^n$ between $g\,\mathrm{d}x$ to $h\,\mathrm{d}x$ is a *c-volume-contraction* map if there exists a constant $c > 0$ such that

$$\|\det \nabla \mathrm{T}\|_{L^\infty(g\,\mathrm{d}x)} \leq c.$$

Theorem 1.8 (Melbourne-Roberto [34])**.** *Suppose there exists a transport map $\mathrm{T} : \mathbb{R}^n \to \mathbb{R}^n$ between probability measures $g\,\mathrm{d}x$ and $h\,\mathrm{d}x$ which is a 1-volume-contraction map. Then, $h\,\mathrm{d}x$ majorizes $g\,\mathrm{d}x$.*

The connection between volume-contracting transport maps and Theorem 1.4 is a simple consequence of the arithmetic-geometric-mean inequality.

Theorem 1.9 (Volume contraction of the Brenier map)**.** *Let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ and $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ be probability measures on $\mathbb{R}^n$, with μ supported on all of $\mathbb{R}^n$, such that there exist $\alpha > 0, \kappa > 0$ with*

$$\Delta V \leq \alpha n \quad \text{and} \quad \nabla^2 W \succeq \kappa\,\mathrm{Id}_n\,.$$

Then, the Brenier map $\nabla\Phi$ transporting μ to ν is a $\left(\frac{\alpha}{\kappa}\right)^{\frac{n}{2}}$-volume-contraction map,

$$\|\det \nabla^2\Phi\|_{L^\infty(\mathrm{d}x)} \leq \left(\frac{\alpha}{\kappa}\right)^{\frac{n}{2}}. \tag{1.6}$$

Combining Theorem 1.8 and Theorem 1.9 we obtain majorization under log-subharmonicity.

Theorem 1.10 (Majorization)**.** *Let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ and $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ be probability measures on $\mathbb{R}^n$, with μ supported on all of $\mathbb{R}^n$, such that there exist $\alpha > 0, \kappa > 0$ with*

$$\Delta V \leq \alpha n \quad \text{and} \quad \nabla^2 W \succeq \kappa\,\mathrm{Id}_n\,.$$

If $\frac{\alpha}{\kappa} \leq 1$, then ν majorizes μ.

In [34] the authors combined Theorem 1.8 and Theorem 1.2 to establish majorization under log-concavity assumptions. Theorem 1.10 improves on this result by relaxing to a log-subharmonicity assumption.

So far one might be under the impression that in the context of majorization only the weaker determinant bound (1.6), as opposed to the stronger trace bound (1.4), plays a role. This is not the case however as the following results show. As a first consequence of the trace bound (1.4) we show that we have monotonicity of majorization along Wasserstein geodesics.

Theorem 1.11 (Monotonicity along Wasserstein geodesics)**.** *Let* $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ *and* $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ *be probability measures on* $\mathbb{R}^n$*, with* μ *supported on all of* $\mathbb{R}^n$*, such that there exist* $\alpha > 0, \kappa > 0$ *with*

$$\Delta V \leq \alpha n \quad \textit{and} \quad \nabla^2 W \succeq \kappa \operatorname{Id}_n .$$

Let $(\rho_t)_{t\in[0,1]}$ *be the geodesic in Wasserstein space*[1] *connecting* μ *and* ν*. Then, if* $\frac{\alpha}{\kappa} \leq 1$*,*

$$[0,1] \ni t \mapsto \int_{\mathbb{R}^n} \varphi(\rho_t(x))\,\mathrm{d}x \text{ is monotonically non-decreasing} \tag{1.7}$$

for every convex function $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}$*.*

The second consequence of the trace bound (1.4) pertains to the important special case of Theorem 1.10 regarding differential entropy. Letting $\varphi(x) = x \log x$ in the definition of majorization we get that, if $\frac{\alpha}{\kappa} \leq 1$, then the target measure must have a bigger differential entropy than the source measure,

$$\mathrm{H}(\nu) \geq \mathrm{H}(\mu), \tag{1.8}$$

where

$$\mathrm{H}(\rho) := \int_{\mathbb{R}^n} \rho \log \rho \quad \text{for absolutely continuous probability measures } \rho \text{ on } \mathbb{R}^n.$$

Our next result, which requires the trace bound (1.4) as opposed to the weaker determinant bound (1.6), provides a quantitative improvement of (1.8).

Theorem 1.12 (Stability of entropy domination)**.** *Let* $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ *and* $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ *be probability measures on* $\mathbb{R}^n$*, with* μ *supported on all of* $\mathbb{R}^n$*, such that there exist* $\alpha > 0, \kappa > 0$ *with*

$$\Delta V \leq \alpha n \quad \textit{and} \quad \nabla^2 W \succeq \kappa \operatorname{Id}_n .$$

Let $\nabla\Phi : \mathbb{R}^n \to \mathbb{R}^n$ *be the Brenier map transporting* μ *to* ν*. If* $\frac{\alpha}{\kappa} \leq 1$*, then*

$$\mathrm{H}(\nu) - \mathrm{H}(\mu) \geq \frac{1}{2n^2} \int_{\mathbb{R}^n} \|\nabla^2\Phi - \operatorname{Id}_n\|_F^2 \,\mathrm{d}\mu,$$

where $\|\cdot\|_F$ *is the Frobenius norm. In particular,*

$$\mathrm{H}(\nu) = \mathrm{H}(\mu) \qquad \Longrightarrow \qquad \mu = \nu. \tag{1.9}$$

Theorem 1.12 implies that if the differential entropies of μ and ν are close, then μ and ν themselves must be close, in the sense that the transport map $\nabla\Phi$ between them must be close to the identity. In particular, the next corollary shows that if the entropies of μ and ν match then $\mu = \nu$.

[1] The Wasserstein space is the space of probability measures on $\mathbb{R}^n$, with finite second moment, endowed with metric $(\mu,\nu) \mapsto \inf_{\pi \text{ joint couplings of } \mu \text{ and } \nu} \int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \,\mathrm{d}\pi(x,y)$ [40].

1.2.1. *Volume contraction of the Kim-Milman map.* Our discussion so far focused on the Brenier map of optimal transport. While the regularity properties of the Brenier map are interesting in their own right, for the purpose of majorization any volume-contracting transport map would yield the same result (note however that this would not be the case for Theorem 1.11 and Theorem 1.12). The only other transport map for which regularity properties are known, in particular Lipschitz regularity, is the *Kim-Milman heat flow map* [26]. To define the Kim-Milman map let $(\mathrm{P}_t)_{t\geq 0}$ be the Langevin semigroup associated to ν, and define the family of diffeomorphisms $\mathrm{F}_t : \mathbb{R}^n \to \mathbb{R}^n$ as the solution to the ordinary differential equation,

$$\partial_t \mathrm{F}_t(x) = -\nabla \log \mathrm{P}_t f(\mathrm{F}_t(x)), \quad \mathrm{F}_0(x) = x, \quad \text{for } t > 0 \text{ and } x \in \mathbb{R}^n, \tag{1.10}$$

where $f := \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$. One can check that F_t transports μ to $\rho_t := (\mathrm{P}_t f)\,\mathrm{d}\nu$. In particular, the Kim-Milman map $\mathrm{F} := \lim_{t\to\infty} \mathrm{F}_t$ transports μ to ν. Even assuming the existence of F (which is unclear in our setting), the proof techniques that establish Theorem 1.9 are not applicable for the Kim-Milman map. However, when the target measure ν is $\gamma := \mathcal{N}(0, \mathrm{Id}_n)$, the standard Gaussian measure on $\mathbb{R}^n$, we can in fact establish the analogue of (1.6).

Theorem (informal) 1.13. *Let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ be a probability measure on $\mathbb{R}^n$, with μ supported on all of $\mathbb{R}^n$, such that there exists $\alpha > 0$ with*

$$\Delta V(x) \leq \alpha n, \quad \text{for every } x \in \mathbb{R}^n.$$

Then, under sufficient regularity, the Kim-Milman heat flow map F *between μ and the standard Gaussian γ on $\mathbb{R}^n$ satisfy*

$$\|\det \nabla \mathrm{F}\|_{L^\infty(\mathrm{d}x)} \leq \alpha^{\frac{n}{2}}. \tag{1.11}$$

We have purposely left Theorem 1.13 vague in terms of regularity, and we refer the reader to Section 4 below for a more precise discussion. Note however that in contrast to Theorem 1.4 where the convexity of the potentials ensures that a bound on the Laplacian implies a bound on the Hessian, which is helpful in approximation arguments (cf. Proposition 2.3), in the case of the Kim-Milman map, passing to the limit in (1.11) seems to require stronger assumptions on the densities of μ and ν.

Remark 1.14. For a source measure μ and a target measure ν the Kim-Milman heat flow map is based on the Langevin dynamics whose invariant measure is ν, starting the dynamics at μ. (In the original work of Kim and E. Milman [26] the reverse flow is considered.) Theorem 1.13 only applies to the flow whose invariant measure ν is equal to γ, the Gaussian measure. However, by composing two Kim-Milman maps we could get a $\left(\frac{\alpha}{\kappa}\right)^{\frac{n}{2}}$-volume-contraction map between μ ans ν under the assumptions of Theorem 1.4 (assuming sufficient regularity). Indeed, the Kim-Milman map between γ and $\nu = e^{-W}\,\mathrm{d}x$, with $\nabla^2 W \succeq \kappa\,\mathrm{Id}_n$, is known to be $\frac{1}{\sqrt{\kappa}}$-Lipschitz [26, Theorem 1.1], [35, Theorem 1], so it suffices to compose this map with the Kim-Milman map of Theorem 1.13.

1.3. **Growth estimates in Fock spaces and log-subharmonic functions.** Let us now turn to the applications of majorization. We start by deriving some classical results on growth estimates of log-subharmonic functions. In particular, as mentioned above, the complex-analytic setting naturally gives rise to log-subharmonic functions. Given $\sigma > 0$ let $\gamma_\sigma := \mathcal{N}(0, \sigma\,\mathrm{Id}_2)$ be the centered Gaussian measure with covariance $\sigma\,\mathrm{Id}_2$ on the complex plane $\mathbb{C}$. For functions $f : \mathbb{C} \to \mathbb{R}$, for which the following is finite, let

$$\|f\|_{p,\sigma} := \left(\frac{p}{2\pi\sigma}\int_{\mathbb{C}} |f(x)\gamma_\sigma(z)|^p\,\mathrm{d}z\right)^{\frac{1}{p}}, \quad p > 0,$$

and

$$\|f\|_{\infty,\sigma} := \operatorname{esssup}\left\{|f(z)|\gamma_\sigma(z) : z \in \mathbb{C}\right\}.$$

The following estimate is classical [42, Theorem 2.7].

Theorem 1.15 (Growth estimates in Fock space). *Fix $0 < p \le \infty$ and $z \in \mathbb{C}$. Then,*

$$f(z) \le e^{\frac{|z|^2}{2\sigma}} \quad \text{for all } f \text{ which are entire and satisfy } \|f\|_{p,\sigma} \le 1. \tag{1.12}$$

Proof. Fix $0 < p \le \infty$ and $z \in \mathbb{C}$. Let $f : \mathbb{C} \to \mathbb{R}$ be an entire function satisfying $\|f\|_{p,\sigma} \le 1$. Define the probability measures μ and ν on $\mathbb{C}$ by

$$\mathrm{d}\mu(z) := \frac{|f(z)|^p}{\|f\|_{p,\sigma}}\gamma_{\frac{\sigma}{p}}(z)\,\mathrm{d}z, \quad \text{and} \quad \mathrm{d}\nu(z) := \gamma_{\frac{\sigma}{p}}(z)\,\mathrm{d}z. \tag{1.13}$$

Since $|f|$ is log-subharmonic we get that μ and ν satisfy the assumptions of Theorem 1.9 with $\alpha = \kappa = \frac{p}{\sigma}$. Hence, (1.6) yields that $|\det \nabla\Phi(z)| \le 1$ for almost all $z \in \mathbb{C}$ where $\nabla\Phi$ is the Brenier map between μ and ν. By the Monge-Ampère equation,

$$\frac{|f(z)|^p}{\|f\|_{p,\sigma}}\gamma_{\frac{\sigma}{p}}(z) = \gamma_{\frac{\sigma}{p}}(\nabla\Phi(z)) \det \nabla^2\Phi(z), \tag{1.14}$$

we get

$$|f(z)|^p\gamma_{\frac{\sigma}{p}}(z) \overset{\det \nabla^2\Phi \le 1}{\le} \gamma_{\frac{\sigma}{p}}(\nabla\Phi(z))\|f\|_{p,\sigma} \overset{\|f\|_{p,\sigma}\le 1}{\le} \gamma_{\frac{\sigma}{p}}(\nabla\Phi(z)) \overset{\gamma_{\frac{\sigma}{p}} \le \frac{p}{2\pi\sigma}}{\le} \frac{p}{2\pi\sigma}.$$

Hence,

$$|f(z)|^p \le e^{p\frac{|z|^2}{2\sigma}},$$

which completes the proof. □

Inspecting the proof of Theorem 1.15 it is immediate that it is a statement about cn-log-subharmonic functions, where $f : \mathbb{R}^n \to \mathbb{R}$ is cn-log-subharmonic, for $c \in \mathbb{R}$, if $x \mapsto \log f(x) - c\frac{|x|^2}{2}$ is subharmonic.

Theorem 1.16 (Growth estimates for log-subharmonic functions). *Fix $\beta \ge 0$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be a $(-\beta n)$-log-subharmonic function such that $\int_{\mathbb{R}^n} f\,\mathrm{d}\gamma = 1$, where γ is the standard Gaussian measure on $\mathbb{R}^n$. Then,*

$$f(x) \le (\beta+1)^{\frac{n}{2}} e^{\frac{|x|^2}{2}}. \tag{1.15}$$

Proof. Define the probability measures $\mathrm{d}\mu := f\,\mathrm{d}\gamma$ and $\nu := \gamma$, and note that they satisfy the assumptions of Theorem 1.9 with $\alpha = (\beta+1)$ and $\kappa = 1$. Hence, by (1.6), the Brenier map $\nabla\Phi$ between μ and ν satisfies

$$|\det \nabla^2\Phi| \le (\beta+1)^{\frac{n}{2}}. \tag{1.16}$$

By the Monge-Ampère equation,

$$f(x)\gamma(x) = \gamma(\nabla\Phi(z)) \det \nabla^2\Phi(z) \overset{(1.16)}{\le} \gamma(\nabla\Phi(z))(\beta+1)^{\frac{n}{2}} \overset{\gamma \le (2\pi)^{-\frac{n}{2}}}{\le} \left(\frac{(\beta+1)}{2\pi}\right)^{\frac{n}{2}},$$

so

$$f(x) \le (\beta+1)^{\frac{n}{2}} e^{\frac{|x|^2}{2}}.$$

□

In [18, Lemma 2.1] Theorem 1.16 is obtained under the stronger log-convexity assumption. The log-subharmonic case was treated in [20] (in the print version, see [18, Remark 2.2]), but only for 0-log-subharmonic functions. Theorem 1.16, which is probably already known, generalizes to cn-log-subharmonic functions for any $c \le 0$.

1.4. **The Wehrl conjecture for Glauber states.** Next we apply our majorization results to provide a new proof and stability results for the Wehrl conjecture for Glauber states. The Wehrl entropy is a classical notion of entropy defined for quantum mechanical states. Wehrl presented this notion in [41] and conjectured that this entropy is bounded by 1 (in dimension 1). The conjecture was proven by Lieb [32], and the characterization of the minimizers was established by Carlen [6]. The validity of Wehrl-type conjectures in other settings has been actively investigated, but there are still questions which remain open [31, 15]. In this section we propose a new transport approach towards this problem. In particular, we will provide another proof of the (generalized) Wehrl conjecture for Glauber states using volume-contracting transport maps, as well as stability results.

We will now present the setting of the original Wehrl conjecture following [6]. Fix $h > 0$ to be the Planck constant and let $\hbar := \frac{h}{2\pi}$ be the reduced Planck constant. Given a Schrödinger wave function $\psi \in L^2(\mathbb{R}^d, \mathrm{d}x)$ define the *coherent state transform* $\mathcal{L} : L^2(\mathbb{R}^d, \mathrm{d}x) \to L^2(\mathbb{R}^{2d}, h^{-d}\,\mathrm{d}q\,\mathrm{d}p)$, where $\mathrm{d}x, \mathrm{d}q, \mathrm{d}p$ denote the standard Lebesgue measure on $\mathbb{R}^d$, by

$$\mathcal{L}\psi(q,p) := e^{\frac{i}{2\hbar}\langle q,p\rangle} \int_{\mathbb{R}^d} e^{\frac{i}{\hbar}\langle x,p\rangle} e^{-\frac{|x-q|^2}{2\hbar}} \psi(x)\,\mathrm{d}x. \tag{1.17}$$

The coherent state transform maps a wave function to a function on the phase space, whose modulus $|\mathcal{L}\psi(q,p)|^2$ represents the density (not necessarily normalized) of phase space states. The coherent state transform is isometric, so if $|\psi|_{L^2(\mathbb{R}^d,\mathrm{d}x)} = 1$, hence

$$\rho_\psi(q,p) := h^{-d}|\mathcal{L}\psi(q,p)|^2 \tag{1.18}$$

is a probability measure on the phase space $\mathbb{R}^{2d}$. Wehrl conjectured that the differential entropy of ρ_ψ (the negative of the Wehrl entropy), satisfies the bound

$$\mathrm{H}(\rho_\psi) = \int_{\mathbb{R}^{2d}} \log \rho_\psi \,\mathrm{d}\rho_\psi \le -d\log(2\pi e\hbar). \tag{1.19}$$

Equality is attained in (1.19) for the *Glauber states*,

$$\psi_{\mathrm{Gaussian}}(x) := e^{i\theta}\psi_{q_0,p_0}(x), \tag{1.20}$$

where $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, $q_0, p_0 \in \mathbb{R}^d$, and

$$\psi_{q_0,p_0}(x) := (2\pi\hbar)^{-d} e^{\frac{i}{\hbar}\langle x,p_0\rangle} e^{-\frac{|x-q_0|^2}{2\hbar}}.$$

In other words, according to the conjecture Glauber states maximize the differential entropy (equivalently minimize the Wehrl entropy),

$$\mathrm{H}(\rho_\psi) \le \mathrm{H}(\rho_{\psi_{\mathrm{Gaussian}}}) \quad \text{for all wave functions } \psi \in L^2(\mathbb{R}^d, \mathrm{d}x). \tag{1.21}$$

As mentioned above, the conjectured inequality (1.21) and the characterization of its equality cases (only Glauber states are minimizers) were proven in [32, 6]. In fact, a stronger statement can be made, namely the *generalized Wehrl conjecture for Glauber states*, where the entropy is replaced by other convex functions, which in our terminology means that Glauber states majorize; see [31, 15] for proofs of the inequalities and [16] for their stability. Our next result applies the tools developed in the previous sections to provide a new proof and stability results for the generalized Wehrl conjecture for Glauber states. In fact, we can also address *mixed states*,

$$\rho_{\psi_1,\dots,\psi_k}(q,p) := h^{-d}\sum_{j=1}^k \lambda_j |\mathcal{L}\psi_j(q,p)|^2, \tag{1.22}$$

where $\{\lambda_j\}_{j=1}^k$ are nonnegative weights which sum up to 1, and $\{\psi_j\}_{j=1}^k$ form an orthonormal system in $L^2(\mathbb{R}^d, \mathrm{d}x)$. (The measure constructed in (1.18) is a *pure* states where all the weights but one

vanish.) Our approach treats mixed and pure states in the same way which simplifies the analysis of stability, which in general is more difficult for mixed states [16].

Theorem 1.17 (Generalized Wehrl conjecture for Glauber states)**.** *Set $h = 1$ and let k be a positive integer. Let $\{\psi_j\}_{j=1}^k$ be an orthonormal system in $L^2(\mathbb{R}^d, \mathrm{d}x)$, and let $\{\lambda_j\}_{j=1}^k$ be nonnegative weights which sum up to 1. Define the probability measure*

$$\rho_{\psi_1,\dots,\psi_k}(q,p) := h^{-d} \sum_{j=1}^k \lambda_j |\mathcal{L}\psi_j(q,p)|^2,$$

and let $\nabla\Phi$ be the Brenier map between $\rho_{\psi_1,\dots,\psi_k}$ and $\rho_{\psi_{\mathrm{Gaussian}}}$. Then,

$$\|\Delta\Phi\|_{L^\infty(\mathrm{d}x)} \leq 2d. \tag{1.23}$$

Moreover,

$$\int_{\mathbb{R}^{2d}} \varphi(\rho_{\psi_1,\dots,\psi_k})\,\mathrm{d}p\,\mathrm{d}q \leq \int_{\mathbb{R}^{2d}} \varphi(\rho_{\psi_{\mathrm{Gaussian}}})\,\mathrm{d}p\,\mathrm{d}q \quad \text{for every convex function } \varphi : [0,1] \to \mathbb{R}, \tag{1.24}$$

$$[0,1] \ni t \mapsto \int_{\mathbb{R}^n} \varphi(\rho_t(x))\,\mathrm{d}x \text{ is monotonically non-decreasing}, \tag{1.25}$$

where $(\rho_t)_{t\in[0,1]}$ is the geodesic in Wasserstein space connecting $\rho_{\psi_1,\dots,\psi_k}$ and $\rho_{\psi_{\mathrm{Gaussian}}}$, $\rho_t := (\nabla\Phi_t)_\sharp \rho_{\psi_1,\dots,\psi_k}$ where $\nabla\Phi_t(x) := (1-t)x + t\nabla\Phi(x)$. Finally,

$$\mathrm{H}(\rho_{\psi_{\mathrm{Gaussian}}}) - \mathrm{H}(\rho_{\psi_1,\dots,\psi_k}) \geq \frac{1}{8d^2} \int_{\mathbb{R}^{2d}} \|\nabla^2\Phi - \mathrm{Id}_{2d}\|_F^2\,\mathrm{d}\rho_{\psi_1,\dots,\psi_k}. \tag{1.26}$$

Proof. First note that from the definition of $\rho_{\psi_1,\dots,\psi_k}$, and as $\{\psi_j\}_{j=1}^k$ is an orthonormal system, we have that $\rho_{\psi_1,\dots,\psi_k} \leq 1$. This will allow us later to restrict to convex functions $\varphi : [0,1] \to \mathbb{R}$ in (1.24). To prove (1.23) first note that direct computation shows that $\rho_{\psi_{\mathrm{Gaussian}}}$ is a Gaussian on $\mathbb{R}^{2d}$ with mean $(q_0, -p_0)$ and covariance $\frac{1}{2\pi}\mathrm{Id}_{2d}$,

$$\rho_{\psi_{\mathrm{Gaussian}}} = \mathcal{N}\left((q_0,-p_0), \frac{1}{2\pi}\mathrm{Id}_{2d}\right).$$

On the other hand, the relation between the coherent state transform and the Bargmann transform [22, Proposition 3.4.1] yields (when $h = 1$), for each $j = 1,\dots,k$,

$$|\mathcal{L}\psi_j(q,p)|^2 = 2^{-\frac{d}{2}} \left|\tilde{f}_j(q+ip)\right|^2 e^{-\pi(|q|^2+|p|^2)}, \tag{1.27}$$

where $\tilde{f}_j : \mathbb{C}^d \to \mathbb{C}^d$ is an entire function. In particular,

$$\rho_{\psi_1,\dots,\psi_k}(q,p) = 2^{-\frac{d}{2}} e^{-\pi(|q|^2+|p|^2)} \sum_{j=1}^k \lambda_j \left|\tilde{f}_j(q+ip)\right|^2. \tag{1.28}$$

The logarithm of the modulus of an entire function is subharmonic, so for each $j = 1,\dots,k$, $|\tilde{f}_j(q+ip)|^2$ is log- subharmonic. Since the mixture of log-subharmonic functions is log-subharmonic [20, Proposition 2.2], it follows that $\sum_{j=1}^k \lambda_j \left|\tilde{f}_j(q+ip)\right|^2$ is log-subharmonic. Hence, we can apply Theorem 1.4 and Theorem 1.9, (in fact also Theorem 1.13), with

$$n := 2d, \quad \mu := \rho_{\psi_1,\dots,\psi_k}, \quad \nu := \rho_{\psi_{\mathrm{Gaussian}}}, \quad \kappa := 2\pi, \quad \alpha := 2\pi.$$

In particular, $\sqrt{\frac{\alpha}{\kappa}} = 1$ which yields the bound (1.23) and also means that we can apply Theorem 1.10, Theorem 1.11, and Theorem 1.12 to get (1.24), (1.25), and (1.26), respectively. □

1.5. **Two-dimensional Coulomb gases.** Coulomb gases are important models in mathematical physics with deep connections to random matrix theory; see [7] for a brief survey. In this section we review the two-dimensional (one-component) Coulomb gas, to which our results above will apply. We consider N particles in $\mathbb{R}^2$ of identical charge in a fixed neutralizing background at inverse temperature β, whose interactions are logarithmic. We identify $\mathbb{R}^2 \simeq \mathbb{C}$ and denote the positions of the particles as $z := (z_1, \dots, z_N) \in \mathbb{C}^N$, for $z_i \in \mathbb{C}$. Let $\mu_{\beta,N}$ be the probability measure on $\mathbb{C}^N$ defined by

$$\mathrm{d}\mu_{\beta,N,Q}(z) := \frac{e^{-Q_{\beta,N}(z)-G_{\beta,N}(z)}}{\mathcal{Z}_{\beta,N}}\,\mathrm{d}z \quad \text{with} \quad \mathcal{Z}_{\beta,N} := \int_{\mathbb{C}^N} e^{-Q_{\beta,N}(z)-G_{\beta,N}(z)}\,\mathrm{d}z, \tag{1.29}$$

where

$$Q_{\beta,N}(z) := \beta N \sum_{j=1}^{N} Q(z_j) \quad \text{with} \quad Q : \mathbb{C} \to \mathbb{R}, \tag{1.30}$$

and

$$G_{\beta,N}(z) := -\beta \sum_{i<j=1}^{N} \log|z_i - z_j|. \tag{1.31}$$

The term $Q_{\beta,N}$ models the particle-background interaction, and the term $G_{\beta,N}$ models the particle-particle logarithmic interaction. A common choice for $Q_{\beta,N}$ is taking Q to be quadratic, $Q(z_j) = \frac{|z_j|^2}{2}$. Since $\Delta G_{\beta,N} \le 0$, we can apply our results to the two-dimensional (one-component) Coulomb gas $\mu_{\beta,N,Q}$ and the pure background model $\nu_{\beta,N,Q}$ given by

$$\mathrm{d}\nu_{\beta,N,Q}(z) := \frac{e^{-Q_{\beta,N}(z)}}{\int e^{-Q_{\beta,N}(z)}\,\mathrm{d}z}\,\mathrm{d}z.$$

Theorem 1.18 (Two-dimensional one-component gas)**.** *Suppose there exists $\alpha_2, \kappa_2 > 0$ such that*

$$\Delta Q \le 2\alpha_2 \quad \textit{and} \quad \nabla^2 Q \succeq \kappa_2 \,\mathrm{Id}_2,$$

where $\Delta Q, \nabla^2 Q$ are the distributional Laplacian and Hessian on $\mathbb{R}^2$, respectively. Let $\nabla\Phi_{\beta,N,Q}$ be the optimal transport map between $\mu_{\beta,N,Q}$ and $\nu_{\beta,N,Q}$. Then,

$$\|\Delta\Phi_{\beta,N,Q}\|_{L^\infty(\mathrm{d}x)} \le 2N\sqrt{\frac{\alpha_2}{\kappa_2}}, \quad \textit{and in particular} \quad \|\det\nabla^2\Phi_{\beta,N,Q}\|_{L^\infty(\mathrm{d}x)} \le \left(\frac{\alpha_2}{\kappa_2}\right)^N. \tag{1.32}$$

Moreover, when

$$Q(z_i) = \frac{|z_j|^2}{4} \quad \text{for all} \quad j = 1, \dots, N, \tag{1.33}$$

we have

$$\int_{\mathbb{C}^N} \varphi(\mu_{\beta,N,Q})\,\mathrm{d}z \le \int_{\mathbb{C}^N} \varphi(\nu_{\beta,N,Q})\,\mathrm{d}z \quad \textit{for every convex function } \varphi : \mathbb{R}_{\ge 0} \to \mathbb{R}, \tag{1.34}$$

$$[0,1] \ni t \mapsto \int_{\mathbb{R}^n} \varphi(\rho_{t,\beta,N,Q}(x))\,\mathrm{d}x \text{ is monotonically non-decreasing}, \tag{1.35}$$

where $(\rho_{t,\beta,N,Q})_{t\in[0,1]}$ is the geodesic in Wasserstein space connecting $\mu_{\beta,N,Q}$ and $\nu_{\beta,N,Q}$, $\rho_{t,\beta,N,Q} := (\nabla\Phi_{t,\beta,N,Q})_\sharp\mu_{\beta,N,Q}$ where $\nabla\Phi_{t,\beta,N,Q}(x) := (1-t)x + t\nabla\Phi_{\beta,N,Q}(x)$, and

$$\mathrm{H}(\nu_{\beta,N,Q}) - \mathrm{H}(\mu_{\beta,N,Q}) \ge \frac{1}{8N^2}\int_{\mathbb{C}^N} \|\nabla^2\Phi_{\beta,N,Q} - \mathrm{Id}_{2N}\|_F^2\,\mathrm{d}\mu_{\beta,N,Q}. \tag{1.36}$$

Proof. Denote $\mathrm{d}\mu_{\beta,N,Q} =: e^{-V_{\beta,N,Q}}\,\mathrm{d}z$. The assumption $\nabla^2 Q \succeq \kappa_2 \,\mathrm{Id}_2$ implies that $\nabla^2 Q_{\beta,N} \succeq \kappa_2\beta N\,\mathrm{Id}_{2N}$. Since $\Delta G_{\beta,N} \le 0$, the assumption $\Delta Q \le 2\alpha_2$ implies $\Delta V_{\beta,N,Q} = \Delta Q_{\beta,N} + \Delta G_{\beta,N} \le 2\alpha_2\beta N^2$. Hence, setting

$$n := 2N, \quad \mu := \mu_{\beta,N,Q}, \quad \nu := \nu_{\beta,N,Q}, \quad \kappa := \kappa_2\beta N, \quad \alpha := \alpha_2\beta N,$$

we can apply Theorem 1.4 and Theorem 1.9 to get (1.32). The remaining assertions of the theorem hold only when Q is quadratic, where $\alpha_2 = \kappa_2 = 1$, by Theorem 1.10, Theorem 1.11, and Theorem 1.12, respectively. □

Organization of paper. Section 2 proves Theorem 1.4, while Section 3 focuses on majorization, proving in particular Theorem 1.11 and Theorem 1.12. Section 4 is dedicated to the Kim-Milman heat flow map. Section 5 is an appendix containing a number of technical results.

Acknowledgments. We thank Alessio Figalli and Cole Graham for useful conversations. We thank Rupert Frank for explaining to us the mixed state case of the Wehrl conjecture, Emanuel Milman for Remark 1.14, Ramon van Handel for pointing the applicability of our results to two-dimensional Coulomb gases, and Ammari Bader for pointing out a mistake in a previous version of Theorem 1.18. We also thank the anonymous referee for their helpful comments.

This material is based upon work supported by the National Science Foundation under Award Number DMS-2331920 and DMS 2055686. G.D.P. also acknowledges the support of the Simons Foundation.

2. The divergence of the Brenier map

In this section we will prove L^p estimates on $\Delta\Phi$ (Theorem 2.1), which in particular will imply Theorem 1.4. While the proof of our main result below Theorem 2.1 is quite technical, at the formal level it follows the original (formal) derivation of Caffarelli [5], as we now explain. Our goal is to derive the bound

$$\|\Delta\Phi\|_{L^\infty(\mathrm{d}x)} \le n\sqrt{\frac{\alpha}{\kappa}} \quad \text{when} \quad \Delta V \le \alpha n \text{ and } \nabla^2 W \succeq \kappa\,\mathrm{Id}_n\,. \tag{2.1}$$

The derivation is based on differentiating the Monge-Ampère equation twice, and using the optimality conditions at the point where $\Delta\Phi$ attains its maximum. To this end, given a unit vector $e \in \mathbb{R}^n$ and a function $\xi : \mathbb{R}^n \to \mathbb{R}$, denote by ξ_e (res. ξ_{ee}) the first (res. second) directional derivative of ξ in the direction e. The Monge-Ampère equation reads

$$e^{-V} = e^{-W}(\nabla\Phi)\det\nabla^2\Phi, \tag{2.2}$$

so take the logarithm on both sides of (2.2), and differentiate twice in a fixed direction e, to get

$$V_{ee} = \langle\nabla^2 W(\nabla\Phi)\nabla^2\Phi e, \nabla^2\Phi e\rangle + \langle\nabla W, \nabla\Phi_{ee}\rangle - \partial_{ee}\log\det\nabla^2\Phi. \tag{2.3}$$

Using the identity

$$\partial_{ee}\log\det\nabla^2\Phi = \mathrm{Tr}[(\nabla^2\Phi)^{-1}\nabla^2\Phi_{ee}] - \mathrm{Tr}\left[\left((\nabla^2\Phi)^{-1}\nabla^2\Phi_e\right)^2\right], \tag{2.4}$$

the identity (2.3) becomes

$$V_{ee} = \langle\nabla^2 W(\nabla\Phi)\nabla^2\Phi e, \nabla^2\Phi e\rangle + \langle\nabla W, \nabla\Phi_{ee}\rangle - \mathrm{Tr}[(\nabla^2\Phi)^{-1}\nabla^2\Phi_{ee}] + \mathrm{Tr}\left[\left((\nabla^2\Phi)^{-1}\nabla^2\Phi_e\right)^2\right]. \tag{2.5}$$

Using $\nabla^2 W \succeq \kappa\,\mathrm{Id}_n$, (2.5) implies

$$V_{ee} \ge \kappa|\nabla^2\Phi e|^2 + \langle\nabla W, \nabla\Phi_{ee}\rangle - \mathrm{Tr}[(\nabla^2\Phi)^{-1}\nabla^2\Phi_{ee}] + \mathrm{Tr}\left[\left((\nabla^2\Phi)^{-1}\nabla^2\Phi_e\right)^2\right], \tag{2.6}$$

so summing on both sides of (2.6) over a basis $\{e_i\}$ of $\mathbb{R}^n$ yields

$$\Delta V \geq \kappa \sum_{i=1}^{n} |\nabla^2 \Phi e_i|^2 + \langle \nabla W, \nabla \Delta \Phi \rangle - \mathrm{Tr}[(\nabla^2 \Phi)^{-1} \nabla^2 \Delta \Phi] + \sum_{i=1}^{n} \mathrm{Tr}\left[\left((\nabla^2 \Phi)^{-1} \nabla^2 \Phi_{e_i} \right)^2 \right]. \tag{2.7}$$

Suppose $\Delta \Phi$ attains its maximum x_0. Then the optimality conditions give $\nabla \Delta \Phi(x_0) = 0$ and $\nabla^2 \Delta \Phi(x_0) \preceq 0$, so

$$\langle \nabla W(x_0), \nabla \Delta \Phi(x_0) \rangle - \mathrm{Tr}[(\nabla^2 \Phi(x_0))^{-1} \nabla^2 \Delta \Phi(x_0)] \geq 0. \tag{2.8}$$

On the other hand,

$$\mathrm{Tr}\left[\left((\nabla^2 \Phi)^{-1} \nabla^2 \Phi_e \right)^2 \right] = \mathrm{Tr}[A^2] \geq 0 \quad \text{with} \quad A := (\nabla^2 \Phi)^{-\frac{1}{2}} \nabla^2 \Phi_e (\nabla^2 \Phi)^{-\frac{1}{2}}. \tag{2.9}$$

Hence, combining (2.7), (2.8), (2.9), and applying the Cauchy-Schwarz inequality, shows that

$$\alpha n \geq \Delta V(x_0) \geq \kappa \sum_{i=1}^{n} |\nabla^2 \Phi(x_0) e_i|^2 \geq \kappa \sum_{i=1}^{n} \langle \nabla^2 \Phi(x_0) e_i, e_i \rangle^2. \tag{2.10}$$

By Jensen's inequality,

$$\alpha n \geq \Delta V(x_0) \geq \kappa \sum_{i=1}^{n} \langle \nabla^2 \Phi(x_0) e_i, e_i \rangle^2 \geq \frac{\kappa}{n} [\Delta \Phi(x_0)]^2, \tag{2.11}$$

so it follows that

$$\Delta \Phi(x_0) \leq n \sqrt{\frac{\alpha}{\kappa}}, \tag{2.12}$$

where we used that $\Delta \Phi(x_0) \geq 0$ since Φ is convex. Since x_0 is a point where $\Delta \Phi$ attains its maximum, we conclude (2.1). Making the above argument rigorous is difficult due to the need to show the existence of a point x_0 where $\Delta \Phi$ attains its maximum. Instead, we follow the L^p approach of Kolesnikov [29], [30, §6], but with some necessary modifications (cf. Remark 2.12).

Theorem 2.1. *Let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ and $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ be probability measures on $\mathbb{R}^n$, with μ supported on all of $\mathbb{R}^n$, such that there exist $\alpha > 0, \kappa > 0$ with*

$$\Delta V \leq \alpha n \quad \text{and} \quad \nabla^2 W \succeq \kappa \,\mathrm{Id}_n .$$

Let $\nabla \Phi : \mathbb{R}^n \to \mathbb{R}^n$ be the Brenier map transporting μ to ν. Then,

$$\|\Delta \Phi\|_{L^\infty(\mathrm{d}x)} \leq n \sqrt{\frac{\alpha}{\kappa}}. \tag{2.13}$$

We start by showing that it suffices to prove Theorem 2.1 assuming sufficient regularity for μ and ν. These regularity assumptions are captured as follows.

Assumption 2.2.

(1) V is smooth.
(2) There exists a constant $c > 0$ such that $\nabla^2 V(x) \preceq c \,\mathrm{Id}_n$ for all $x \in \mathbb{R}^n$.
(3) There exist constants $a, b > 0$ such that $\lim_{|x| \to \infty} V(x) - (b|x|^2 - a) \geq 0$.
(4) W is smooth.
(5) There exists a constant $d > 0$ such that $\nabla^2 W(x) \preceq d \,\mathrm{Id}_n$ for all $x \in \mathbb{R}^n$.

Note that Assumption 2.2(3) implies that μ has a finite second moment while the assumption $\nabla^2 W \succeq \kappa \,\mathrm{Id}_n$ implies that ν has a finite second moment (by the Poincaré inequality for ν).

Proposition 2.3. *It suffices to prove Theorem 2.1 for μ, ν which in addition to satisfying the assumptions of Theorem 2.1 also satisfy Assumption 2.2.*

Proof. The proof is based on the composition of the following two steps.
Step 1. Given μ, ν which satisfy the assumptions of Theorem 2.1 we show that there exist sequences of probability measures $\{\mu_k\}, \{\nu_k\}$ converging weakly to μ, ν, respectively, such that, for all k, each pair μ_k, ν_k satisfies the assumptions of Theorem 2.1 and Assumption 2.2.
Step 2. Assume that Theorem 2.1 holds true for each pair μ_k, ν_k and then pass to the limit to deduce (2.13).

Before we start with step 1 let us state a number of important smoothing properties of the Orenstein-Uhlenbeck semigroup which we will use to smooth out the measures μ and ν.

Proposition 2.4. *Let γ be the standard Gaussian measure in $\mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be a nonnegative function in $L^1(\gamma)$. Let $(\mathrm{P}_t)_{t\geq 0}$ be the Orenstein-Uhlenbeck semigroup,*

$$\mathrm{P}_t f(x) := \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\,\mathrm{d}\gamma(y), \quad t \geq 0, \quad x \in \mathbb{R}^n. \tag{2.14}$$

(i) For every $x \in \mathbb{R}^n$ and $t > 0$,

$$\nabla^2 \log \mathrm{P}_t f(x) \succeq -\frac{e^{-2t}}{1-e^{-2t}}\,\mathrm{Id}_n \quad \text{for all } x \in \mathbb{R}^n \text{ and } t > 0. \tag{2.15}$$

(ii) Suppose f is c-log-concave for $c \in \mathbb{R}$ (i.e., $x \mapsto \log f(x) - c\frac{|x|^2}{2}$ is concave). Then, for every $x \in \mathbb{R}^n$,

$$\nabla^2 \log \mathrm{P}_t f(x) \preceq \frac{e^{-2t}c}{1-c(1-e^{-2t})}\,\mathrm{Id}_n \begin{cases} \text{for any } t \in [0,\infty) & \text{if } c \leq 1 \\ \text{for any } t \in \left[0, \log\left(\sqrt{\frac{c}{c-1}}\right)\right] & \text{if } c > 1. \end{cases} \tag{2.16}$$

(iii) Suppose f is cn-log-subharmonic for some $c \leq 0$ (i.e., $x \mapsto \log f(x) - c\frac{|x|^2}{2}$ is subharmonic). Then, for every $x \in \mathbb{R}^n$ and $t > 0$,

$$\Delta \log \mathrm{P}_t f(x) \geq \frac{e^{-2t}cn}{1-c(1-e^{-2t})} \geq cn. \tag{2.17}$$

The proof of Proposition 2.4 (with some extra results) is given in Section 5.2.

Let us begin with **Step 1**. We first construct the sequence $\{\mu_k \propto e^{-V_k}\,\mathrm{d}x\}$ whose members all satisfy the assumptions of Theorem 2.1 as well as Assumption 2.2(1-3). Given $\mu = e^{-V}\,\mathrm{d}x$ let

$$\tilde{V}_k := -\log \mathrm{P}_{\frac{1}{k}} e^{-V}.$$

Then $\tilde{V}_k$ is smooth, and by Proposition 2.4(iii) it satisfies $\Delta\tilde{V}_k(x) \leq \alpha n$ for all $x \in \mathbb{R}^n$. This shows that we can construct a measure $\propto e^{-\tilde{V}_k}\,\mathrm{d}x$ which satisfies the assumptions of Theorem 2.1 as well as Assumption 2.2(1). Moreover, Assumption 2.2(2) is also satisfied for some c_k by Proposition 2.4(i). Next we modify $\tilde{V}_k$ so that Assumption 2.2(3) is also satisfied. Define

$$V_k(x) := \left(1 - \frac{1}{k}\right)\tilde{V}_k(x) + \frac{1}{k}\alpha\frac{|x|^2}{2}$$

and note that $\Delta V_k \leq \alpha n$, V_k is smooth, and

$$\nabla^2 V_k \preceq \left(\left(1-\frac{1}{k}\right)c_k + \frac{1}{k}\alpha\right)\mathrm{Id}_n,$$

so the measure $\propto e^{-V_k}\,\mathrm{d}x$ satisfies the assumptions of Theorem 2.1 as well as Assumptions 2.2(1-2). Let us show that V_k also satisfies Assumption 2.2(3). To this end we first need to argue that V is positive outside of some ball. Indeed, since $\Delta V \leq \alpha n$, we have that, for any $x_0 \in \mathbb{R}^n$, the function

$$\mathbb{R}^n \ni x \mapsto e^{\alpha\frac{|x-x_0|^2}{2} - V(x)}$$

is subharmonic. Hence, choosing x_0 such that $|x_0| > R$ for some $R > 0$, we have

$$e^{-V(x_0)} \le \fint_{B_1(x_0)} e^{\alpha\frac{|x-x_0|^2}{2}-V(x)}\,\mathrm{d}x \le \frac{e^{\frac{\alpha}{2}}}{\mathrm{Vol}(B_1)}\int_{(B_{R-1}(0))^c} e^{-V(x)}\,\mathrm{d}x. \tag{2.18}$$

The right-hand side of (2.18) converges to 0 as $R \to \infty$ since $\int e^{-V}$ is finite. Hence, $\lim_{x\to\infty} V(x) = 0$ and Assumption 2.2(3) follows by the construction of V_k. Finally, let $\mu_k \propto e^{-V_k}\,\mathrm{d}x$ and note that $\mu_k \to \mu$ weakly.

Next we construct the sequence $\{\nu_k \propto e^{-W_k}\,\mathrm{d}x\}$ whose members all satisfy the assumptions of Theorem 2.1 as well as Assumption 2.2(4-5). Given $\nu = e^{-W}\,\mathrm{d}x$ let

$$W_k := -\log \mathrm{P}_{\frac{1}{k}} e^{-W}$$

and let $\{\nu_k \propto e^{-W_k}\,\mathrm{d}x\}$. Then, as in the construction of V_k, W_k is smooth and there exists a constant d_k such that $\nabla^2 W \preceq d_k\,\mathrm{Id}_n$. Hence, W_k satisfies Assumption 2.2(4-5). Finally, by Proposition 2.4(ii), W_k satisfies the assumptions of Theorem 2.1 with $\nabla^2 W_k \succeq \kappa_k$ where

$$\kappa_k := \frac{\kappa e^{-\frac{2}{k}}}{1+\kappa(1-e^{-\frac{2}{k}})} = \kappa - \frac{\kappa(\kappa+1)(1-e^{-\frac{2}{k}})}{1+\kappa(1-e^{-\frac{2}{k}})}. \tag{2.19}$$

We now move to **Step 2**. Assume that Theorem 2.1 holds true for each pair μ_k, ν_k, and let $\nabla\Phi_k$ be the Brenier map between μ_k and ν_k. Then, for k large enough, by Theorem 2.1 and Remark 1.5,

$$\|\Delta\Phi_k\|_{L^\infty(\mathrm{d}x)} \le n\sqrt{\frac{\alpha}{\kappa_k}} \quad \text{and} \quad \nabla^2\Phi_k(x) \preceq n\sqrt{\frac{\alpha}{\kappa_k}}\,\mathrm{Id}_n \ \text{ for every } x \in \mathbb{R}^n. \tag{2.20}$$

By [35, Lemma 1], which follows [38, Lemma 2.1] building on [26, Lemma 3.3], we get that, up to a subsequence, $\{\nabla\Phi_k\}$ converges almost everywhere to some transport map T between μ and ν. Since $\{\nabla\Phi_k\}$ converges so does $\{\Phi_k\}$ to some convex function Φ. It follows that $T = \nabla\Phi$ is the Brenier map between μ and ν. Finally, the proof is complete since Φ_k converges to Φ locally uniformly, $\Delta\Phi_k$ converges to $\Delta\Phi$ as distribution, and as

$$\begin{aligned}
\|\Delta\Phi\|_{L^\infty(\mathrm{d}x)} &= \sup_{\eta\in C_c^\infty, \|\eta\|_{L^1}=1} \int \Delta\Phi(x)\eta(x)\,\mathrm{d}x = \sup_{\eta\in C_c^\infty, \|\eta\|_{L^1}=1} \int \Phi(x)\Delta\eta(x)\,\mathrm{d}x \\
&= \sup_{\eta\in C_c^\infty, \|\eta\|_{L^1}=1} \lim_{k\to\infty} \int \Phi_k(x)\Delta\eta(x)\,\mathrm{d}x = \sup_{\eta\in C_c^\infty, \|\eta\|_{L(\mathrm{d}x)}=1} \lim_{k\to\infty} \int \Delta\Phi_k(x)\eta(x)\,\mathrm{d}x \\
&\le \lim_{k\to\infty} \|\Delta\Phi_k\|_{L^\infty(\mathrm{d}x)} = n\sqrt{\frac{\alpha}{\kappa}}.
\end{aligned}$$

□

Proof of Theorem 2.1. In light of Proposition 2.3 we may assume Assumption 2.2 from here on. Under the assumptions of Theorem 2.1, together with Assumption 2.2, we have that there exists an optimal transport map $\nabla\Phi$ between μ and ν which is smooth [9, Theorem 1.1], satisfies the Monge-Ampère equation,

$$e^{-V(x)} = e^{-W(\nabla\Phi(x))}\det\nabla^2\Phi(x), \quad \text{for all } x \in \mathbb{R}^n, \tag{2.21}$$

and, by Theorem 1.2, satisfies the bound

$$0 \preceq \nabla^2\Phi(x) \preceq C\,\mathrm{Id}_n \quad \text{for all } x \in \mathbb{R}^n, \tag{2.22}$$

for some constant $C > 0$. Taking the logarithm in (2.21) we get

$$V(x) = W(\nabla\Phi(x)) - \log\det\nabla^2\Phi(x), \tag{2.23}$$

so, for every $x, y \in \mathbb{R}^n$,

$$V(x+y)+V(x-y)-2V(x) = \{W(\nabla\Phi(x+y))+W(\nabla\Phi(x-y))-2W(\nabla\Phi(x))\} - \log\left[\frac{\det\nabla^2\Phi(x+y)\det\nabla^2\Phi(x-y)}{(\det\nabla^2\Phi(x))^2}\right]. \tag{2.24}$$

Equation (2.24) is the finite difference analogue of equation (2.3). The left-hand side of (2.24) is a second-order finite difference of V which we wish to relate to ΔV, so to this end we introduce the following Δ_ϵ operator and its properties; see Section 5.1 for the proof.

Lemma 2.5. *Fix $\epsilon > 0$ and for a function $f : \mathbb{R}^n \to \mathbb{R}$ denote*

$$\Delta_\epsilon f(x) := \fint_{\partial B_\epsilon(0)} [f(x+y)-f(x)]\,\mathrm{d}y, \tag{2.25}$$

where

$$\fint_{\partial B_\epsilon(0)} f := \frac{1}{|\partial B_\epsilon(0)|}\int_{\partial B_\epsilon(0)} f.$$

Then,

$$\lim_{\epsilon\to 0}\frac{\Delta_\epsilon f(x)}{\epsilon^2} = \frac{\Delta f(x)}{2n}, \tag{2.26}$$

where Δf is the distributional Laplacian of f. Further, if $\Delta f \le \ell$, then

$$\Delta_\epsilon f \le \frac{\ell}{n}\frac{\epsilon^2}{2}, \qquad \forall\ \epsilon > 0. \tag{2.27}$$

We now integrate (2.24) over y with respect to the uniform measure on the sphere $\partial B_\epsilon(x)$ of radius ϵ centered at x. By the assumption $\Delta V \le \alpha n$, together with (2.27), we have

$$\epsilon^2\alpha \ge 2\Delta_\epsilon V(x) = 2\fint_{\partial B_\epsilon(0)} [W(\nabla\Phi(x+y))-W(\nabla\Phi(x))]\,\mathrm{d}y - \fint_{\partial B_\epsilon(0)} \log\left[\frac{\det\nabla^2\Phi(x+y)\det\nabla^2\Phi(x-y)}{(\det\nabla^2\Phi(x))^2}\right]\mathrm{d}y. \tag{2.28}$$

On the other hand, by the assumption $\nabla^2 W \succeq \kappa\,\mathrm{Id}_n$,

$$W(\nabla\Phi(x+y))-W(\nabla\Phi(x)) \ge \langle(\nabla W)(\nabla\Phi(x)), \nabla\Phi(x+y)-\nabla\Phi(x)\rangle + \frac{\kappa}{2}|\nabla\Phi(x+y)-\nabla\Phi(x)|^2, \tag{2.29}$$

which is the analogue of (2.6). Combining (2.28) and (2.29) implies

$$\epsilon^2\alpha \ge \kappa\fint_{\partial B_\epsilon(0)} |\nabla\Phi(x+y)-\nabla\Phi(x)|^2\,\mathrm{d}y + \fint_{\partial B_\epsilon(0)}\left\{2\langle(\nabla W)(\nabla\Phi(x)), \nabla\Phi(x+y)-\nabla\Phi(x)\rangle - \log\left[\frac{\det\nabla^2\Phi(x+y)\det\nabla^2\Phi(x-y)}{(\det\nabla^2\Phi(x))^2}\right]\right\}\mathrm{d}y. \tag{2.30}$$

Next we will multiply both sides of (2.30) by $(\Delta_\epsilon\Phi(x))^p$ and integrate against $e^{-V(x)}\,\mathrm{d}x$. This requires showing that some integrals are finite.

Lemma 2.6. *For every $p > 0$ we have*

(2.31)

$$\epsilon^2\alpha \int (\Delta\Phi_\epsilon(x))^p e^{-V(x)}\,\mathrm{d}x \geq \kappa \int \fint_{\partial B_\epsilon(0)} \left[|\nabla\Phi(x+y) - \nabla\Phi(x)|^2\right] \mathrm{d}y (\Delta_\epsilon\Phi(x))^p e^{-V(x)}\,\mathrm{d}x$$

$$+ \int (\Delta_\epsilon\Phi(x))^p e^{-V(x)}$$

$$\cdot \left\{ \fint_{\partial B_\epsilon(0)} \left\{ 2\langle \nabla W(\nabla\Phi(x)), \nabla\Phi(x+y) - \nabla\Phi(x)\rangle - \log\left[\frac{\det\nabla^2\Phi(x+y)\det\nabla^2\Phi(x-y)}{(\det\nabla^2\Phi(x))^2}\right]\right\} \mathrm{d}y \right\} \mathrm{d}x,$$

and all the integrals are finite.

Proof. Once we show that the integrals are finite, Equation (2.31) follows by multiplying both sides of (2.30) by $(\Delta_\epsilon\Phi)^p$ and integrating against $e^{-V}\,\mathrm{d}x$. Note that at the moment we are only interested in these quantities to be finite and not uniformly bounded in ϵ. The integral $\int(\Delta_\epsilon\Phi)^p e^{-V}\,\mathrm{d}x$ is finite by (2.22) and (2.27). The integral $\int \fint_{\partial B_\epsilon(0)} \left[|\nabla\Phi(x+y) - \nabla\Phi(x)|^2\right] \mathrm{d}y(\Delta_\epsilon\Phi)^p e^{-V}\,\mathrm{d}x$ is finite since by the fundamental theorem of calculus we can write the difference $\nabla\Phi(x+y)-\nabla\Phi(x)$ as an integral along a path between $x+y$ and x of $\nabla^2\Phi$ multiplied by a vector of length $\leq\epsilon$, so the difference is bounded since $\nabla^2\Phi$ is bounded by (2.22). To bound the term $\langle\nabla W(\nabla\Phi(x)), \nabla\Phi(x+y)-\nabla\Phi(x)\rangle$ use Assumption 2.2(5) according to which ∇W grows at most linearly, and the fact that again $\nabla\Phi(x+y)-\nabla\Phi(x)$ can be bounded by a constant, to see that $\int(\Delta_\epsilon\Phi)^p \fint_{\partial B_\epsilon(0)} 2\langle\nabla W(\nabla\Phi(x)), \nabla\Phi(x+y)-\nabla\Phi(x)\rangle e^{-V}$ is finite since ν has a finite first moment. Finally, to bound the term

$$\int \fint_{\partial B_\epsilon(0)} \left|\log\left[\frac{\det\nabla^2\Phi(x+y)\det\nabla^2\Phi(x-y)}{(\det\nabla^2\Phi(x))^2}\right]\right| e^{-V(x)}\,\mathrm{d}x\,\mathrm{d}y$$

we use identity (2.28) so we need to bound

$$\int \fint_{\partial B_\epsilon(0)} |W(\nabla\Phi(x+y)) - W(\nabla\Phi(x))|(\Delta_\epsilon\Phi)^p e^{-V}\,\mathrm{d}y\,\mathrm{d}x$$

as well as

$$\int |\Delta_\epsilon V| e^{-V}\,\mathrm{d}x.$$

For the first term, again by (2.22), the distance between $\nabla\Phi(x+y)$ and $\nabla\Phi(x)$ is bounded by a constant, so by Assumption 2.1(5), $\fint_{\partial B_\epsilon(0)} |W(\nabla\Phi(x+y)) - W(\nabla\Phi(x))|$ is linear in $\nabla\Phi$ (at some point between $x+y$ and x) and, since $\nabla^2\Phi$ is bounded, $\int \fint_{\partial B_\epsilon(0)} |W(\nabla\Phi(x+y)) - W(\nabla\Phi(x))|(\Delta_\epsilon\Phi)^p e^{-V}\,\mathrm{d}y\,\mathrm{d}x$ is finite since ν has a finite first moment. Finally, for the term $\int|\Delta_\epsilon V|e^{-V}\,\mathrm{d}x$, we use the definition of Δ_ϵ and note that at infinity V grows quadratically so $\int|\Delta_\epsilon V|e^{-V}\,\mathrm{d}x$ is finite since μ has a finite second moment. □

We will show in Proposition 2.8 below that the second term in (2.31) is nonnegative, which is the replacement of (2.8)-(2.9). Let us complete the proof of Theorem 2.1 assuming the validity of Proposition 2.8.

Lemma 2.7. *Suppose that for every $p > 0$,*

$$\epsilon^2\alpha \int (\Delta_\epsilon\Phi(x))^p e^{-V(x)}\,\mathrm{d}x \geq \kappa \int \fint_{\partial B_\epsilon(0)} \left[|\nabla\Phi(y+x) - \nabla\Phi(x)|^2\right] \mathrm{d}y (\Delta_\epsilon\Phi(x))^p e^{-V(x)}\,\mathrm{d}x. \tag{2.32}$$

Then,

$$\|\Delta\Phi\|_{L^\infty(\mu)} \leq n\sqrt{\frac{\alpha}{\kappa}}. \tag{2.33}$$

Proof. Dividing both sides of (2.32) by ϵ^{2p+2}, and using (2.26), we get

$$\begin{aligned}\alpha \int \left(\frac{\Delta\Phi}{2n}\right)^p e^{-V}\,\mathrm{d}x &\geq \kappa \int \left\{ \lim_{\epsilon\to 0} \frac{1}{\epsilon^2} \fint_{\partial B_\epsilon(0)} \left[|\nabla\Phi(y+x) - \nabla\Phi(x)|^2\right] \mathrm{d}y \right\} \left(\frac{\Delta\Phi}{2n}\right)^p e^{-V}\,\mathrm{d}x \\ &= \kappa \int \left\{ \fint_{\partial B_1(0)} \lim_{\epsilon\to 0} \left| \frac{\nabla\Phi(x+\epsilon y) - \nabla\Phi(x)}{\epsilon} \right|^2 \mathrm{d}y \right\} \left(\frac{\Delta\Phi}{2n}\right)^p e^{-V}\,\mathrm{d}x \\ &= \kappa \int \left\{ \fint_{\partial B_1(0)} \left|\nabla^2\Phi(x) y\right|^2 \mathrm{d}y \right\} \left(\frac{\Delta\Phi}{2n}\right)^p e^{-V}\,\mathrm{d}x.\end{aligned}$$

Since

$$\fint_{\partial B_1(0)} \left|\nabla^2\Phi(x) y\right|^2 \mathrm{d}y = \frac{\mathrm{Tr}[(\nabla^2\Phi(x))^2]}{n} \geq \frac{(\Delta\Phi(x))^2}{n^2},$$

it follows that

$$\alpha \int \left(\frac{\Delta\Phi}{2n}\right)^p e^{-V}\,\mathrm{d}x \geq \kappa \int \frac{(\Delta\Phi)^2}{n^2} \left(\frac{\Delta\Phi}{2n}\right)^p e^{-V}\,\mathrm{d}x,$$

and hence,

$$\int (\Delta\Phi)^p e^{-V} \geq \frac{1}{n^2}\frac{\kappa}{\alpha} \int (\Delta\Phi)^{p+2} e^{-V}\,\mathrm{d}x. \tag{2.34}$$

By Hôlder's inequality, with exponents $\frac{p+2}{2}, \frac{p+2}{p}$, we have

$$\int (\Delta\Phi)^p e^{-V} \leq \left(\int \left[(\Delta\Phi)^p\right]^{\frac{p+2}{p}} e^{-V}\right)^{\frac{p}{p+2}} \left(\int 1^{\frac{p}{p+2}} e^{-V}\right)^{\frac{p+2}{p}} = \left(\int (\Delta\Phi)^{p+2} e^{-V}\right)^{\frac{p}{p+2}}, \tag{2.35}$$

so combining (2.35) and (2.34) we get

$$\left(\int (\Delta\Phi)^{p+2} e^{-V}\right)^{\frac{p}{p+2}} \geq \frac{1}{n^2}\frac{\kappa}{\alpha} \int (\Delta\Phi)^{p+2} e^{-V}.$$

Relabeling $p \mapsto 2p$, the above can be written as

$$\|(\Delta\Phi)^2\|_{L^{p+1}(\mu)} \leq n^2 \frac{\alpha}{\kappa}. \tag{2.36}$$

Inequality (2.33) follows by taking $p \to \infty$,

$$\|\Delta\Phi\|_{L^\infty(\mu)} \leq n\sqrt{\frac{\alpha}{\kappa}}. \tag{2.37}$$

□

It remains to show that the second term in (2.31) is nonnegative, which was first observed by Kolesnikov in [29]. Here we essentially follow his proof with some minor adaptations to our case.

Proposition 2.8.

(2.38)

$$\int (\Delta_\epsilon\Phi(x))^p e^{-V(x)} \cdot \left\{ \fint_{\partial B_\epsilon(0)} \left\{ 2\langle \nabla W(\nabla\Phi(x)), \nabla\Phi(x+y) - \nabla\Phi(x)\rangle - \log\left[\frac{\det\nabla^2\Phi(x+y)\det\nabla^2\Phi(x-y)}{(\det\nabla^2\Phi(x))^2}\right] \right\} \mathrm{d}y \right\} \mathrm{d}x \geq 0.$$

Proof. Let $\nabla\Psi = (\nabla\Phi)^{-1}$ be the optimal transport map between $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ to $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$. The key estimate required to obtain (2.38) is contained in the following proposition.

Proposition 2.9.

$$
\begin{aligned}
&\int (\Delta_\epsilon \Phi(x))^p e^{-V(x)} \fint_{\partial B_\epsilon(0)} 2\langle \nabla W(\nabla\Phi(x)), \nabla\Phi(x+y) - \nabla\Phi(x)\rangle \,\mathrm{d}y \,\mathrm{d}x \\
&\geq 2 \fint_{\partial B_\epsilon(0)} \left\{ \int (\Delta_\epsilon \Phi(\nabla\Psi(x)))^p \left\{ \mathrm{Tr}\left[\nabla^2\Phi(\nabla\Psi(x)+y)\nabla^2\Psi(x)\right] - n \right\} \mathrm{d}\nu(x) \right\} \mathrm{d}y.
\end{aligned} \tag{2.39}
$$

Assuming the estimate (2.39) let us prove the estimate (2.38). Using the symmetry of $\partial B_\epsilon(0)$ under reflection $y \mapsto -y$ we have

$$
\begin{aligned}
&\int (\Delta_\epsilon \Phi)^p e^{-V} \\
&\cdot \left\{ \fint_{\partial B_\epsilon(0)} \left\{ 2\langle \nabla W(\nabla\Phi), \nabla\Phi(x+y) - \nabla\Phi(x)\rangle - \log\left[\frac{\det \nabla^2\Phi(x+y) \det \nabla^2\Phi(x-y)}{(\det \nabla^2\Phi(x))^2} \right] \right\} \mathrm{d}y \right\} \mathrm{d}x \\
&\overset{(2.39)}{\geq} 2 \fint_{\partial B_\epsilon(0)} \left\{ \int (\Delta_\epsilon \Phi(\nabla\Psi))^p \left\{ \mathrm{Tr}\left[\nabla^2\Phi(\nabla\Psi(x)+y)\nabla^2\Psi(x)\right] - n \right\} \mathrm{d}\nu \right\} \mathrm{d}y \\
&- 2 \int (\Delta_\epsilon \Phi)^p \fint_{\partial B_\epsilon(0)} \log\det\left[\nabla^2\Phi(x+y)(\nabla^2\Phi(x))^{-1}\right] \mathrm{d}y \, e^{-V} \,\mathrm{d}x \\
&= 2 \fint_{\partial B_\epsilon(0)} \left\{ \int \left\{ \mathrm{Tr}\left[\nabla^2\Phi(\nabla\Psi(x)+y)\nabla^2\Psi(x)\right] - n \right\} (\Delta_\epsilon \Phi(\nabla\Psi))^p \,\mathrm{d}\nu \right\} \mathrm{d}y \\
&- 2 \fint_{\partial B_\epsilon(0)} \left\{ \int \log\det\left[\nabla^2\Phi(\nabla\Psi(x)+y)(\nabla^2\Phi(\nabla\Psi(x)))^{-1}\right] (\Delta_\epsilon \Phi(\nabla\Psi))^p \,\mathrm{d}\nu \right\} \mathrm{d}y \\
&= 2 \fint_{\partial B_\epsilon(0)} \left\{ [\mathrm{Tr}[A(x)] - n - \log\det A(x)] \, (\Delta_\epsilon \Phi(\nabla\Psi))^p \,\mathrm{d}\nu \right\} \mathrm{d}y,
\end{aligned}
$$

where $A(x) = \nabla^2\Phi(\nabla\Psi(x)+y)(\nabla^2\Phi(\nabla\Psi(x)))^{-1}$. By [30, p. 8], $\mathrm{Tr}[A(x)] - n - \log\det A(x) \geq 0$ for every x, which completes the proof. □

It remains to prove Proposition 2.9.

Proof of Proposition 2.9. Let $\nabla\Psi = (\nabla\Phi)^{-1}$ be the optimal transport map between $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ to $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$. The estimate (2.39) will obtained as a consequence of integration by parts. But in order to justify the vanishing of the boundary terms in the integration by parts we need to work with cutoff functions. Let $\{\eta_k\}$ be a sequence of compactly supported smooth functions $\eta_k : \mathbb{R}^n \to \mathbb{R}$ such that

- $0 \leq \eta_k \leq 1$ for every k.
- $\lim_{k\to\infty} \eta_k = 1$ uniformly on every compact set.
- $\lim_{k\to\infty} \int |\nabla\eta_k(x)|^2 \,\mathrm{d}\nu(x) = 0$.

In the computations below we will simplify the notation and write

$$
\delta\Phi_y(x) := \nabla\Phi(x+y) - \nabla\Phi(x),
$$

omitting the dependence on x when it is clear from context, $\delta\Phi_y := \delta\Phi_y(x)$.

By the Monge-Ampère equation (2.21) (and since the integrals can be exchanged by the proof of Lemma 2.6), multiplying the integrand in the left-hand side of (2.39) by $\eta_k(\nabla\Phi(x))$, and integrating

against $e^{-V(x)}\,\mathrm{d}x$, yields

$$\begin{aligned}
&\int \left\{ (\Delta_\epsilon \Phi)^p \fint_{\partial B_\epsilon(0)} 2\langle \nabla W(\nabla\Phi), \delta\Phi_y\rangle \,\mathrm{d}y \right\} \eta_k(\nabla\Phi)\,\mathrm{d}\mu \\
&= \int \left\{ (\Delta_\epsilon \Phi)^p \fint_{\partial B_\epsilon(0)} 2\langle \nabla W(\nabla\Phi), \delta\Phi_y\rangle \,\mathrm{d}y \right\} \eta_k(\nabla\Phi) e^{-W(\nabla\Phi)} \det\nabla^2\Phi\,\mathrm{d}x \\
&= -\int \left\{ (\Delta_\epsilon \Phi)^p \fint_{\partial B_\epsilon(0)} 2\langle \nabla(e^{-W(\nabla\Phi)}), (\nabla^2\Phi)^{-1}\delta\Phi_y\rangle \,\mathrm{d}y \right\} \eta_k(\nabla\Phi) \det\nabla^2\Phi\,\mathrm{d}x \\
&= -\fint_{\partial B_\epsilon(0)} \left\{ \int 2\langle \nabla(e^{-W(\nabla\Phi)}), \mathrm{Cof}(\nabla^2\Phi)\delta\Phi_y\rangle (\Delta_\epsilon\Phi)^p \eta_k(\nabla\Phi)\,\mathrm{d}x \right\} \mathrm{d}y.
\end{aligned}$$

By integration by parts in the x variable (which has no boundary terms because η_k is compactly supported), and using the fact that the cofactor matrix is divergence free, we get

$$\begin{aligned}
&-\int \langle \nabla(e^{-W(\nabla\Phi)}), \mathrm{Cof}(\nabla^2\Phi)\delta\Phi_y\rangle (\Delta_\epsilon\Phi)^p \eta_k(\nabla\Phi) \\
&\quad = \int e^{-W(\nabla\Phi)} \,\mathrm{div}\left[\mathrm{Cof}(\nabla^2\Phi)\delta\Phi_y (\Delta_\epsilon\Phi)^p \eta_k(\nabla\Phi)\right] \\
&\quad = \int e^{-W(\nabla\Phi)} \,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi)\nabla[\delta\Phi_y (\Delta_\epsilon\Phi)^p \eta_k(\nabla\Phi)]\right] \\
&\quad = \int e^{-W(\nabla\Phi)} \,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi)\nabla[\delta\Phi_y (\Delta_\epsilon\Phi)^p]\right] \eta_k(\nabla\Phi) \\
&\qquad + \int e^{-W(\nabla\Phi)} \,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi)\left\{((\Delta_\epsilon\Phi)^p\delta\Phi_y)\otimes\nabla[\eta_k(\nabla\Phi)]\right\}\right],
\end{aligned}$$

where we use the convention $(\nabla u(x))_{ij} = \partial_i u_j(x)$ for a vector field u, and $(w\otimes v)_{ij} = w_j v_i$ for vectors v, w. Hence, our goal is to lower bound the term

$$\begin{aligned}
&\int \left\{ (\Delta_\epsilon \Phi)^p \fint_{\partial B_\epsilon(0)} 2\langle \nabla W(\nabla\Phi), \delta\Phi_y\rangle \,\mathrm{d}y \right\} \eta_k(\nabla\Phi)\,\mathrm{d}\mu \\
&= 2\fint_{\partial B_\epsilon(0)} \int e^{-W(\nabla\Phi)} \,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi)\nabla[\delta\Phi_y (\Delta_\epsilon\Phi)^p]\right] \eta_k(\nabla\Phi)\,\mathrm{d}x\,\mathrm{d}y \\
&\quad + 2\fint_{\partial B_\epsilon(0)} \int e^{-W(\nabla\Phi)} \,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi)\left\{((\Delta_\epsilon\Phi)^p\delta\Phi_y)\otimes\nabla[\eta_k(\nabla\Phi)]\right\}\right] \mathrm{d}x\,\mathrm{d}y.
\end{aligned} \tag{2.40}$$

The second term on the right-hand side of (2.40) will be shown to vanish as $k\to\infty$ so it suffices to lower bound the first term on the right-hand side of (2.40), and then take the limit $k\to\infty$. We start with the first term in (2.40).

Lemma 2.10.

$$\begin{aligned}
&\fint_{\partial B_\epsilon(0)} \int e^{-W(\nabla\Phi(x))} \,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi(x))\nabla[(\nabla\Phi(x+y)-\nabla\Phi(x))\,(\Delta_\epsilon\Phi(x))^p]\right] \eta_k(\nabla\Phi(x))\,\mathrm{d}x\,\mathrm{d}y \\
&\geq \fint_{\partial B_\epsilon(0)} \int \left\{\mathrm{Tr}\left[\nabla^2\Phi(\nabla\Psi(x)+y)\nabla^2\Psi(x)\right] - n\right\} (\Delta_\epsilon\Phi(\nabla\Psi(x)))^p \eta_k(x)\,\mathrm{d}\nu(x)\,\mathrm{d}y.
\end{aligned} \tag{2.41}$$

Proof. Let $\nabla\Psi = (\nabla\Phi)^{-1}$ be the optimal transport map between $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ to $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$, and recall $\nabla^2\Psi(x) = (\nabla^2\Phi)^{-1}(\nabla\Psi(x))$. Using the chain rule we have, for any vector field u in $\mathbb{R}^n$,

$\operatorname{div}_x[u(\nabla\Psi(x))] = \operatorname{Tr}[\nabla^2\Psi(x)\nabla u(\nabla\Psi(x))]$, so with

$$\delta\Phi_y := \nabla\Phi(x+y) - \nabla\Phi(x), \quad \delta\Phi_y(\nabla\Psi) := \delta\Phi_y(\nabla\Psi(x)) = \nabla\Phi(\nabla\Psi(x)+y) - \nabla\Phi(\nabla\Psi(x)),$$

we have

$$\begin{aligned}
&\int e^{-W(\nabla\Phi)} \operatorname{Tr}\left[\operatorname{Cof}(\nabla^2\Phi)\nabla[(\Delta_\epsilon\Phi)^p\delta\Phi_y]\right]\eta_k(\nabla\Phi)\,\mathrm{d}x = \int \operatorname{Tr}\left[\nabla^2\Psi(\nabla\Phi)\nabla[(\Delta_\epsilon\Phi)^p\delta\Phi_y]\right]\eta_k(\nabla\Phi)\,\mathrm{d}\mu\\
&= \int \operatorname{Tr}\left[\nabla^2\Psi\nabla[\delta\Phi(\nabla\Psi)(\Delta_\epsilon\Phi(\nabla\Psi))^p]\right]\eta_k\,\mathrm{d}\nu = \int \operatorname{div}[(\Delta_\epsilon\Phi(\nabla\Psi))^p\delta\Phi_y(\nabla\Psi)]\eta_k\,\mathrm{d}\nu\\
&= \int \operatorname{div}[(\nabla\Phi(\nabla\Psi(x)+y)-x)](\Delta_\epsilon\Phi(\nabla\Psi))^p\eta_k\,\mathrm{d}\nu + \int\langle\delta\Phi_y(\nabla\Psi), \nabla(\Delta_\epsilon\Phi(\nabla\Psi))^p\rangle\eta_k\,\mathrm{d}\nu.
\end{aligned}$$

Hence,

(2.42)
$$\begin{aligned}
&⨍_{\partial B_\epsilon(0)}\int e^{-W(\nabla\Phi)} \operatorname{Tr}\left[\operatorname{Cof}(\nabla^2\Phi)\nabla[(\Delta_\epsilon\Phi)^p\delta\Phi_y]\right]\eta_k(\nabla\Phi)\,\mathrm{d}x\,\mathrm{d}y\\
&= ⨍_{\partial B_\epsilon(0)}\int \operatorname{div}[(\nabla\Phi(\nabla\Psi(x)+y)-x)](\Delta_\epsilon\Phi(\nabla\Psi))^p\eta_k\,\mathrm{d}\nu\,\mathrm{d}y + ⨍_{\partial B_\epsilon(0)}\int\langle\delta\Phi_y(\nabla\Psi), \nabla(\Delta_\epsilon\Phi(\nabla\Psi))^p\rangle\eta_k\,\mathrm{d}\nu\,\mathrm{d}y.
\end{aligned}$$

Let us bound the two terms in (2.42) separately. For the first term in (2.42) we use the chain rule to get

$$\operatorname{div}[\nabla\Phi(\nabla\Psi(x)+y)-x] = \operatorname{div}[\nabla\Phi(\nabla\Psi(x)+y)] - n \geq \operatorname{Tr}\left[\nabla^2\Phi(\nabla\Psi(x)+y)\nabla^2\Psi(x)\right] - n.$$

Using $\nabla^2\Psi(x) = (\nabla^2\Phi)^{-1}(\nabla\Psi(x))$ gives

(2.43)
$$\begin{aligned}
&⨍_{\partial B_\epsilon(0)}\int \operatorname{div}[(\nabla\Phi(\nabla\Psi(x)+y)-x)](\Delta_\epsilon\Phi(\nabla\Psi))^p\eta_k\,\mathrm{d}\nu\,\mathrm{d}y\\
&\geq ⨍_{\partial B_\epsilon(0)}\int \left\{\operatorname{Tr}\left[\nabla^2\Phi(\nabla\Psi(x)+y)\nabla^2\Psi(x)\right] - n\right\}(\Delta_\epsilon\Phi(\nabla\Psi))^p\eta_k\,\mathrm{d}\nu\,\mathrm{d}y.
\end{aligned}$$

For the second term in (2.42) we use the definition of Δ_ϵ to write

(2.44)
$$\begin{aligned}
&⨍_{\partial B_\epsilon(0)}\int \langle\nabla\delta\Phi_y(\nabla\Psi), \nabla\left((\Delta_\epsilon\Phi(\nabla\Psi))^p\right)\rangle\,\eta_k\,\mathrm{d}\nu\,\mathrm{d}y\\
&= p⨍_{\partial B_\epsilon(0)}\int (\Delta_\epsilon\Phi(\nabla\Psi))^{p-1}\left\langle\delta\Phi_y(\nabla\Psi), \nabla^2\Psi\nabla\Delta_\epsilon\Phi(\nabla\Psi)\right\rangle\eta_k\,\mathrm{d}\nu\,\mathrm{d}y\\
&= p\int (\Delta_\epsilon\Phi(\nabla\Psi))^{p-1}\left\langle⨍_{\partial B_\epsilon(0)}\delta\Phi_y(\nabla\Psi)\,\mathrm{d}y, \nabla^2\Psi\nabla\Delta_\epsilon\Phi(\nabla\Psi)\right\rangle\eta_k\,\mathrm{d}\nu\\
&\overset{(2.25)}{=} p\int (\Delta_\epsilon\Phi(\nabla\Psi))^{p-1}\left\langle(\Delta_\epsilon\nabla\Phi)(\nabla\Psi), \nabla^2\Psi\nabla\Delta_\epsilon\Phi(\nabla\Psi)\right\rangle\eta_k\,\mathrm{d}\nu\\
&= p\int (\Delta_\epsilon\Phi(\nabla\Psi))^{p-1}\left\langle\nabla\Delta_\epsilon\Phi(\nabla\Psi), \nabla^2\Psi\nabla\Delta_\epsilon\Phi(\nabla\Psi)\right\rangle\eta_k\,\mathrm{d}\nu \geq 0,
\end{aligned}$$

where we used the convexity of Φ and Ψ as well as $\eta_k \geq 0$. Combining (2.43) and (2.44) yields (2.41). □

We now turn to the second term in (2.40).

Lemma 2.11.

(2.45)

$$2\int e^{-W(\nabla\Phi(x))}\,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi(x))\left\{\left((\nabla\Phi(x+y)-\nabla\Phi(x))\,(\Delta_\epsilon\Phi(x))^p\right)\otimes\nabla[\eta_k(\nabla\Phi(x))]\right\}\right]\mathrm{d}x$$
$$\leq 2\left(\int\left|\left(\nabla\Phi(\nabla\Psi(x)+y)-\nabla\Phi(\nabla\Psi(x))\right)(\Delta_\epsilon\Phi(\nabla\Psi(x)))^p\right|^2\mathrm{d}\nu(x)\right)^{\frac{1}{2}}\left(\int|\nabla\eta_k(x)|^2\,\mathrm{d}\nu(x)\right)^{\frac{1}{2}}.$$

Proof. Given a symmetric matrix M and vectors u, v we have $\mathrm{Tr}[M^{-1}\{u\otimes(Mv)\}] = \langle u, v\rangle$, so applying this identity we get, with $\delta\Phi_y := \nabla\Phi(x+y)-\nabla\Phi(x)$,

$$\int e^{-W(\nabla\Phi)}\,\mathrm{Tr}\left[\mathrm{Cof}(\nabla^2\Phi)\left\{(\delta\Phi_y(\Delta_\epsilon\Phi)^p)\otimes\nabla[\eta_k(\nabla\Phi)]\right\}\right]\mathrm{d}x$$
$$=\int\mathrm{Tr}\left[\left\{[\delta\Phi(\Delta_\epsilon\Phi)^p]\otimes[(\nabla\eta_k)(\nabla\Phi)]\right\}\right]\mathrm{d}\mu=\int\mathrm{Tr}\left[\left\{[\delta\Phi_y(\nabla\Psi)(\Delta_\epsilon\Phi(\nabla\Psi))^p]\right\}\otimes\nabla\eta_k\right]\mathrm{d}\nu$$
$$\leq\left(\int|\delta\Phi_y(\nabla\Psi)(\Delta_\epsilon\Phi(\nabla\Psi))^p|^2\,\mathrm{d}\nu\right)^{\frac{1}{2}}\left(\int|\nabla\eta_k|^2\,\mathrm{d}\nu\right)^{\frac{1}{2}}.$$

□

Let us now take the limits $k\to\infty$ of the two terms in (2.40). For the first term in (2.40) we first use (2.41) and then take the limit $k\to\infty$ in (2.41). We can move the limit past the integrals in (2.41) by dominated convergence theorem using $\eta_k\leq 1$, and the fact that the rest of the integrand is integrable (as in the proof of Lemma 2.6). Thus, if the second term in (2.40) vanishes in the limit $k\to\infty$ we will get (2.39). To show that the second term in (2.40) vanishes in the limit $k\to\infty$ it suffices to show that the first term on the right-hand side of (2.45) is finite since by assumption $\lim_{k\to\infty}\int|\nabla\eta_k(x)|^2\,\mathrm{d}\nu(x)=0$. The first term on the right-hand side of (2.45) is finite by the same argument as in the proof of Lemma 2.6. □

To summarize, the combination of Lemma 2.10 and Lemma 2.11 yields Proposition 2.9, which in turns implies Proposition 2.8. It follows that the second term in (2.31) is nonnegative and the proof of Theorem 2.1 is complete by Lemma 2.7. □

Remark 2.12. Our proof of Theorem 2.1 follows the proof of [30, Theorem 6.2]. However, one important modification we need to make is the introduction of the Δ_ϵ operator. In [30], the lack of regularity of V and $\nabla\Phi$ is remedied by considering finite differences (rather than actual derivatives), $V(x+e)-V(x)$ and $\nabla\Phi(x+e)-\nabla\Phi(x)$, for vectors $e\in\mathbb{R}^n$. The finite difference $V(x+e)-V(x)$ is then controlled by explicit assumptions on the second directional derivatives of V, which in turn leads to control on the second directional derivatives of Φ. In contrast, we only have at our disposal the control of ΔV, so the finite differences approach just outlined does not work. The operator Δ_ϵ is the appropriate replacement to the finite differences scheme.

3. Majorization

Theorem 1.9 and Theorem 1.10 were established in Section 1.2, so with regard to majorization it remains to prove Theorem 1.11 and Theorem 1.12. We begin with Theorem 1.11.

Theorem 3.1 (Monotonicity along Wasserstein geodesics)**.** *Let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ and $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ be probability measures on $\mathbb{R}^n$, with μ supported on all of $\mathbb{R}^n$, such that there exist $\alpha>0,\kappa>0$ with*

$$\Delta V\leq\alpha n\quad\text{and}\quad\nabla^2 W\succeq\kappa\,\mathrm{Id}_n\,.$$

Let $\nabla\Phi:\mathbb{R}^n\to\mathbb{R}^n$ be the Brenier map transporting μ to ν, and let $(\rho_t)_{t\in[0,1]}$ be the geodesic in Wasserstein space connecting μ and ν, $\rho_t := (\nabla\Phi_t)_\sharp\mu$ where $\nabla\Phi_t(x) := (1-t)x+t\nabla\Phi(x)$. Then,

if $\frac{\alpha}{\kappa} \le 1$,

$$[0,1] \ni t \mapsto \int_{\mathbb{R}^n} \varphi(\rho_t(x))\,\mathrm{d}x \text{ is monotonically non-decreasing} \tag{3.1}$$

for every convex function $\varphi : \mathbb{R}_{\ge 0} \to \mathbb{R}$.

Proof. Let $\nabla\Psi := (\nabla\Phi)^{-1}$ be Brenier map between ν and μ, and let $\nabla\Psi_t(x) := (1-t)\nabla\Psi(x) + tx$ be the Brenier map between $\nu = \rho_1$ to ρ_t. Let us show that the Brenier map $(\nabla\Psi_t)^{-1}(x) = [(1-t)\nabla\Psi(x) + tx]^{-1}$ between ρ_t and $\rho_1 = \nu$ satisfies $\Delta[(\nabla\Psi_t)^{-1}] \le n$. Indeed, the eigenvalues of $\nabla[(\nabla\Psi_t)^{-1}(x)]$ are

$$\left\{ \frac{\lambda_i(x)}{1 + t(\lambda_i(x) - 1)} \right\}_{i=1}^n,$$

where $\{\lambda_i(x)\}_{i=1}^n$ are the eigenvalues of $\nabla^2\Phi(x)$. Given a fixed $x \in \mathbb{R}^n$ define the function $\theta : [0,1] \to \mathbb{R}$ by $\theta(t) := \sum_{i=1}^n \frac{\lambda_i(x)}{1+t(\lambda_i(x)-1)}$. The function θ is concave and satisfies $\theta(0) \le n$, by (2.13), and $\theta(1) = n$. It follows that $\theta(t) \le n$ for all $t \in [0,1]$, i.e., $\Delta[(\nabla\Psi_t)^{-1}] \le n$. In particular, using $\frac{\alpha}{\kappa} \le 1$, $\det[(\nabla\Psi_t)^{-1}] \le 1$ so $\rho_1 = \nu$ majorizes ρ_t.

Moreover, setting $\nabla\Phi_s(x) := (1-s)x + s\nabla\Phi(x)$ to be the Brenier map between ρ_0 and ρ_s we clearly have that $\Delta\Phi_s \le n$ and so $\rho_0 = \mu$ is majorized by ρ_s. To show that for $0 < r \le s$ the measure ρ_r is majorized by ρ_s, we can apply the first step with ρ_1 replaced by ρ_s and Φ replaced by Φ_s. Notice indeed that the only property we have used of Φ is that $\Delta\Phi \le n$, which is true for Φ_s as well. □

Let us now prove Theorem 1.12.

Theorem 3.2. *Let* $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ *and* $\mathrm{d}\nu = e^{-W}\,\mathrm{d}x$ *be probability measures on* $\mathbb{R}^n$*, with* μ *supported on all of* $\mathbb{R}^n$*, such that there exist* $\alpha > 0, \kappa > 0$ *with*

$$\Delta V \le \alpha n \quad \text{and} \quad \nabla^2 W \succeq \kappa\,\mathrm{Id}_n\,.$$

Let $\nabla\Phi : \mathbb{R}^n \to \mathbb{R}^n$ *be the Brenier map transporting* μ *to* ν*. If* $\frac{\alpha}{\kappa} \le 1$*, then*

$$\mathrm{H}(\nu) - \mathrm{H}(\mu) \ge \frac{1}{2n^2} \int_{\mathbb{R}^n} \|\nabla^2\Phi - \mathrm{Id}_n\|_F^2\,\mathrm{d}\mu, \tag{3.2}$$

where $\|\cdot\|_F$ *is the Frobenius norm. In particular,*

$$\mathrm{H}(\nu) = \mathrm{H}(\mu) \qquad \Longrightarrow \qquad \mu = \nu. \tag{3.3}$$

Proof. By (2.13), and the assumption $\frac{\alpha}{\kappa} \le 1$,

$$\|\Delta\Phi\|_{L^\infty(\mathrm{d}x)} \le n\sqrt{\frac{\alpha}{\kappa}} \le n.$$

Abusing notation we identify between the measures μ, ν and their densities, so the Monge-Ampère equation reads

$$\mu = \nu(\nabla\Phi)\det\nabla^2\Phi.$$

By the change of variables formula,

$$\begin{aligned}\mathrm{H}(\mu) &= \int \left(\nu(\nabla\Phi)\det\nabla^2\Phi\right)\log\nu(\nabla\Phi) + \int \left(\nu(\nabla\Phi)\det\nabla^2\Phi\right)\log\det\nabla^2\Phi \\ &= \mathrm{H}(\nu) + \int \log\det\left(\nabla^2\Phi \circ (\nabla\Phi)^{-1}\right)\mathrm{d}\nu = \mathrm{H}(\nu) + \int \log\det\left(\nabla^2\Phi\right)\mathrm{d}\mu,\end{aligned}$$

so

$$\mathrm{H}(\nu) - \mathrm{H}(\mu) = \sum_{k=1}^n \int [-\log\lambda_k(x)]\,\mathrm{d}\mu(x),$$

where $0 \le \lambda_1(x) \le \cdots \le \lambda_n(x)$ are the eigenvalues of $\nabla^2\Phi(x)$. Using

$$-\log s \ge -\log t + \frac{t-s}{t} + \frac{(s-t)^2}{2\max\{s,t\}^2}, \quad s,t \in (0,\infty),$$

(see [14, Proof of Lemma 2.5]), with $s = \lambda_k(x)$ and $t = 1$, we get

$$\begin{aligned}
\mathrm{H}(\nu) - \mathrm{H}(\mu) &\ge \sum_{k=1}^n \int \left\{ (1-\lambda_k(x)) + \frac{1}{2}\frac{(\lambda_k-1)^2}{\max\{\lambda_n,1\}^2} \right\} \mathrm{d}\mu(x) \\
&\ge \int [n - \Delta\Phi]\,\mathrm{d}\mu + \frac{1}{2n^2}\sum_{k=1}^n \int (\lambda_k(x)-1)^2\,\mathrm{d}\mu(x) \\
&\ge \frac{1}{2n^2}\int \|\nabla^2\Phi(x) - \mathrm{Id}_n\|_{\mathrm{F}}^2\,\mathrm{d}\mu(x),
\end{aligned}$$

where we used $\Delta\Phi \le n \Rightarrow \lambda_n \le n$ in the second inequality, and $\Delta\Phi \le n$ in the last inequality. This establishes (3.2). Equation (3.3) follows from (3.2) since

$$\mathrm{H}(\nu) = \mathrm{H}(\mu) \quad \Longrightarrow \quad \nabla^2\Phi(x) = \mathrm{Id} \quad \text{for almost-eveywhere } x,$$

which implies that $\nabla\Phi$ is the identity map. □

4. The Kim-Milman heat flow transport map

When the target measure ν is the standard Gaussian measure γ we can construct a volume-contracting map based on the heat flow map of Kim and E. Milman [26] (following Otto and Villani [39]). While this transport map exists outside the Gaussian setting [26, 13], our techniques to establish volume contraction are restricted to Gaussian targets. In addition, in contrast to the Brenier map, we can only establish volume contraction rather than control of the divergence of the map. Finally, to establish the existence of the heat flow map we require further regularity than those for the optimal transport map. For these reasons, we will keep this section brief and assume whatever regularity is needed.

We recall the definition of the Orenstein-Uhlenbeck semigroup (P_t),

$$\mathrm{P}_t f(x) := \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\,\mathrm{d}\gamma(y), \quad t \ge 0, \quad x \in \mathbb{R}^n, \quad f \in L^1(\gamma). \tag{4.1}$$

The construction of the heat flow map between a source measure μ and a Gaussian target γ is based on the following flow.

Proposition 4.1. *Let ρ be an absolutely continuous probability measure on $\mathbb{R}^n$, and let $f := \frac{\mathrm{d}\rho}{\mathrm{d}\gamma}$. Suppose the ordinary differential equation*

$$\begin{cases} \partial_t \mathrm{F}_t(x) = -\nabla \log \mathrm{P}_t f(\mathrm{F}_t(x)), & \text{for } t > 0 \text{ and } x \in \mathbb{R}^n, \\ \mathrm{F}_0(x) = x, & \text{for } x \in \mathbb{R}^n, \end{cases} \tag{4.2}$$

has a unique smooth solution. Then, the probability measures $\rho_t := (\mathrm{F}_t)_\sharp \rho$ satisfy

$$\mathrm{d}\rho_t(x) = \mathrm{P}_t f(x)\,\mathrm{d}\gamma(x), \quad \text{for } t \ge 0 \text{ and } x \in \mathbb{R}^n. \tag{4.3}$$

The identity (4.3) follows from the standard switch between the Eulerian and Lagrangian perspectives, as well as the partial differential equation satisfied by $(t,x) \mapsto \mathrm{P}_t f(x)$ (cf. [26, §1.2]). The key point is that as $t \to \infty$, $\mathrm{P}_t f \to 1$, so $\rho_t \to \gamma$. In particular, whenever $\mathrm{F} := \lim_{t\to\infty} \mathrm{F}_t$ exists, we find that F transports ρ to γ. The map F is the *Kim-Milman heat flow map*. The main result of this section is that, when the target measure is Gaussian, the Kim-Milman heat flow map also achieves the bound (1.6).

Theorem 4.2. *Let $\mathrm{d}\mu = e^{-V}\,\mathrm{d}x$ be a probability measure on $\mathbb{R}^n$ such that there exists $\alpha > 0$ with*

$$\Delta V(x) \le \alpha n \quad \text{for every } x \in \mathbb{R}^n.$$

Suppose that the flow (4.2), with $f := \frac{\mathrm{d}\mu}{\mathrm{d}\gamma}$, converges to a differentiable limit $\mathrm{F} := \lim_{t\to\infty} \mathrm{F}_t$, and that $\lim_{t\to\infty} \det \nabla \mathrm{F}_t = \det \nabla \mathrm{F}$. Then, the Kim-Milman heat flow map between μ and the standard Gaussian γ on $\mathbb{R}^n$ satisfy

$$\| \det \nabla \mathrm{F} \|_{L^\infty(\mathrm{d}x)} \le \alpha^{\frac{n}{2}}. \tag{4.4}$$

The observation behind the proof of Theorem 4.2 is that $\det \nabla \mathrm{F}_t$ can be controlled if $\Delta \log \mathrm{P}_t f$ can be controlled. On other hand, by Proposition 2.4(ii), $\Delta \log \mathrm{P}_t f$ can be lower bounded provided that $\Delta \log f$ can be lower bounded. The next result implements these observations.

Proposition 4.3. *Fix $c \ge 1$ and let $\mathrm{d}\rho = f\,\mathrm{d}\gamma$ be a $(-cn)$-log-subharmonic measure. Fix $t \ge 0$ and let F_t be the heat flow map from ρ to $(\mathrm{P}_t f)\gamma$. Then, for all $x \in \mathbb{R}^n$ and $t \ge 0$,*

$$|\det \nabla \mathrm{F}_t(x)| \le \left[(1 - e^{-2t})(c-1) + 1\right]^{\frac{n}{2}}, \quad \text{for every } x \in \mathbb{R}^n.$$

Proof. We start by recalling the Jacobi formula. Let $(M_t)_{t\ge 0}$ be a family of invertible $n \times n$ matrices such that the map $t \mapsto \det M_t$ is differentiable. Then,

$$\partial_t \det M_t = \mathrm{Tr}\left[M_t^{-1} \partial_t M_t\right] \det M_t. \tag{4.5}$$

In order to apply (4.5) we first need to derive the evolution equation for $(\nabla \mathrm{F}_t)$. The evolution of (F_t) is determined by the equation

$$\partial_t \mathrm{F}_t(x) = -\nabla \log \mathrm{P}_t f(\mathrm{F}_t(x)), \quad \mathrm{F}_0(x) = x, \quad \forall\, x \in \mathbb{R}^n, \tag{4.6}$$

so differentiating (4.6) with respect to x yields the evolution equation of $(\nabla \mathrm{F}_t)$,

$$\partial_t \nabla \mathrm{F}_t(x) = -\nabla^2 \log \mathrm{P}_t f(\mathrm{F}_t(x)) \nabla \mathrm{F}_t(x), \quad \nabla \mathrm{F}_0 = \mathrm{Id}, \quad \forall\, x \in \mathbb{R}^n. \tag{4.7}$$

Hence, by the Jacobi formula (4.5) and the cyclic property of the trace,

$$\begin{aligned}
\partial_t \det \nabla \mathrm{F}_t(x) &= \mathrm{Tr}\left[(\nabla \mathrm{F}_t(x))^{-1} \partial_t \nabla \mathrm{F}_t(x)\right] \det \nabla \mathrm{F}_t(x) \\
&= \mathrm{Tr}\left[(\nabla \mathrm{F}_t(x))^{-1} \left\{-\nabla^2 \log \mathrm{P}_t f(\mathrm{F}_t(x)) \nabla \mathrm{F}_t(x)\right\}\right] \det \nabla \mathrm{F}_t(x) \\
&= \left[-\Delta \log \mathrm{P}_t f(\mathrm{F}_t(x))\right] \det \nabla \mathrm{F}_t(x).
\end{aligned}$$

We conclude that

$$\partial_t \det \nabla \mathrm{F}_t(x) = \left[-\Delta \log \mathrm{P}_t f(\mathrm{F}_t(x))\right] \det \nabla \mathrm{F}_t(x), \quad \det \nabla \mathrm{F}_0 = 1, \quad \forall\, x \in \mathbb{R}^n. \tag{4.8}$$

Since ρ is a $(-cn)$-log-subharmonic measure it follows that f is a $-(c-1)n$-log-subharmonic function. Hence, by Proposition 2.4(iii),

$$-\Delta \log \mathrm{P}_t f(\mathrm{F}_t(x)) \le \frac{n(c-1)e^{-2t}}{1 + (1 - e^{-2t})(c-1)},$$

and it follows from Grönwall's inequality that

$$\det \nabla \mathrm{F}_t(x) \le e^{\int_0^t \frac{n(c-1)e^{-2s}}{1+(1-e^{-2s})(c-1)}\,\mathrm{d}s}.$$

To conclude the proof note that, for every $x \in \mathbb{R}^n$, $\det \nabla \mathrm{F}_0(x) = 1$ and $\det \nabla \mathrm{F}_t(x) \ne 0$ for every $t \ge 0$ since F_t is a diffeomorphism. Hence, for every $x \in \mathbb{R}^n$ and $t \ge 0$, $\det \nabla \mathrm{F}_t(x) > 0$, so

$$|\det \nabla \mathrm{F}_t(x)| = \det \nabla \mathrm{F}_t(x) \le e^{\int_0^t \frac{n(c-1)e^{-2s}}{1+(1-e^{-2s})(c-1)}\,\mathrm{d}s}.$$

The indefinite integral of the integrand inside the exponential is $\log[1 + (c-1)(1 - e^{-2t})]^{n/2}$, which completes the proof. □

Proof of Theorem 4.2. By Proposition 4.3 and the assumption on V,

$$|\det \nabla \mathrm{F}_t(x)| \le \left[(1-e^{-2t})(\alpha-1)+1\right]^{\frac{n}{2}}. \tag{4.9}$$

Taking $t \to \infty$ yields

$$|\det \nabla \mathrm{F}(x)| \le \alpha^{\frac{n}{2}}, \quad \text{for all } x \in \mathbb{R}^n.$$

□

5. Appendix

5.1. **The operator Δ_ϵ.** In this section we prove some of the properties of the operator Δ_ϵ. In particular, let us prove Lemma 2.5.

Lemma 5.1. *Fix $\epsilon > 0$ and for a function $f : \mathbb{R}^n \to \mathbb{R}$ denote*

$$\Delta_\epsilon f(x) := \fint_{\partial B_\epsilon(0)} [f(x+y) - f(x)]\, \mathrm{d}y,$$

where

$$\fint_{\partial B_\epsilon(0)} f := \frac{1}{|\partial B_\epsilon(0)|} \int_{\partial B_\epsilon(0)} f.$$

Then,

$$\lim_{\epsilon \to 0} \frac{\Delta_\epsilon f(x)}{\epsilon^2} = \frac{\Delta f(x)}{2n}, \tag{5.1}$$

where Δf is the distributional Laplacian of f. Further, if $\Delta f \le \ell$, then

$$\Delta_\epsilon f \le \frac{\ell}{n}\frac{\epsilon^2}{2}, \qquad \forall\ \epsilon > 0. \tag{5.2}$$

Proof. To prove (5.1) first note that letting $\nabla^2 f$ be the distributional derivative of f we have

$$\lim_{\epsilon \to 0} \frac{f(x+\epsilon y) + f(x-\epsilon y) - 2f(x)}{\epsilon^2} = \langle \nabla^2 f(x) y, y\rangle,$$

in the sense of distributions, where the distribution $\langle \nabla^2 f(x) y, y\rangle$ is defined as

$$\big\langle \langle \nabla^2 f(x) y, y\rangle, \eta \big\rangle = \int f(x) \langle \nabla^2 \eta(x) y, y\rangle \, \mathrm{d}x,$$

for test functions η. Hence, using the symmetry of $\partial B_\epsilon(0)$ under reflection, and changing variables $y \mapsto \epsilon y$,

$$\begin{aligned}
\lim_{\epsilon \to 0} \frac{\Delta_\epsilon f(x)}{\epsilon^2} &= \lim_{\epsilon \to 0} \frac{1}{2\epsilon^2} \left\{ \fint_{\partial B_\epsilon(0)} 2[f(x+y) - f(x)]\, \mathrm{d}y \right\} \\
&= \frac{1}{2} \lim_{\epsilon \to 0} \left\{ \fint_{\partial B_\epsilon(0)} \frac{f(x+y) + f(x-y) - 2f(x)}{\epsilon^2}\, \mathrm{d}y \right\} \\
&= \frac{1}{2} \lim_{\epsilon \to 0} \left\{ \frac{1}{\epsilon^{n-1}} \fint_{\partial B_1(0)} \frac{f(x+\epsilon y) + f(x-\epsilon y) - 2f(x)}{\epsilon^2} \epsilon^{n-1}\, \mathrm{d}y \right\} \\
&= \frac{1}{2} \fint_{\partial B_1(0)} \langle \nabla^2 f(x) y, y\rangle \, \mathrm{d}y.
\end{aligned}$$

For each $i, j \in [n]$, by the symmetry of $\partial B_\epsilon(0)$ under $y_j \mapsto -y_j$,

$$\fint_{\partial B_1(0)} y_i y_j \, \mathrm{d}y = \delta_{ij} \fint_{\partial B_1(0)} y_i^2 \, \mathrm{d}y = \frac{1}{n}.$$

It follows that

$$\lim_{\epsilon\to 0}\frac{\Delta_\epsilon f(x)}{\epsilon^2} = \frac{1}{2}\sum_{i,j=1}^n \fint_{\partial B_1(0)} \partial^2_{ij} f(x) y_i y_j \,\mathrm{d}y = \frac{1}{2}\sum_{i=1}^n \partial^2_{ii} f(x) \fint_{\partial B_1(0)} y_i^2 \,\mathrm{d}y$$
$$= \frac{1}{2n}\sum_{i=1}^n \partial^2_{ii} f(x) = \frac{\Delta f(x)}{2n}.$$

Next we move to the proof of (5.2). We have

$$\Delta_\epsilon f(x) = \int_0^\epsilon \frac{\mathrm{d}}{\mathrm{d}r}\left\{\fint_{\partial B_r(0)} [f(x+y)-f(x)]\,\mathrm{d}y\right\}\mathrm{d}r$$
$$= \int_0^\epsilon \frac{\mathrm{d}}{\mathrm{d}r}\left[\fint_{\partial B_r(0)} f(x+y)\,\mathrm{d}y\right]\mathrm{d}r.$$

Since, by the change of variables $y \mapsto ry$,

$$\fint_{\partial B_r(0)} f(x+y)\,\mathrm{d}y = \fint_{\partial B_1(0)} f(x+ry)\,\mathrm{d}y,$$

we have, by integration by parts,

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\fint_{\partial B_r(0)} f(x+y)\,\mathrm{d}y\right] = \fint_{\partial B_1(0)} \frac{\mathrm{d}}{\mathrm{d}r} f(x+ry)\,\mathrm{d}y = \fint_{\partial B_1(0)} \langle \nabla f(x+ry), y\rangle \,\mathrm{d}y$$
$$= \fint_{\partial B_1(0)} \frac{1}{r}\langle \nabla_y [f(x+ry)], y\rangle \,\mathrm{d}y = \fint_{B_1(0)} \frac{1}{r}\Delta_y [f(x+ry)]\,\mathrm{d}y$$
$$= \frac{r}{n}\fint_{B_1(0)} \Delta f(x+ry)\,\mathrm{d}y \le \frac{\ell r}{n},$$

where the last inequality used $\Delta f \le \ell$. It follows that

$$\Delta_\epsilon f(x) \le \int_0^\epsilon r\,\mathrm{d}r = \frac{\ell}{n}\frac{\epsilon^2}{2}.$$

□

5.2. **Smoothing under Ornstein-Uhlenbeck and heat semigroups.** In this section we discuss Proposition 2.4. Since most of the results in the proposition are standard, when relevant we will only sketch the arguments and give references for detailed proofs. For the sake of completeness we will prove some additional smoothing properties beyond those stated in Proposition 2.4.

Proposition 5.2. *Let γ be the standard Gaussian measure on $\mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathbb{R}_{\ge 0}$ be a nonnegative function in $L^1(\gamma)$. Let $(\mathrm{P}_t)_{t\ge 0}$ be the Orenstein-Uhlenbeck semigroup,*

$$\mathrm{P}_t f(x) := \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\,\mathrm{d}\gamma(y), \quad t \ge 0, \quad x \in \mathbb{R}^n, \tag{5.3}$$

and let $(\mathrm{H}_t)_{t\ge 0}$ be the heat semigroup

$$\mathrm{H}_t f(x) := \int_{\mathbb{R}^n} f(x+\sqrt{t}y)\,\mathrm{d}\gamma(y), \quad t \ge 0, \quad x \in \mathbb{R}^n. \tag{5.4}$$

(1) For any $x \in \mathbb{R}^n$ and $t > 0$,

$$\nabla^2 \log \mathrm{P}_t f(x) \succeq -\frac{e^{-2t}}{1-e^{-2t}}\,\mathrm{Id}_n \quad \text{and} \quad \nabla^2 \log \mathrm{H}_t f(x) \succeq -\frac{1}{t}\,\mathrm{Id}_n\,. \tag{5.5}$$

(2) Suppose that f is c-log-concave for some $c \in \mathbb{R}$ (i.e., $x \mapsto \log f(x) - c\frac{|x|^2}{2}$ is concave). Then, for every $x \in \mathbb{R}^n$,

$$\nabla^2 \log \mathrm{P}_t f(x) \preceq \frac{e^{-2t}c}{1-c(1-e^{-2t})} \mathrm{Id}_n \begin{cases} \text{for any } t \in [0,\infty) & \text{if } c \le 1 \\ \text{for any } t \in \left[0, \log\left(\sqrt{\frac{c}{c-1}}\right)\right] & \text{if } c > 1 \end{cases}, \tag{5.6}$$

and

$$\nabla^2 \log \mathrm{H}_t f(x) \preceq \frac{c}{1-ct} \mathrm{Id}_n \begin{cases} \text{for any } t \in [0,\infty) & \text{if } c \le 0 \\ \text{for any } t \in \left[0, \frac{1}{c}\right] & \text{if } c > 0 \end{cases}. \tag{5.7}$$

(3) Suppose that f is c-log-convex for some $c \le 0$ (i.e., $x \mapsto \log f(x) - c\frac{|x|^2}{2}$ is convex). Then, for every $x \in \mathbb{R}^n$ and $t > 0$,

$$\nabla^2 \log \mathrm{P}_t f(x) \succeq \frac{e^{-2t}c}{1-c(1-e^{-2t})} \mathrm{Id}_n \quad \text{and} \quad \nabla^2 \log \mathrm{H}_t f(x) \succeq \frac{c}{1-ct} \mathrm{Id}_n \,. \tag{5.8}$$

Suppose f is cn-log-subharmonic for some $c \le 0$ (i.e., $x \mapsto \log f(x) - c\frac{|x|^2}{2}$ is subharmonic). Then, for every $x \in \mathbb{R}^n$ and $t > 0$,

$$\Delta \log \mathrm{P}_t f(x) \ge \frac{e^{-2t}cn}{1-c(1-e^{-2t})} \quad \text{and} \quad \Delta \log \mathrm{H}_t f(x) \ge \frac{cn}{1-ct}. \tag{5.9}$$

Remark 5.3. It was shown in [3, Remark] that there exists a 0-log-superharmonic f such that, for any $t > 0$, $\mathrm{P}_t f$ is strictly log-subharmonic. Concretely, let $n = 2$ and $f(x) = f(x_1, x_2) := e^{x_1 x_2}$. Then, $\Delta \log f(x) = 0$ for all $x \in \mathbb{R}^n$ while $\Delta \log \mathrm{P}_t f(x) > 0$ for all $x \in \mathbb{R}^n$ and $t > 0$. We conclude that log-superharmonic functions are not preserved under the Orenstein-Uhlenbeck/heat semigroups.

Proof. We will present a number of proofs for the various parts in Proposition 5.2. Since the Ornstein-Uhlenbeck semigroup and heat semigroup are related by the change of variables

$$\mathrm{P}_t f(x) = \mathrm{H}_{1-e^{-2t}} f(e^{-t}x), \tag{5.10}$$

we will use whichever one is convenient for the specific proof. The most robust proof technique (which works also on manifolds) is due to Hamilton [24, 1] based on deriving a partial differential equation for $M(t,x) := \nabla^2 \log \mathrm{H}_t f(x)$, and then using the maximum principle. Standard computation (e.g. [37, Proposition 2.11]) shows that

$$\partial_t M(t,x) = \frac{1}{2}\Delta M(t,x) + L(M(t,x)) + M^2(t,x), \tag{5.11}$$

where $L(M(t,x))$ is a *linear* differential operator in $M(t,x)$. Below we will apply the maximum principle to $(t,x) \mapsto M(t,x)$ to deduce Proposition 5.2(1-3).

Another proof technique, which is more probabilistic in nature, uses the specific from of the Ornstein-Uhlenbeck/heat semigroup in Euclidean space. Specifically, it is based on the following covariance identity

$$\nabla^2 \log \mathrm{P}_t f(x) = \frac{e^{-2t}}{1-e^{-2t}} \left(\frac{\mathrm{Cov}\left[p_{e^{-t}x, 1-e^{-2t}}\right]}{1-e^{-2t}} - \mathrm{Id}_n \right), \quad \forall\, x \in \mathbb{R}^n,\ \forall\, t > 0, \tag{5.12}$$

where $\mathrm{Cov}\,[p_{z,s}]$ is the covariance matrix of the probability measure

$$\mathrm{d}p_{z,s}(y) \propto f(y) e^{-\frac{|y-z|^2}{2s}}\, \mathrm{d}y. \tag{5.13}$$

The identity (5.12) is standard, e.g., [36, Equation (3.2)], [28, Equation (3.4)].

Proofs of Proposition 5.2(1).

Proof 1. This is Hamilton's matrix inequality [1] which is proven via the maximum principle. Indeed, by (5.11), since the smallest eigenvalue λ is a concave function of $M(x,t)$, it satisfies

$$\partial_t \lambda(x,t) \geq \frac{1}{2}\Delta\lambda(x,t) + L(\lambda(x,t)) + \lambda^2(x,t). \tag{5.14}$$

Hence, $\lambda(x,t) \geq g(t)$ where $g(t) := -1/t$ solves the equation,

$$\partial_t g(t) = g^2(t), \qquad g(0) = -\infty.$$

Proof 2. Use (5.12) and $\mathrm{Cov}\left[p_{e^{-t}x,e^{-2t}}\right] \succeq 0$.

Proof 3. Consequence of the intrinsic dimensional local logarithmic Sobolev inequality: The term inside the logarithm in [12, Equation 29] must be nonnegative. (This is analogous to the way in which the Li-Yau inequality is deduced from the dimensional local logarithmic Sobolev inequality [2].)

Proof 4. Mixture of log-convex densities is log-convex: For any fixed $y \in \mathbb{R}^n$ the function $x \mapsto \frac{e^{-2t}}{1-e^{-2t}}\frac{|x|^2}{2} + \log \mathrm{P}_t\delta_y(x)$ is convex, where δ_y is a point mass at y. Hence, the mixture $x \mapsto \frac{e^{-2t}}{1-e^{-2t}}\frac{|x|^2}{2} + \log \int f(y)\mathrm{P}_t\delta_y(x)\,\mathrm{d}y$ is also convex for any probability density $f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$. The proof is complete since $\int f(y)\mathrm{P}_t\delta_y(x)\,\mathrm{d}y = \mathrm{P}_t f(x)$ [11, Lemma 1.3, Appendix].

Proofs of Proposition 5.2(2).

Proof 1. We will use the maximum principle. Let $\Lambda(t,x)$ be the maximal eigenvalue of $M(t,x) := \nabla^2 \log \mathrm{H}_t f(x)$. Since the maximal eigenvalue Λ is a convex function of M, (5.11) implies that

$$\partial_t \Lambda(t,x) \leq \frac{1}{2}\Delta\Lambda(t,x) + L(\Lambda(t,x)) + \Lambda^2(t,x). \tag{5.15}$$

Fix $x \in \mathbb{R}^n$. Then,

$$\Lambda(t,x) \leq g(t) \quad \text{where } g \text{ solves the equation} \quad \partial_t g(t) = g^2(t), \quad g(0) = \Lambda(0,x),$$

for every t for which g is well-defined. Since the solution of the ordinary differential equation for g is $g(t) = \frac{\Lambda(0,x)}{1-\Lambda(0,x)t}$ for all t where the denominator does not vanish, we see that if $\Lambda(0,x) \leq 0$, then g is well-defined for all $t \geq 0$, and if $\Lambda(0,x) > 0$, then g is well-defined for all $t \in \left[0, \frac{1}{\Lambda(0,x)}\right)$. Hence, since $\Lambda(0,x) \leq c$,

$$\Lambda(t,x) \leq \frac{\Lambda(0,x)}{1-\Lambda(0,x)t} \begin{cases} \text{for any } t \in [0,\infty) & \text{if } c \leq 0 \\ \text{for any } t \in \left[0, \frac{1}{c}\right] & \text{if } c > 0, \end{cases}$$

which completes the proof of (5.7).

Proof 2. We will use the covariance identity following the argument in [36, Lemma 3.4(2)]. If f is c-log-concave then the measure $p_{z,s}$ in (5.13) is $\left(\frac{1}{s} - c\right)$-log-concave. In particular, $p_{e^{-t}x,1-e^{-2t}}$ is $\left(\frac{1}{1-e^{-2t}} - c\right)$-log-concave so by the Brascamp-Lieb inequality [4], as long as $\left(\frac{1}{1-e^{-2t}} - c\right) \geq 0$,

$$\mathrm{Cov}\left[p_{e^{-t}x,1-e^{-2t}}\right] \preceq \left(\frac{1}{1-e^{-2t}} - c\right)^{-1} \mathrm{Id}_n\,.$$

Hence, by (5.12),

$$\nabla^2 \log \mathrm{P}_t f(x) \preceq \frac{ce^{-2t}}{1-c(1-e^{-2t})}\,\mathrm{Id}_n \begin{cases} \text{for any } t \in [0,\infty) & \text{if } c \leq 1 \\ \text{for any } t \in \left[0, \log\left(\sqrt{\frac{c}{c-1}}\right)\right] & \text{if } c > 1, \end{cases}$$

which proves (5.6).

Proof 3. We will use the Prékopa-Leindler inequality. We need to show that $x \mapsto \mathrm{H}_t f(x) e^{-c_t \frac{|x|^2}{2}}$ is log-concave were $c_t := \frac{c}{1-ct}$. We have

$$\mathrm{H}_t f(x) e^{-\frac{c}{1-ct}\frac{|x|^2}{2}} = \int f(z) e^{-\frac{|z-x|^2}{2t}} e^{-\frac{c}{1-ct}\frac{|x|^2}{2}} \,\mathrm{d}z$$
$$= \int f(z) \exp\left(-\frac{1}{2}\left[\frac{1}{t(1+tc_t)}\,|z-(1+tc_t)x|^2 + \frac{c_t}{1+tc_t}|z|^2\right]\right) \mathrm{d}z = \mathrm{H}_{t(1+tc_t)} f_{c_t}((1+tc_t)x),$$

where $f_{c_t}(z) = f(z) e^{-\frac{1}{2}\frac{c_t}{1+tc_t}|z|^2}$, as long as $1+tc_t \geq 0$. Since f is c-log-concave, the function f_{c_t} is 0-log-concave as $c - \frac{c_t}{1+tc_t} = c - \frac{\frac{c}{1-ct}}{1+t\frac{c}{1-ct}} = 0$. Hence, $\mathrm{H}_{t(1+tc_t)} f_{c_t}$ is log-concave because it is the convolution of log-concave functions, which is log-cocnave by the Prékopa-Leindler inequality [17, §9]. Hence, as long as $1+tc_t \geq 0$, $\mathrm{H}_t f(x) e^{-\frac{c}{1-ct}\frac{|x|^2}{2}}$ is log-concave, which implies (5.7).

Proofs of Proposition 5.2(3).
Proof 1. We will use the maximum principle. Recall that the minimum eigenvalue $\lambda(t,x)$ of $M(t,x) := \nabla^2 \log \mathrm{H}_t f(x)$ satisfies (5.14). Hence, $\lambda(t,x) \geq g(t)$ where g solves the equation $\partial_t g(t) = g^2(t)$ with $g(0) = \lambda(0,x)$. Since the solution of the ordinary differential equation for g is $g(t) = \frac{\lambda(0,x)}{1-\lambda(0,x)t}$ for all t where the denominator does not vanish, we get that if $\lambda(0,x) \leq 0$, then, for all $t \geq 0$, $\lambda(t,x) \geq \frac{\lambda(0,x)}{1-\lambda(0,x)t}$. Since $\lambda(0,x) \geq c$ the proof of (5.8) is complete.

To prove (5.9) let $m(t,x) := \mathrm{Tr}[M(t,x)]$, and note that, by $\mathrm{Tr}[M^2(t,x)] \geq \frac{\mathrm{Tr}[M(t,x)]^2}{n}$, (5.11) implies that

$$\partial_t m(t,x) \geq \frac{1}{2}\Delta m(t,x) + L(m(t,x)) + \frac{m^2(t,x)}{n}. \tag{5.16}$$

Again, this implies that, for any fixed $x \in \mathbb{R}^n$ and $t \geq 0$, $m(t,x) \geq \frac{m(0,x)}{1-\frac{m(0,x)}{n}t}$. Since $m(t,x) \geq cn$ the proof of (5.9) is complete

Proof 2. The proof is based on the fact that a mixture of log-convex (res. log-subharmonic) functions is log-convex (res. log-subharmonic): The argument we present below is based on the proof of [35, Lemma 5] which treats the c-log-convex case. Here we adapt the proof also to the cn-log-subharmonic case. Our starting point is the following identity for the action of Orenstein-Uhlenbeck semigroup. The point of this identity to separate the effect of the semigroup on the quadratic part of the function.

Lemma 5.4. *Fix $\beta \geq 0$ and let $f(x) = e^{-R(x)-\beta\frac{|x|^2}{2}}$. Let (P_t) be the Orenstein-Uhlenbeck semigroup (5.3). Then,*

$$\mathrm{P}_t f(x) = e^{-\frac{\beta e^{-2t}}{1+(1-e^{-2t})\beta}\frac{|x|^2}{2}} \int e^{-R_{t,z}(x)} e^{-\frac{\beta}{2}|z|^2} \,\mathrm{d}z, \tag{5.17}$$

where

$$R_{t,z}(x) := (2\pi)^{-\frac{n}{2}}(1-e^{-2t}+\beta^{-1})^{-\frac{n}{2}} R\left(\frac{\sqrt{1-e^{-2t}}}{\sqrt{1-e^{-2t}+\beta^{-1}}}z + \frac{e^{-t}}{1+\beta(1-e^{-2t})}x\right). \tag{5.18}$$

Proof. By definition,

$$\begin{aligned}
\mathrm{P}_t f(x) &= \int f(e^{-t}x + \sqrt{1-e^{-2t}}y)\,\mathrm{d}\gamma(y) \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \int e^{-R(e^{-t}x+\sqrt{1-e^{-2t}}y)-\beta\frac{|e^{-t}x+\sqrt{1-e^{-2t}}y|^2}{2}-\frac{|y|^2}{2}}\,\mathrm{d}y \\
&= \frac{e^{-\beta e^{-2t}\frac{|x|^2}{2}}}{(2\pi)^{\frac{n}{2}}} \int e^{-R(e^{-t}x+\sqrt{1-e^{-2t}}y)-\beta\frac{2\langle e^{-t}x,\sqrt{1-e^{-2t}}y\rangle+(1-e^{-2t}+\beta^{-1})|y|^2}{2}}\,\mathrm{d}y \\
&= \frac{e^{-\beta e^{-2t}\frac{|x|^2}{2}}}{(2\pi)^{\frac{n}{2}}} \int e^{-R(e^{-t}x+\sqrt{1-e^{-2t}}y)-\frac{\beta}{2}\left|\sqrt{1-e^{-2t}+\beta^{-1}}y+\frac{\sqrt{1-e^{-2t}}e^{-t}}{\sqrt{1-e^{-2t}+\beta^{-1}}}x\right|^2} e^{-\frac{-(1-e^{-2t})e^{-2t}}{1-e^{-2t}+\beta^{-1}}\frac{\beta|x|^2}{2}}\,\mathrm{d}y \\
&= \frac{e^{-\left(1-\frac{1-e^{-2t}}{1-e^{-2t}+\beta^{-1}}\right)\frac{e^{-2t}\beta}{2}|x|^2}}{(2\pi)^{\frac{n}{2}}} \int e^{-R(e^{-t}x+\sqrt{1-e^{-2t}}y)-\frac{\beta}{2}\left|\sqrt{1-e^{-2t}+\beta^{-1}}y+\frac{\sqrt{1-e^{-2t}}e^{-t}}{\sqrt{1-e^{-2t}+\beta^{-1}}}x\right|^2}\,\mathrm{d}y \\
&= \frac{e^{-\frac{\beta e^{-2t}}{1+(1-e^{-2t})\beta}\frac{|x|^2}{2}}}{(2\pi)^{\frac{n}{2}}} \int e^{-R(e^{-t}x+\sqrt{1-e^{-2t}}y)-\frac{\beta}{2}\left|\sqrt{1-e^{-2t}+c^{-1}}y+\frac{\sqrt{1-e^{-2t}}e^{-t}}{\sqrt{1-e^{-2t}+\beta^{-1}}}x\right|^2}\,\mathrm{d}y.
\end{aligned}$$

Let $z = \sqrt{1-e^{-2t}+\beta^{-1}}y + \frac{\sqrt{1-e^{-2t}}e^{-t}}{\sqrt{1-e^{-2t}+\beta^{-1}}}x$, so that $\mathrm{d}y = (1-e^{-2t}+\beta^{-1})^{-\frac{n}{2}}\,\mathrm{d}z$, and change variables to get

$$\begin{aligned}
\mathrm{P}_t f(x) &= \frac{e^{-\frac{\beta e^{-2t}}{1+(1-e^{-2t})\beta}\frac{|x|^2}{2}}}{(2\pi)^{\frac{n}{2}}} \int e^{-R\left(\frac{\sqrt{1-e^{-2t}}}{\sqrt{1-e^{-2t}+\beta^{-1}}}z+\frac{e^{-t}}{1+\beta(1-e^{-2t})}x\right)-\frac{\beta}{2}|z|^2}(1-e^{-2t}+\beta^{-1})^{-\frac{n}{2}}\,\mathrm{d}z \\
&= e^{-\frac{\beta e^{-2t}}{1+(1-e^{-2t})\beta}\frac{|x|^2}{2}} \int e^{-R_{t,z}(x)}e^{-\frac{\beta}{2}|z|^2}\,\mathrm{d}z,
\end{aligned}$$

where

$$R_{t,z}(x) := (2\pi)^{-\frac{n}{2}}(1-e^{-2t}+\beta^{-1})^{-\frac{n}{2}} R\left(\frac{\sqrt{1-e^{-2t}}}{\sqrt{1-e^{-2t}+\beta^{-1}}}z + \frac{e^{-t}}{1+\beta(1-e^{-2t})}x\right).$$

□

Let us now complete the proof of Proposition 5.2(3) using (5.17). Given f which is c-log-convex (res. cn-log-subharmonic) let $R(x) := -\log f(x) - |c|\frac{|x|^2}{2}$. Then R is concave (res. superharmonic), so with $R_{t,z}$ as in (5.18) we have that $x \mapsto e^{-R_{t,z}(x)}$ is log-convex (res. log-subharmonic). Since the mixture of log-convex functions is log convex [33, Chapter 16.B] (res. the mixture of log-subharmonic functions is log-subharmonic [20, Proposition 2.2]), it follows that $x \mapsto \int e^{-R_{t,z}(x)}e^{-\frac{|c|}{2}|z|^2}\,\mathrm{d}z$ is log-convex (res. log-subharmonic). From (5.17) we see that $\mathrm{P}_t f$ is $\frac{ce^{-2t}}{1-c(1-e^{-2t})}$-log-convex (res. $n\frac{ce^{-2t}}{1-c(1-e^{-2t})}$-log-subharmonic). □

References

[1] Giacomo Ascione, Daniele Castorina, Giovanni Catino, and Carlo Mantegazza, *A matrix Harnack inequality for semilinear heat equations*, Math. Eng. **5** (2023), Paper No. 003, 15.

[2] Dominique Bakry, Fran¸cois Bolley, and Ivan Gentil, *The Li-Yau inequality and applications under a curvature-dimension condition*, Ann. Inst. Fourier (Grenoble) **67** (2017), 397–421.

[3] Neal Bez, Shohei Nakamura, and Hiroshi Tsuji, *Stability of hypercontractivity, the logarithmic Sobolev inequality, and Talagrand's cost inequality*, J. Funct. Anal. **285** (2023), Paper No. 110121, 66.

[4] 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*, J. Functional Analysis **22** (1976), 366–389.

[5] Luis A. Caffarelli, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys. **214** (2000), 547–563.

[6] Eric A. Carlen, *Some integral identities and inequalities for entire functions and their application to the coherent state transform*, J. Funct. Anal. **97** (1991), 231–249.

[7] Djalil Chafaï, *Aspects of Coulomb gases*, Topics in statistical mechanics, Panor. Synthèses, vol. 59, Soc. Math. France, Paris, 2023, pp. 1–40.

[8] Dario Cordero-Erausquin, *Some applications of mass transport to Gaussian-type inequalities*, Archive for rational mechanics and analysis **161** (2002), 257–269.

[9] Dario Cordero-Erausquin and Alessio Figalli, *Regularity of monotone transport maps between unbounded domains*, Discrete and Continuous Dynamical Systems **39** (2019), 7101–7112.

[10] Dario Cordero-Erausquin, Matthieu Fradelizi, and Bernard Maurey, *The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems*, Journal of Functional Analysis **214** (2004), 410–427.

[11] Ronen Eldan and James R. Lee, *Regularization under diffusion and anticoncentration of the information content*, Duke Math. J. **167** (2018), 969–993.

[12] Alexandros Eskenazis and Yair Shenfeld, *Intrinsic dimensional functional inequalities on model spaces*, J. Funct. Anal. **286** (2024), Paper No. 110338, 56.

[13] Max Fathi, Dan Mikulincer, and Yair Shenfeld, *Transportation onto log-Lipschitz perturbations*, Calc. Var. Partial Differential Equations **63** (2024), Paper No. 61, 25.

[14] A. Figalli, F. Maggi, and A. Pratelli, *A mass transportation approach to quantitative isoperimetric inequalities*, Invent. Math. **182** (2010), 167–211.

[15] Rupert L. Frank, *Sharp inequalities for coherent states and their optimizers*, Adv. Nonlinear Stud. **23** (2023), Paper No. 20220050, 28.

[16] Rupert L Frank, Fabio Nicola, and Paolo Tilli, *The generalized Wehrl entropy bound in quantitative form*, arXiv preprint arXiv:2307.14089 (2023).

[17] Richard Gardner, *The Brunn-Minkowski inequality*, Bulletin of the American mathematical society **39** (2002), 355–405.

[18] N. Gozlan, M. Madiman, C. Roberto, and P. M. Samson, *Deviation inequalities for convex functions motivated by the Talagrand conjecture*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **457** (2017), 168–182.

[19] Nathael Gozlan and Maxime Sylvestre, *Global regularity estimates for optimal transport via entropic regularisation*, arXiv preprint arXiv:2501.11382 (2025).

[20] Piotr Graczyk, Todd Kemp, and Jean-Jacques Loeb, *Hypercontractivity for log-subharmonic functions*, J. Funct. Anal. **258** (2010), 1785–1805.

[21] Piotr Graczyk, Todd Kemp, and Jean-Jacques Loeb, *Strong logarithmic Sobolev inequalities for log-subharmonic functions*, Canad. J. Math. **67** (2015), 1384–1410.

[22] Karlheinz Gröchenig, *Foundations of time-frequency analysis*, Springer Science & Business Media, 2001.

[23] Leonard Gross and Martin Grothaus, *Reverse hypercontractivity for subharmonic functions*, Canad. J. Math. **57** (2005), 506–534.

[24] Richard S Hamilton, *Matrix Harnack estimate for the heat equation*, Communications in analysis and geometry **1** (1993), 113–126.

[25] Gilles Hargé, *A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces*, Probability theory and related fields **130** (2004), 415–440.

[26] Young-Heon Kim and Emanuel Milman, *A generalization of Caffarelli's contraction theorem via (reverse) heat flow*, Math. Ann. **354** (2012), 827–862.

[27] Boaz Klartag, *Marginals of geometric inequalities*, Geometric Aspects of Functional Analysis: Israel Seminar 2004–2005, Springer, 2007, pp. 133–166.

[28] Bo'az Klartag and Eli Putterman, *Spectral monotonicity under Gaussian convolution*, Ann. Fac. Sci. Toulouse Math. (6) **32** (2023), 939–967.

[29] A. V. Kolesnikov, *On Sobolev regularity of mass transport and transportation inequalities*, Theory Probab. Appl. **57** (2013), 243–264.

[30] Alexander V Kolesnikov, *Mass transportation and contractions*, Preprint arXiv:1103.1479 (2011).

[31] Aleksei Kulikov, Fabio Nicola, Joaquim Ortega-Cerdà, and Paolo Tilli, *Monotonicity theorem for subharmonic functions on manifolds*, arXiv preprint arXiv:2212.14008 (2022).

[32] Elliott H. Lieb, *Proof of an entropy conjecture of Wehrl*, Comm. Math. Phys. **62** (1978), 35–41.

[33] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold, *Inequalities: theory of majorization and its applications*, second ed., Springer Series in Statistics, Springer, New York, 2011.
[34] James Melbourne and Cyril Roberto, *Transport-majorization to analytic and geometric inequalities*, J. Funct. Anal. **284** (2023), Paper No. 109717, 36.
[35] Dan Mikulincer and Yair Shenfeld, *On the Lipschitz Properties of Transportation Along Heat Flows*, Geometric Aspects of Functional Analysis, Lecture Notes in Math., vol. 2327, Springer, Cham, 2023, pp. 269–290.
[36] Dan Mikulincer and Yair Shenfeld, *The Brownian transport map*, Probab. Theory Related Fields **190** (2024), 379–444.
[37] Reto Müller, *Differential harnack inequalities and the ricci flow*, vol. 5, European Mathematical Society, 2006.
[38] Joe Neeman, *Lipschitz changes of variables via heat flow*, arXiv preprint arXiv:2201.03403 (2022).
[39] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), 361–400.
[40] Cédric Villani, *Topics in optimal transportation*, vol. 58, American Mathematical Soc., 2021.
[41] Alfred Wehrl, *On the relation between classical and quantum-mechanical entropy*, Reports on Mathematical Physics **16** (1979), 353–358.
[42] Kehe Zhu, *Analysis on Fock spaces*, Graduate Texts in Mathematics, vol. 263, Springer, New York, 2012.

Courant Institute of Mathematical Sciences, New York University. 251 Mercer Street, New York, NY 10012-1185
Email address: guido@cims.nyu.edu

Division of Applied Mathematics, Brown University, Providence, RI, USA
Email address: Yair_Shenfeld@Brown.edu