Available online at www.sciencedirect.com
ScienceDirect
Journal of Functional Analysis 257 (2009) 1175–1221
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

# Mass transportation proofs of free functional inequalities, and free Poincaré inequalities

Michel Ledoux [a], Ionel Popescu [b,c,*]

[a] *Institut de Mathématiques de Toulouse, Université de Toulouse, F-31062 Toulouse, France*
[b] *Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332, United States*
[c] *IMAR 21, Calea Grivitei Street 010702-Bucharest, Sector 1, Romania*

Received 13 February 2009; accepted 17 March 2009

Available online 31 March 2009

Communicated by D. Stroock

**Abstract**

This work is devoted to direct mass transportation proofs of families of functional inequalities in the context of one-dimensional free probability, avoiding random matrix approximation. The inequalities include the free form of the transportation, Log-Sobolev, HWI interpolation and Brunn–Minkowski inequalities for strictly convex potentials. Sharp constants and some extended versions are put forward. The paper also addresses two versions of free Poincaré inequalities and their interpretation in terms of spectral properties of Jacobi operators. The last part establishes the corresponding inequalities for measures on $\mathbb{R}_+$ with the reference example of the Marcenko–Pastur distribution.
© 2009 Elsevier Inc. All rights reserved.

*Keywords:* Functional inequalities; Mass transport; Spectral gap; Random matrices

## 1. Introduction

A distinguished role in the world of functional inequalities is played by the logarithmic Sobolev (Log-Sobolev) inequality and the Talagrand or transportation cost inequality. There is an extensive literature dedicated to these inequalities in the classical setting of Euclidean and Riemannian spaces (cf. e.g. [2,23,29,32]).

* Corresponding author at: Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332, United States.
*E-mail address:* ipopescu@math.gatech.edu (I. Popescu).

0022-1236/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfa.2009.03.011

Given a probability measure $\nu$ on $\mathbb{R}^d$, the transportation cost inequality states that for some $\rho > 0$ and any other probability measure $\mu$ on $\mathbb{R}^d$,

$$\rho W_2^2(\mu, \nu) \leqslant E(\mu \mid \nu). \tag{$T(\rho)$}$$

Here $W_2(\mu, \nu)$ is the Wasserstein distance between $\mu$ and $\nu$ of finite second moment defined by

$$W_2(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \left( \iint |x-y|^2 \pi(dx, dy) \right)^{1/2}$$

with $\Pi(\mu, \nu)$ denoting the set of probability measures on $\mathbb{R}^{2d}$ with marginals $\mu$ and $\nu$ and

$$E(\mu \mid \nu) = \int \log \frac{d\mu}{d\nu} \, d\mu$$

is the relative entropy of $\mu$ with respect to $\nu$ if $\mu \ll \nu$ and $+\infty$ otherwise. The Log-Sobolev inequality is that for any $\mu$

$$E(\mu \mid \nu) \leqslant \frac{1}{2\rho} I(\mu \mid \nu) \tag{$LSI(\rho)$}$$

where

$$I(\mu \mid \nu) = \int \left| \nabla \log \frac{d\mu}{d\nu} \right|^2 d\mu$$

is the Fisher information of $\mu$ with respect to $\nu$ which is defined in the case $\mu \ll \nu$ with $\frac{d\mu}{d\nu}$ being differentiable. A more subtle inequality is the HWI inequality relating entropy (notice that $E(\mu \mid \nu)$ is $H(\mu \mid \nu)$ in [25] which explains the H), Wasserstein distance W, and Fisher information I

$$E(\mu \mid \nu) \leqslant \sqrt{I(\mu \mid \nu)} W_2(\mu, \nu) - \frac{\rho}{2} W_2^2(\mu, \nu). \tag{$HWI(\rho)$}$$

Poincaré's inequality in this classical context is that for any compactly supported and smooth function $\psi$ on $\mathbb{R}^d$,

$$\rho \operatorname{Var}_\mu(\psi) \leqslant \int |\nabla \psi|^2 \mu(dx) \tag{$P(\rho)$}$$

where $\operatorname{Var}_\mu(\psi) = \int \psi^2(x) \mu(dx) - (\int \psi(x) \mu(dx))^2$ is the variance of $\psi$ with respect to $\mu$.

Starting with Gaussian measures [14,28], these inequalities were established for measures on $\mathbb{R}^d$ with strictly convex potentials by the Bakry–Émery criterion [2,23,29,32]. More precisely, if $\nu(dx) = e^{-V(x)} dx$, with $V(x) - \rho|x|^2$ convex on $\mathbb{R}^d$ for some $\rho > 0$, both $T(\rho)$ and $LSI(\rho)$ hold true. Otto and Villani generated interest in this topic through their remarkable paper [25], in which they showed that the logarithmic Sobolev inequality implies the trasportation inequality, in a rather general setting. This connection was actually put further through the stronger $HWI(\rho)$ inequality, which was shown in [25] to be valid in the case $V(x) - \rho|x|^2$ is convex for some $\rho \in \mathbb{R}$, When $\rho > 0$, $LSI(\rho)$ is a consequence of $HWI(\rho)$. Subsequently the main result from [25] was simplified and extended, for example [5] and recently [13] to mention only two sources.

Another interesting connection in these families of functional inequalities is that any of $T(\rho)$, $LSI(\rho)$ or $HWI(\rho)$ imply the Poincaré inequality $P(\rho)$.

The work [25] by Otto and Villani input in a powerful way the use of mass transportation ideas in the context of functional inequalities. Starting from this, Cordero-Erausquin used in [9] direct convexity arguments combined with mass transport methods to reprove the Log-Sobolev, transportation and HWI inequalities for measures with strictly convex potentials. The strategy is going back to the original approach of [28] to the transportation inequality (see also [4]).

In the world of free probability, as it was shown by Ben Arous and Guionnet in [1], one can realize the free entropy as the rate function of the large deviations for the distribution of eigenvalues of some $n \times n$ complex random matrix ensembles (see also [19]). To wit a little bit here, let $V : \mathbb{R} \to \mathbb{R}$ be a nice function with enough growth at infinity and define the probability distribution

$$\mathbb{P}_n(dM) = \frac{1}{Z_n} e^{-n \operatorname{Tr}_n(V(M))} dM$$

on the set $\mathcal{H}_n$ of complex Hermitian $n \times n$ matrices where $dM$ is the Lebesgue measure on $\mathcal{H}_n$. For a matrix $M$, let $\mu_n(M) = \frac{1}{n}\sum_{k=1}^n \delta_{\lambda_k(M)}$ be the distribution of eigenvalues of $M$. These are random variables with values in $\mathcal{P}(\mathbb{R})$, the set of probability measures on $\mathbb{R}$ which converge almost surely to a non-random measure $\mu_V$ on $\mathbb{R}$. For a measure $\mu$ on $\mathbb{R}$, its the logarithmic energy with external field $V$ is defined by

$$E(\mu) = \int V(x)\,\mu(dx) - \iint \log|x-y|\,\mu(dx)\,\mu(dy).$$

The minimizer of $E(\mu)$ over all probability measures on $\mathbb{R}$ is exactly the measure $\mu_V$. From [1] we learned that the distributions of $\{\mu_n\}_{n\geqslant 1}$ under $\mathbb{P}_n$ satisfy a large deviations principle with scaling $n^2$ and rate function given by

$$R(\mu) = E(\mu) - E(\mu_V).$$

The example of the quadratic potential $V(x) = x^2$ defining the paradigmatic Gaussian Unitary Ensemble in random matrix theory gives rise to the celebrated semicircular law as equilibrium measure.

Within this random matrix framework, if $V(x) - \rho x^2$ is smooth and convex for some $\rho > 0$, then the function $\Phi(M) = \operatorname{Tr}_n(V(M))$ is strongly convex ($\Phi(M) - n\rho|M|^2$ is convex) on $\mathbb{R}^{n^2} = \mathcal{H}_n$. An application of the classical $LSI(n\rho)$ on $\mathcal{H}_n$ for large $n$ was used by Biane [3] to prove a Log-Sobolev inequality in the context of one-dimensional free probability which holds (cf. [18]) in the following form

$$E(\mu) - E(\mu_V) \leqslant \frac{1}{4\rho} I(\mu) \tag{1.1}$$

for any probability measure $\mu$ on $\mathbb{R}$ whose density with respect to the Lebegue measure is in $L^3(\mathbb{R})$, where

$$I(\mu) = \int \big(H\mu(x) - V'(x)\big)^2\,\mu(dx)$$

with $H\mu = 2\int \frac{1}{x-y}\,\mu(dx)$ being the Hilbert transform of $\mu$.

More precisely, Biane and Voiculescu used the free Ornstein Uhlenbeck process and the complex Burger equation. Using the large random matrix strategy, Hiai Petz and Ueda [18] reproved and extended the result of Biane and Voiculescu in the following form. If $V(x) - \rho x^2$ is convex for some $\rho > 0$, then for every probability measure $\mu$ on $\mathbb{R}$,

$$\rho W_2^2(\mu, \mu_V) \leqslant E(\mu) - E(\mu_V). \tag{1.2}$$

Later, the first author [24] gave a simpler proof of (1.1) and (1.2) based on a free version of the geometric Brunn–Minkowski inequality obtained as a random matrix limiting case of its classical counterpart. He also showed the free analog of the Otto–Villani theorem indicating that the free Log-Sobolev inequality implies the free transportation inequality (1.2).

The first scope of this paper is to provide direct proofs of the preceding functional inequalities in free probability without random matrix approximation. The second author of this paper in [26] gave a simple proof of the transportation inequality (1.2) on the same line of ideas as in [28] for the classical case where random matrix theory is entirely avoided.

In this paper, following the approach of Cordero-Erausquin [9] (see also [4]), we use a combination of mass transport and convex analysis which apply to strictly convex potentials. The methods allow us besides to enlarge the class of potentials under consideration, in particular in instances which lack a proper random matrix approximation. For example, we cover potentials $V$ on the line such that $V(x) - \rho|x|^p$ is convex for some $\rho > 0$ and $p > 1$ as well as a class of bounded perturbations of convex potentials. Using this approach, we present here an HWI free inequality for various cases of potentials. For the case $V(x) - \rho x^2$ convex for some $\rho \in \mathbb{R}$, this is

$$E(\mu) - E(\mu_V) \leqslant \sqrt{I(\mu)}W_2(\mu, \mu_V) - \rho W_2^2(\mu, \mu_V). \tag{1.3}$$

Also a Brunn–Minkowski inequality receives a direct proof as well.

One interesting byproduct of our method is that some constants may be shown to be sharp. For the case of a quadratic $V$, Eqs. (1.1), (1.2) and (1.3) are sharp.

Another topic discussed here in Section 3 is a free form of the transportation inequality which does not depend on the potential and that might be thought of as a version of the celebrated Pinsker inequality comparing total variation distance and entropy between probability measures. As opposed to the classical case, the free counterpart is more delicate.

The second part of this work is devoted to free one-dimensional Poincaré inequalities. Using random matrix approximations and the classical Poincaré inequality, we first give an ansatz to what could be a possible Poincaré inequality in the free probability world. In the case of $V(x) - \rho x^2$ convex for some $\rho > 0$, such that the measure $\mu_V$ has support $[-1, 1]$, this states as,

$$\int \phi'(x)^2 \mu_V(dx) \geqslant \frac{\rho}{2\pi^2} \int_{-1}^{1}\int_{-1}^{1} \left(\frac{\phi(x) - \phi(y)}{x - y}\right)^2 \frac{1 - xy}{\sqrt{1 - x^2}\sqrt{1 - y^2}}\, dx\, dy, \tag{1.4}$$

for any smooth function $\phi$ on the interval $[-1, 1]$.

There is also a second version of the Poincaré which is discussed in [3] for the case of the semicircular law. This inequality has a natural meaning in the context of free probability as the derivative $\nabla\phi$ of a function from the classical $P(\rho)$ is replaced by the noncommutative derivative $\frac{\phi(x)-\phi(y)}{x-y}$, and thus our second version takes the form

$$\iint\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\mu(dx)\,\mu(dy)\geqslant C\,\mathrm{Var}_\mu(\phi)\quad\text{for every }\phi\in C_0^1(\mathbb{R}).\tag{1.5}$$

As opposed to (1.4) which requires certain conditions on the measure $\mu_V$, it turns out that (1.5) is always satisfied for any compactly supported measure $\mu$ with some constant. As was shown in [3] for the semicircular law, one can completely characterize the distribution in terms of the constant $C$.

After the use of convexity, inequality (1.4) may actually be interpreted as a spectral gap as follows. On $L^2\big(\frac{\mathbb{1}_{[-2,2]}(x)dx}{\sqrt{4-x^2}}\big)$ take the Jacobi operator

$$Lf=-\big(1-x^2\big)f''(x)+xf'(x)$$

and the counting number operator defined by

$$NT_n=nT_n$$

where $T_n$ are the Chebyshev polynomials of the first kind, which are orthogonal in $L^2\big(\frac{\mathbb{1}_{[-2,2]}(x)\,dx}{\sqrt{4-x^2}}\big)$. Then, (1.4) for $V(x)=x^2/2$ is equivalent to

$$L\geqslant N.$$

Inequality (1.5) in the case of $V(x)=x^2/2$ can also be seen as the spectral gap for the counting number operator on $L^2(\mathbb{1}_{[-2,2]}(x)\sqrt{4-x^2}\,dx)$ with respect to the basis given by the Chebyshev polynomials of second kind. A more general situation is discussed in Section 9 which includes both versions of the Poincaré inequalities.

As we mentioned already, in the classical setting, the Log-Sobolev and the transportation inequality imply the Poincaré inequalities. We do not have a satisfactory picture of these implications in the free context, for any of the two versions of the Poincaré inequality discussed here.

In the final part, we investigate the preceding families of functional inequalities for probability measures supported on the positive real axis. The random matrix context is the one of Wishart ensembles with reference measure the Marcenko–Pastur distribution as opposed to the semicircular law, and the free functional inequalities correspond formally to the case of potentials $V(x)=rx-s\log(x)$ for $r>0$, $s\geqslant 0$ on $\mathbb{R}_+$. Using the mass transportation method, we prove transportation, Log-Sobolev and HWI inequalities which were not investigated previously. A version of the Poincaré inequality is also discussed.

The structure of the paper is as follows. Sections 2, 4, 5 and 6 deal with the mass transportation proofs of respectively the transportation, Log-Sobolev, HWI and Brunn–Minkowski inequalities. Section 3 studies transportation inequalities which involve some metric on the probabilities and which are independent of the potential $V$. Sections 7 and 8 are devoted to the two versions of the Poincaré inequality in the free context, related in Section 9 through Jacobi operators. Section 10 investigates the preceding inequalities with respect to the Marcenko–Pastur distribution and its convex extensions.

## 2. Transportation inequality

Throughout this paper we consider lower semicontinuous potentials $V:\mathbb{R}\to\mathbb{R}$ such that

$$\lim_{|x|\to\infty}\bigl(V(x)-2\log|x|\bigr)=\infty. \tag{2.1}$$

For a given Borel set $\Gamma\subset\mathbb{R}$, denote by $\mathcal{P}(\Gamma)$ the set of probability measures supported on $\Gamma$.

The logarithmic energy with external potential $V$ is defined by

$$E_V(\mu):=\int V(x)\,\mu(dx)-\iint\log|x-y|\,\mu(dx)\,\mu(dy)$$

whenever both integrals exist and have finite values. In particular for measures $\mu$ which have atoms, $E_V(\mu)=+\infty$ because the second integral is $+\infty$.

It is known (see [27] or [11]) that under condition (2.1) there exists a unique minimizer of $E_V$ in the set $\mathcal{P}(\mathbb{R})$ and the solution $\mu_V$ is compactly supported. The variational characterization of the minimizer $\mu_V$ (cf. [27, Theorem 1.3]) is that for a constant $C\in\mathbb{R}$,

$$\begin{aligned} V(x)&\geqslant 2\int\log|x-y|\,\mu_V(dy)+C\quad\text{for quasi-every } x\in\mathbb{R},\\ V(x)&=2\int\log|x-y|\,\mu_V(dy)+C\quad\text{for quasi-every } x\in\operatorname{supp}(\mu_V), \end{aligned}\tag{2.2}$$

where $\operatorname{supp}(\mu_V)$ stands for the support of $\mu$. If $\mu$ is such that $E_V(\mu)<\infty$, then Borel quasi-everywhere sets have $\mu$ measure 0 and thus the properties above hold almost surely with respect to $\mu$.

For simplicity of the notation, we will drop the subscript $V$ from $E_V$ unless the dependence of the potential has to be highlighted.

Now we summarize some known facts about the equilibrium measure and its support as one can easily deduce them from [27, Chapter IV] and [11, Chapter 6].

**Theorem 1.**

1. *Let $V$ be a potential satisfying* (2.1) *and $\alpha\neq 0$, $\beta\in\mathbb{R}$. Set $V_{\alpha,\beta}(x)=V(\alpha x+\beta)$. Then, $\mu_{V_{\alpha,\beta}}=((\mathrm{id}-\beta)/\alpha)_{\#}\mu_V$ and*

$$E_V(\mu_V)=E_{V_{\alpha,\beta}}(\mu_{V_{\alpha,\beta}})-\log|\alpha|. \tag{2.3}$$

2. *If $V$ is convex satisfying* (2.1)*, then the support of the equilibrium measure $\mu_V$ consists of one interval $[a,b]$ where $a$ and $b$ solve the system*

$$\begin{cases} \dfrac{1}{2\pi}\displaystyle\int_a^b V'(x)\sqrt{\frac{x-a}{b-x}}\,dx=1,\\[2ex] \dfrac{1}{2\pi}\displaystyle\int_a^b V'(x)\sqrt{\frac{b-x}{x-a}}\,dx=-1. \end{cases}\tag{2.4}$$

3. *Let $V$ be either a $C^2$ satisfying* (2.1) *whose equilibrium measure has support* $[a,b]$. *Then the equilibrium measure $\mu_V$ has density $g(x)$, given by*

$$g(x) = \mathbb{1}_{[a,b]}(x)\frac{\sqrt{(x-a)(b-x)}}{2\pi^2}\int_a^b \frac{V'(y)-V'(x)}{(y-x)\sqrt{(y-a)(b-y)}}\,dy. \tag{2.5}$$

4. *If $V$ is $C^2$, then*

$$V'(x) = p.v.\int \frac{2}{x-y}\mu_V(dx) \quad \text{for } \mu_V\text{-a.s. all } x \in \operatorname{supp}(\mu_V), \tag{2.6}$$

*where p.v. stands for the principal value integral. Notice that the principal value makes sense as $\mu_V$ has a continuous density.*

We mention as a basic example that if $V(x) = \rho x^2$ is quadratic, then $\mu_V$ is the semicircular law

$$\mu_V(dx) = \mathbb{1}_{[-\sqrt{2/\rho},\sqrt{2/\rho}]}(x)\sqrt{2\rho - \rho^2 x^2}\,\frac{dx}{\pi}.$$

In this work, for $p \geqslant 1$, we use $W_p(\mu,\nu)$ for the Wasserstein distance on the space of probability measures on $\mathbb{R}$ defined as

$$W_p(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)}\left(\iint |x-y|^p \pi(dx,dy)\right)^{1/p} \tag{2.7}$$

with $\Pi(\mu,\nu)$ denoting the set of probability measures on $\mathbb{R}^2$ with marginals $\mu$ and $\nu$. Note here that if $\theta$ is the (non-decreasing) transport map such that $\theta_{\#}\mu = \nu$, then

$$W_p^p(\mu,\nu) = \int \left|\theta(x) - x\right|^p \nu(dx). \tag{2.8}$$

For a detailed discussion on this topic we refer the reader to [29].

Our first result concerns the free version of the transportation cost inequality. As discussed in the introduction, the first assertion for strictly convex potentials was initially proved by large matrix approximation in [18]. The strategy of proof is inspired from [4,9,28] (see [26]).

**Theorem 2** *(Transportation inequality).*

1. *If $V$ is $C^2$ and $V(x) - \rho x^2$ is convex for some $\rho > 0$, then for any probability measure $\mu$ on $\mathbb{R}$,*

$$\rho W_2^2(\mu,\mu_V) \leqslant E(\mu) - E(\mu_V). \tag{2.9}$$

*If $V(x) = \rho x^2$, then the equality in* (2.9) *is attained for measures $\mu = \theta_{\#}\mu_V$, with $\theta(x) = x + m$, therefore the constant $\rho$ in front of $W_2^2(\mu,\mu_V)$ is sharp.*

2. *Assume that $V$ is $C^2$, convex and $V''(x) \geqslant \rho > 0$ for all $|x| \geqslant r$. Then, there is a constant $C = C(r, \rho, \mu_V, V) > 0$, such that*

$$CW_2^2(\mu, \mu_V) \leqslant E(\mu) - E(\mu_V). \tag{2.10}$$

3. *In the case $V$ is $C^2$ and $V(x) - \rho|x|^p$ is convex for some real number $p > 1$, then, for any probability measure $\mu$ on $\mathbb{R}$,*

$$c_p \rho W_p^p(\mu, \mu_V) \leqslant E(\mu) - E(\mu_V) \tag{2.11}$$

*where $c_p = \inf_{x\in\mathbb{R}}(|1+x|^p - |x|^p - p\,\mathrm{sign}(x)|x|^{p-1}) > 0$.*

**Proof.** 1. Since there is nothing to prove in the case $E(\mu) = \infty$, we assume that $E(\mu) < \infty$. In this case we also have that the measure $\mu$ and $\mu_V$ both have second finite moments.

Now we take the non-decreasing transportation map $\theta$ such that $\theta_{\#}\mu_V = \mu$ which exists due to the lack of atoms of $\mu_V$. Using the transport map $\theta$, we first write

$$\begin{aligned} E(\mu) - E(\mu_V) = {} & \int \big(V\big(\theta(x)\big) - V(x) - V'(x)\big(\theta(x) - x\big)\big)\,\mu_V(dx) \\ & + \iint \left(\frac{\theta(x)-\theta(y)}{x-y} - 1 - \log\frac{\theta(x)-\theta(y)}{x-y}\right)\mu_V(dx)\,\mu_V(dy) \end{aligned} \tag{2.12}$$

where in between we used the variational equation (2.6) to justify that

$$\begin{aligned} \int V'(x)\big(\theta(x) - x\big)\,\mu_V(dx) &= 2\iint \frac{\theta(x)-x}{x-y}\,\mu_V(dy)\,\mu_V(dx) \\ &= \iint \frac{(\theta(x)-x) - (\theta(y)-y)}{x-y}\,\mu_V(dy)\,\mu_V(dx). \end{aligned}$$

Since $V(x) - \rho x^2$ is convex, for any $x, y$ the following holds

$$V(y) - V(x) - V'(x)(y-x) \geqslant \rho\big(y^2 - x^2 - 2x(y-x)\big) = \rho(y-x)^2.$$

On the other hand since $a - 1 \geqslant \log(a)$ for any $a \geqslant 0$, Eqs. (2.12) and (2.8) yield (2.9).

In the case $V(x) = \rho x^2$ it is easy to see that for $\theta(x) = x + m$, all inequalities involved become equalities, thus we attain equality in (2.9) for translations of $\mu_V$.

2. We start the proof with (2.12), whereas this time we need to exploit the logarithmic term to get our inequality. The idea is to use the strong convexity where $\psi(x) := \theta(x) - x$ takes large values and for small values of $\psi(x)$ we try to compensate this with the second integral of (2.12).

Notice in the first place that by Taylor's theorem we have that

$$V(y) - V(x) - V'(x)(y-x) = (y-x)^2 \int_0^1 V''\big((1-\tau)x + \tau y\big)(1-\tau)\,d\tau. \tag{2.13}$$

Now, let us assume that the support of the equilibrium measure $\mu_V$ is $[a, b]$. Next, $V''(x) \geqslant 0$ and $V''(x) \geqslant \rho$ for $|x| \geqslant r$, implies that for $|y| \geqslant 2r + 2\max\{|a|, |b|\}$, we obtain that

$$\begin{aligned} V(y)-V(x)-V'(x)(y-x) &\geqslant (y-x)^2\int_{1/2}^{1} V''\big((1-\tau)x+\tau y\big)(1-\tau)\,d\tau \\ &\geqslant \rho(y-x)^2/8 \quad \text{for any } x\in[a,b]. \end{aligned}$$

Now write $\theta(x)=x+\psi(x)$. Thus using (2.12), and denoting $R=2r+2\max\{|a|,|b|\}$ we continue with

$$\begin{aligned} &\int \big(V\big(\theta(x)\big)-V(x)-V'(x)\big(\theta(x)-x\big)\big)\,\mu_V(dx) \\ &\quad\geqslant \frac{1}{2}\int \psi^2(x)\int_0^1 V''\big(x+\tau\psi(x)\big)(1-\tau)\,d\tau\,\mu_V(dx) \\ &\quad\geqslant \frac{\rho}{16}\int_{|\psi|\geqslant R}\psi^2(x)\,\mu_V(dx). \end{aligned} \tag{2.14}$$

This inequality provides a lower bound of the first term in (2.12). Further, it is not hard to check that

$$\begin{aligned} &\int_{|\psi|\geqslant R}\psi^2(x)\,\mu_V(dx) \\ &\quad= \frac{1}{2}\int \mathbb{1}_{|\psi|\geqslant R}(x)\psi^2(x)\,\mu_V(dx)+\frac{1}{2}\int \mathbb{1}_{|\psi|\geqslant R}(y)\psi^2(y)\,\mu_V(dy) \\ &\quad\geqslant \frac{1}{8}\iint \mathbb{1}_{|\psi(x)-\psi(y)|\geqslant 2R}(x,y)\big|\psi(x)-\psi(y)\big|^2\,\mu_V(dx)\,\mu_V(dy). \end{aligned} \tag{2.15}$$

Now we treat the second integral on the left-hand side of (2.12). Use that $t-\log(1+t)\geqslant |t|-\log(1+|t|)$ for any $t>-1$ together with the fact that $t-\log(1+t)$ is an increasing function for $t\geqslant 0$ to argue that

$$\begin{aligned} &\iint\bigg(\frac{\psi(x)-\psi(y)}{x-y}-\log\bigg(1+\frac{\psi(x)-\psi(y)}{x-y}\bigg)\bigg)\,\mu_V(dx)\,\mu_V(dy) \\ &\quad\geqslant \iint\bigg(\frac{|\psi(x)-\psi(y)|}{b-a}-\log\bigg(1+\frac{|\psi(x)-\psi(y)|}{b-a}\bigg)\bigg)\,\mu_V(dx)\,\mu_V(dy). \end{aligned} \tag{2.16}$$

Further, for $s\geqslant 0$ and $u,v>0$ we have

$$us^2+s-\log(1+s)\geqslant \begin{cases} \frac{v-\log(1+v)}{v^2}s^2, & 0\leqslant s\leqslant v \\ us^2, & v\leqslant s \end{cases} \geqslant \min\bigg\{u,\frac{v-\log(1+v)}{v^2}\bigg\}s^2.$$

This inequality used for $u=\frac{\rho(b-a)^2}{128}$ and $v=\frac{2R}{b-a}$ in combination with (2.15) and (2.16) yields for the choice of $c=\min\{u,(v-\log(1+v))/v^2\}$ that

$$\frac{\rho}{16}\int\limits_{|\psi|\geqslant R}\psi^2(x)\,\mu_V(dx)+\iint\left(\frac{\psi(x)-\psi(y)}{x-y}-\log\left(1+\frac{\psi(x)-\psi(y)}{x-y}\right)\right)\mu_V(dx)\,\mu_V(dy)$$
$$\geqslant c\iint\big(\psi(x)-\psi(y)\big)^2\,\mu_V(dx)\,\mu_V(dy)$$
$$=c\left[\int\psi^2(x)\,\mu_V(dx)-\left(\int\psi(x)\mu_V(dx)\right)^2\right]. \tag{2.17}$$

This shows that $E(\mu)-E(\mu_V)$ is bounded below by a constant times the variance of $\psi$. Notice that $W_2^2(\mu,\mu_V)=\int\psi^2(x)\mu_V(dx)$ and in order to complete the proof we have to replace the variance of $\psi$ by the integral of $\psi^2$ with respect to $\mu_V$. This boils down to estimating the $\mu_V$ integral of $\psi$ in terms of the integral of $\psi^2$.

To this end, use Cauchy's inequality:

$$\left(\int\psi(x)\,\mu_V(dx)\right)^2\leqslant\int\psi^2(x)\left(1+\frac{1}{2c}\int\limits_0^1V''\big(x+\tau\psi(x)\big)(1-\tau)\,d\tau\right)\mu_V(dx)$$
$$\times\int\frac{1}{1+\frac{1}{2c}\int_0^1V''(x+\tau\psi(x))(1-\tau)\,d\tau}\,\mu_V(dx).$$

This inequality combined with Eqs. (2.12), (2.14) and (2.17), results with

$$E(\mu)-E(\mu_V)\geqslant\int\psi^2(x)\left(c+\frac{1}{2}\int\limits_0^1V''\big(x+\tau\psi(x)\big)(1-\tau)\,d\tau\right)\mu_V(dx)$$
$$\times\int\frac{\int_0^1V''(x+\tau\psi(x))(1-\tau)d\tau}{2c+\int_0^1V''(x+\tau\psi(x))(1-\tau)\,d\tau}\,\mu_V(dx)$$
$$\geqslant c\int\frac{\int_0^1V''(x+\tau\psi(x))(1-\tau)\,d\tau}{2c+\int_0^1V''(x+\tau\psi(x))(1-\tau)\,d\tau}\,\mu_V(dx)W_2^2(\mu,\mu_V),$$

where here we used the convexity encoded into $V''\geqslant 0$ and the fact that $W_2^2(\mu,\mu_V)=\int\psi^2(x)\mu_V(dx)$ to get the lower bound of the first integral.

From the previous inequality, it becomes clear that we are done as soon as we prove that the quantity in front of $W_2^2(\mu,\mu_V)$ is bounded from below by a positive constant uniformly in $\psi$. To carry this out, notice that $V''$ can not be identically zero on $[a,b]$. Indeed, if $V''$ were identically zero on $[a,b]$, then we would have that $V'(x)=K$ for all $x\in[a,b]$, and this plugged into Eq. (2.4), yields that $K(b-a)=2$ and $K(b-a)=-2$, a system without a solution. Therefore $V''$ is not identically 0 on $[a,b]$. If $|\psi(x)|>R$, then $V''(x+\tau\psi(x))\geqslant\rho$ for $1/2\leqslant\tau<1$, which implies $\int_0^1V''(x+\tau\psi(x))(1-\tau)d\tau\geqslant\rho/8$. On the other hand, if $|\psi(x)|\leqslant R$, then

$$\int\limits_0^1V''\big(x+\tau\psi(x)\big)(1-\tau)\,d\tau\geqslant\int\limits_0^\delta V''\big(x+\tau\psi(x)\big)(1-\tau)\,d\tau\geqslant\frac{\delta}{2}\inf_{|y-x|\leqslant\delta R}V''(y)$$

for all $0 \leqslant \delta \leqslant 1$. Define

$$w(x) = \sup_{\delta \in [0,1]} \min\left\{\frac{\rho}{8}, \frac{\delta}{2} \inf_{|y-x| \leqslant \delta R} V''(y)\right\}.$$

Since $V''$ is not identically 0 on $[a,b]$, it follows that $w$ is not identically zero on $[a,b]$. With this we obtain that

$$\int_0^1 V''\big(x + \tau\psi(x)\big)(1-\tau)\,d\tau \geqslant w(x) \geqslant 0,$$

and then that

$$c\int \frac{\int_0^1 V''(x+\tau\psi(x))(1-\tau)\,d\tau}{2c + \int_0^1 V''(x+\tau\psi(x))(1-\tau)\,d\tau}\,\mu_V(dx) \geqslant C = \int \frac{cw(x)}{2c+w(x)}\,\mu_V(dx) > 0$$

which finishes the proof of (2.10) with this choice of $C$.

3. For the inequality (2.11), we follow the same route as in the proof of (2.9), the only change this time being that $V(x) - \rho|x|^p$ is convex, and thus we obtain

$$V(y) - V(x) - V'(x)(y-x) \geqslant \rho\big(|y|^p - |x|^p - p\,\mathrm{sign}(x)|x|^{p-1}(y-x)\big). \tag{2.18}$$

Writing $\theta(x) = x + \psi(x)$, and using (2.12) together with $a - 1 \geqslant \log(a)$ for $a \geqslant 0$, one arrives at

$$E(\mu) - E(\mu_V) \geqslant \rho \int \big(|x+\psi(x)|^p - |x|^p - p\,\mathrm{sign}(x)|x|^{p-1}\psi(x)\big)\,\mu_V(dx).$$

Now we use the fact that for all $a, b \in \mathbb{R}$,

$$|a+b|^p - |b|^p - p\,\mathrm{sign}(b)|b|^{p-1}a \geqslant c_p|a|^p, \tag{2.19}$$

which applied to the above inequality in conjunction to (2.8), yields inequality (2.11). □

**Remark 1.**

1. The $C^2$ regularity of $V$ for (2.9) can be dropped (see [26]) but to simplify the presentation here we decided to consider only this case.
2. If $V(x) - \rho|x|^p$ is convex, then using inequalities (2.11), (2.10) and Young's inequality we obtain that for any $2 \leqslant k \leqslant p$, there exists a constant $c = c(k, p, \rho, \mu_V, V)$ such that

$$cW_k^k(\mu, \mu_V) \leqslant E(\mu) - E(\mu_V).$$

3. We want to point out that the inequalities (2.11) and (2.10) are somehow complementary to each other. For example, if we take $V(x) = \rho|x|^p$ with $p > 1$ and the measure $\mu = \theta_{\#}\mu_V$ for $\theta(x) = x + m$, then Eq. (2.11) takes the form

$$c_p m^p \leqslant \int \bigl(|x+m|^p - |x|^p\bigr)\,\mu_V(dx) \tag{2.20}$$

while Eq. (2.10) becomes

$$Cm^2 \leqslant \int \bigl(|x+m|^p - |x|^p\bigr)\,\mu_V(dx),$$

which, because it is easy to check that $\mu_V$ is symmetric, is the same as

$$Cm^2 \leqslant \int \bigl(|x+m|^p - |x|^p - p\,\mathrm{sign}(x)|x|^{p-1}m\bigr)\,\mu_V(dx). \tag{2.21}$$

Notice here that (2.20) is in the right scale for large $m$ as (2.21) is in the right scale for $m$ close to 0, because in this case the integrand is of the size $m^2$. It seems that Talagrand's transportation inequality in this context has two aspects, one is the large $W_p(\mu, \mu_V)$ which is dictated by the potential $V$ for large values and results with Eq. (2.11) and the small $W_2(\mu, \mu_V)$ regime which is dictated by the repulsion effect of the logarithm and results with Eq. (2.10).
4. It is not clear whether inequality (2.10) still holds for the case of a potential $V$ which is not convex. Of interest would be the particular case $V(x) = ax^4 + bx^2$ for some $a > 0$ and $b < 0$. This example actually raises the question of the stability of transportation inequality under bounded perturbations.
5. Very likely the constant $c_p$ in (2.11) is not sharp.

## 3. Potential independent transportation inequalities

In this section, we investigate some potential independent transportation inequalities. A transportation inequality in the form of (2.10) can not possibly hold without a quadratic growth at infinity. Also, the proof of (2.10) might lead to the conclusion that the logarithmic term plays a more important role. Therefore the natural question one may ask is whether there is a manifestation of this fact in some sort of transportation type inequality which is independent of the potential involved. The main question reduces to hint some appropriate distance one needs to use to replace the Wasserstein distance in Theorem 2. We investigate in this section several possibilities, starting with the free version of the classical Pinsker's inequality.

The Pinsker's inequality classically states that (cf. [10] and [21])

$$2\|\mu - \nu\|_v^2 \leqslant E(\mu \mid \nu) \quad \text{for any } \mu, \nu \text{ probability measures on } \mathbb{R},$$

where $\|\mu - \nu\|_v$ is the total variation distance between $\mu$ and $\nu$ and $E(\mu \mid \nu)$ is the relative entropy between $\mu$ and $\nu$. This in particular shows that if $\mu_n$ convergence to $\mu$ in entropy, then $\mu_n$ converges to $\mu$ is a very strong sense.

The same natural question can be posed in the logarithmic entropy context. For a given potential $V$, is there an inequality of the form

$$C\|\mu-\mu_V\|_v^2 \leqslant E(\mu)-E(\mu_V)$$

for a given constant $C>0$ and any probability distribution $\mu$ on $\mathbb{R}$?

It turns out that these inequalities do not hold for the logarithmic energy. In fact, we will show that even a weaker inequality of the form

$$C|F_\mu-F_{\mu_V}|_u^2 \leqslant E(\mu)-E(\mu_V) \tag{3.1}$$

does not hold, where $F_\mu$ denotes the cumulative function of a probability measure $\mu$ on the line. Even though the uniform distance does not have the same widespread use in probability it appears for example in the Berry–Esseen type estimates for the convergence in the central limit theorem. This is the reason why we consider this distance as the first next best candidate wherever the total variation fails. Clearly this metric gives a stronger topology as the topology of weak convergence.

We will construct a counterexample to (3.1) in the case of $V(x)=2x^2$, for which the equilibrium measure is

$$\mu_V(dx)=\mathbb{1}_{[-1,1]}(x)\frac{2\sqrt{1-x^2}}{\pi}\,dx,$$

the semicircular law on $[-1,1]$. Consider now the sequence

$$\mu_n(dx)=\mathbb{1}_{[-1,1]}(x)\frac{2\sqrt{1-x^2}}{\pi}\,dx+\frac{\sum_{k=2}^{2n-1}(-1)^k T_{2k+1}(x)}{4(n^2-1)\pi\sqrt{1-x^2}}\,dx$$

where $T_k$ is the $k$th Chebyshev polynomial of the first kind. With these choices we have that

$$E(\mu_n)-E(\mu_V)\leqslant\frac{\pi^2}{\log(n/3)}|F_{\mu_n}-F_{\mu_V}|_u^2 \quad \text{for all } n\geqslant 4. \tag{3.2}$$

Let us point out that $\mu_n$ is indeed a probability measure. This requires a little proof but it is entirely elementary and is left to the reader.

To prove (3.1), notice that since the support of $\mu_n$ is the same as the support of $\mu_V$, we have from (2.2) that

$$E(\mu_n)-E(\mu_V)=-\iint\log|x-y|(\mu_n-\mu_V)(dx)(\mu_n-\mu_V)(dy). \tag{3.3}$$

Next remark that $\mu_n=\cos_\#(f_n\lambda)$ and $\mu_V=\cos_\#(g\lambda)$, where $\lambda$ is the Lebesgue measure on $[0,\pi]$ and

$$f_n(t)=\frac{1-\cos(2t)}{\pi}+\frac{1}{4\pi(n^2-1)}\sum_{k=2}^{2n-1}(-1)^k\cos\big((2k+1)t\big), \qquad g(t)=\frac{1-\cos(2t)}{\pi}$$

and further

$$
\begin{aligned}
&-\iint \log|x-y|(\mu_n-\mu_V)(dx)(\mu_n-\mu_V)(dy)\\
&\quad = -\int_0^\pi\int_0^\pi \log|\cos t-\cos s|h_n(t)h_n(s)\,dt\,ds \quad \text{where } h_n=f_n-g.
\end{aligned}
$$

Now we provide a formula for the logarithmic energy we learnt from [15] and have not seen it elsewhere. Here is a quick description. Write first $\cos t=(e^{it}+e^{-it})/2$ and $\cos s=(e^{is}+e^{-is})/2$ so $|\cos t-\cos s|=|(e^{it}+e^{-it})/2-(e^{is}+e^{-is})|/2=|1-e^{i(t+s)}||1-e^{i(t-s)}|/2$ and so, for $t\neq s$, and $t$ or $s$ not equal to $\pi$,

$$
\begin{aligned}
\log|\cos t-\cos s| &= -\log 2+\operatorname{Re}\big(\log\big(1-e^{i(t+s)}\big)+\log\big(1-e^{i(t-s)}\big)\big)\\
&= -\log 2-\sum_{\ell=1}^{\infty}\operatorname{Re}\big(e^{i\ell(t+s)}/\ell+e^{i\ell(t-s)}/\ell\big)\\
&= -\log 2-\sum_{\ell=1}^{\infty}\frac{2}{\ell}\cos(\ell t)\cos(\ell s).
\end{aligned}
$$

From this, one gets to

$$
-\int_0^\pi\int_0^\pi \log|\cos t-\cos s|h_n(t)h_n(s)\,dt\,ds=\sum_{\ell=1}^{\infty}\frac{2}{\ell}\left(\int_0^\pi \cos(\ell t)h_n(t)\,dt\right)^2. \tag{3.4}
$$

But now,

$$
\begin{aligned}
\int_0^\pi \cos(\ell t)h_n(t)\,dt &= \frac{1}{4\pi(n^2-1)}\sum_{k=2}^{2n-1}(-1)^k\int_0^\pi \cos(\ell t)\cos\big((2k+1)t\big)\,dt\\
&= \begin{cases} \frac{(-1)^{(\ell-1)/2}}{8(n^2-1)}, & 4\leqslant \ell\leqslant 4n \text{ and odd}\\ 0 & \text{otherwise}\end{cases}
\end{aligned}
$$

and thus

$$
\begin{aligned}
-\int_0^\pi\int_0^\pi \log|\cos t-\cos s|h_n(t)h_n(s)\,dt\,ds &= \sum_{\ell=1}^{\infty}\frac{2}{\ell}\left(\int_0^\pi \cos(\ell t)h_n(t)\,dt\right)^2\\
&= \frac{1}{32(n^2-1)^2}\sum_{\ell=2}^{2n-1}\frac{1}{2\ell+1}. \qquad (3.5)
\end{aligned}
$$

On the other hand $|F_{\mu_n}-F_{\mu_V}|_u=|F_{f_n\lambda}-F_{g\lambda}|_u=\sup_{x\in[0,\pi]}|\int_0^x h_n(t)\,dt|$ and

$$
\int_0^x h_n(t)\,dt=\frac{1}{4\pi(n^2-1)}\sum_{\ell=2}^{2n-1}\frac{(-1)^\ell\sin((2\ell+1)x)}{2\ell+1},
$$

from which for $x = \pi/4$, we obtain

$$|F_{\mu_n} - F_{\mu_V}|_u = \sup_{x\in[0,\pi]} \left| \int_0^x h_n(t)\,dt \right| \geqslant \frac{1}{4\pi(n^2-1)} \sum_{\ell=2}^{2n-1} \frac{1}{2\ell+1}. \tag{3.6}$$

Combining (3.5) and (3.6) we get

$$\frac{\pi^2}{2\sum_{\ell=2}^{2n-1}\frac{1}{2\ell+1}} |F_{\mu_n} - F_{\mu_V}|_u^2 \geqslant -\iint \log|x-y|(\mu_n-\mu_V)(dx)(\mu_n-\mu_V)(dy) \tag{3.7}$$

which together with the fact that $\sum_{\ell=2}^{2n-1}\frac{1}{2\ell+1} \geqslant \frac{1}{2}\log(n/3)$ for $n \geqslant 4$ and (3.3), we finally arrive at (3.2).

The example shown above has the property that $E(\mu_n) - E(\mu_V)$ converges to 0 when $n$ goes to infinity, and also that $|F_{\mu_n} - F_{\mu_V}|_u$ converges to zero. Despite the fact that (3.1) does not hold, we will see below in Corollary 1 that if $E(\mu_n) - E(\mu_V)$ converges to 0, then $|F_{\mu_n} - F_{\mu_V}|_u$ always converges to 0.

We consider now a weak form of (3.1). To do this we define the distance

$$d(\mu,\nu) = \sup_{a,b\in\mathbb{R}} \left| \int e^{-|ax+b|}\,\mu(dx) - \int e^{-|ax+b|}\,\nu(dx) \right|. \tag{3.8}$$

With this definition we have the following result.

**Theorem 3.** *For any potential $V$ satisfying* (2.1)*, we have that for any compactly supported measure $\mu$,*

$$4\pi^3 d^2(\mu,\mu_V) \leqslant E(\mu) - E(\mu_V). \tag{3.9}$$

**Proof.** Using Eqs. (2.1) and (2.2), we get for any compactly supported measure $\mu$ with $E(\mu)$ finite,

$$E(\mu) - E(\mu_V) \geqslant -\iint \log|x-y|(\mu-\mu_V)(dx)(\mu-\mu_V)(dy).$$

We will prove that for any measures $\mu$ and $\nu$ with compact support such that $-\iint \log|x-y|\,\mu(dx)\,\mu(dy) < \infty$ and $-\iint \log|x-y|\,\nu(dx)\,\nu(dy) < \infty$, we have that

$$4\pi^3 d^2(\mu,\nu) \leqslant -\iint \log|x-y|(\mu-\nu)(dx)(\mu-\nu)(dy), \tag{3.10}$$

which shows that (3.10) implies (3.9).

Now we use [11, Eq. (6.45)] to write

$$-\iint \log|x-y|(\mu-\mu_V)(dx)(\mu-\mu_V)(dy) = \int_0^\infty \frac{|\hat{\mu}(t) - \hat{\mu}_V(t)|^2}{t}\,dt \tag{3.11}$$

where the hat stands for the Fourier transform, and continue with

$$\int_0^\infty \frac{|\hat{\mu}(t)-\hat{\nu}(t)|^2}{t}\,dt = \frac{1}{2}\int_{-\infty}^{\infty} \frac{|\hat{\mu}(t)-\hat{\nu}(t)|^2}{|t|}\,dt \geqslant |a| \int_{-\infty}^{\infty} \frac{|\hat{\mu}(t)-\hat{\nu}(t)|^2}{a^2+t^2}\,dt$$
$$\geqslant \frac{a^2}{\pi}\left|\int_{-\infty}^{\infty} \frac{(\hat{\mu}(t)-\hat{\nu}(t))e^{-ict}}{a^2+t^2}\,dt\right|^2$$

for any $a, c \in \mathbb{R}$ with $a \neq 0$. Further, using the inversion formula for the Fourier transform, one has

$$\int_{-\infty}^{\infty} \frac{(\hat{\mu}(t)-\hat{\nu}(t))e^{-ict}}{a^2+t^2}\,dt = 2\pi \int \hat{\phi}(x)(\mu-\nu)(dx) = \frac{2\pi^2}{|a|}\int e^{-|a(x+c)|}(\mu-\nu)(dx) \tag{3.12}$$

because for $\phi(t) = \frac{e^{ict}}{a^2+t^2}$,

$$\hat{\phi}(x) = \int \frac{e^{i(x+c)t}}{t^2+a^2}\,dt = \frac{\pi e^{-|a(x+c)|}}{|a|}.$$

From here, (3.10) follows immediately. □

**Remark 2.** From Eq. (3.11) it seems that the distance one should consider should be the Sobolev norm with exponent $-1/2$. This is another possible candidate to the role of $d$ played here, however not always finite. We chose the metric $d$ as it's definition is somehow close to uniform norm of the difference of the Laplace transforms of the measures. It is also always defined and bounded by 1, thus resembling the total variation distance.

The next result is collecting facts about how strong the topology induced by $d$ is.

**Proposition 1.**

1. *$d$ is a distance on $\mathcal{P}(\mathbb{R})$ and if $d(\mu_n, \mu) \underset{n\to\infty}{\longrightarrow} 0$, then $\mu_n \to_{n\to\infty} \mu$ in the weak topology. In addition $d(\delta_a, \delta_b) = 1$ for $a \neq b$, thus the topology induced by $d$ is strictly stronger than the weak convergence topology.*
2. *For any two probability measures $\mu$ and $\nu$,*

$$d(\mu,\nu) \leqslant 2|F_\mu - F_\nu|_u. \tag{3.13}$$

3. *If $V$ satisfies condition* (2.1)*, then $E_V(\mu_n) \underset{n\to\infty}{\longrightarrow} E_V(\mu_V)$ implies $|F_{\mu_n} - F_{\mu_V}|_u \underset{n\to\infty}{\longrightarrow} 0$.*

**Proof.** 1. To prove that $d$ is a distance the only non trivial fact is that for two probability measures $\mu$ and $\nu$, $d(\mu,\nu)=0$ implies $\mu=\nu$. Thus from Eq. (3.12), we obtain for $a=1$ that for all $c\in\mathbb{R}$,

$$\int_{-\infty}^{\infty}\frac{(\hat{\mu}(t)-\hat{\nu}(t))e^{-ict}}{1+t^2}\,dt=0.$$

Since this holds true for any $c\in\mathbb{R}$, it implies that the Fourier transform of the function $t\to\frac{\hat{\mu}(t)-\hat{\nu}(t)}{1+t^2}$ is 0, which means that the function in discussion must be 0. This means that $\hat{\mu}=\hat{\nu}$, or equivalently that $\mu=\nu$.

Let $\mathcal{L}(\mu,\nu)$ stand for the Levy distance which induces the weak topology on $\mathcal{P}(\mathbb{R})$. Let $d(\mu_n,\mu)\underset{n\to\infty}{\longrightarrow}0$. Assume now that there exists $\epsilon>0$ and a subsequence such that $\mathcal{L}(\mu_{n_k},\mu)\geqslant\epsilon$. Otherwise said, the sequence $\mu_n$ has a subsequence which is not convergent to $\mu$. Since, we are dealing with probability measures, there is a subsequence $\mu_{n_{k_l}}$ which is vaguely convergent to a measure $\nu$ with total mass less than 1. This means that for any continuous function $\phi$ which is vanishing at infinity, we have that

$$\int\phi\,d\mu_{n_{k_l}}\underset{l\to\infty}{\longrightarrow}\int\phi\,d\nu.$$

We can apply this for functions $\phi(x)=e^{-|ax+b|}$ where $a\neq0$ and infer that

$$\int e^{-|ax+b|}\mu_{n_{k_l}}(dx)\underset{l\to\infty}{\longrightarrow}\int e^{-|ax+b|}\nu(dx)\quad\text{for all } a\neq0,\ b\in\mathbb{R}.$$

On the other hand, because $d(\mu_{n_{k_l}},\mu)\underset{l\to\infty}{\longrightarrow}0$, these considerations result with

$$\int e^{-|ax+b|}\mu(dx)=\int e^{-|ax+b|}\nu(dx)\quad\text{for all } a\neq0,\ b\in\mathbb{R}.$$

Further, using the dominated convergence for $b=0$ and $a\to0$, we obtain that $\nu$ is a probability measure. From the discussion at the beginning of this proof, it also follows that $\nu=\mu$ and this in turn results with $\mu_{n_{k_l}}$ being weakly convergent to $\mu$, a contradiction. This proves that the convergence in the metric $d$ implies weak convergence.

It is obvious that $d(\mu,\nu)\leqslant1$ for any measures $\mu$ and $\nu$. For the case of discrete measures, we also have that $1\geqslant d(\delta_a,\delta_b)\geqslant\int e^{-\alpha|x-a|}\delta_a(dx)-\int e^{-\alpha|x-a|}\delta_b(dx)$ for any $\alpha>0$, which yields that $1\geqslant d(\delta_a,\delta_b)\geqslant1-e^{-\alpha|b-a|}$ for all $\alpha>0$. Letting $\alpha\to\infty$, we get that $d(\delta_a,\delta_b)=1$ for $a\neq b$ which shows that convergence in $d$ is strictly stronger than convergence in the weak topology.

2. From the fact that for any finite positive measure $\mu$,

$$\int_{(0,\infty)}\big(1-e^{-\alpha y}\big)\mu(dx)=\int_{(0,\infty)}\alpha e^{-\alpha y}\mu\big((y,\infty)\big)\,dy,$$

we deduce that

$$\int e^{-\alpha|x-a|}(\mu-\nu)(dx) = \int_{(0,\infty)} \alpha e^{-\alpha y}\big[F_\mu(a-y) - F_\mu(a+y) - F_\nu(a-y) + F_\nu(a+y)\big]\,dy$$

which easily yields (3.13).

3. We actually show that if $\mu_n$ and $\mu$ are compactly supported probability measures such that

$$-\iint \log|x-y|\,\mu(dx)\mu(dy) < \infty, \qquad -\iint \log|x-y|\,\mu_n(dx)\,\mu_n(dy) < \infty$$

and

$$\lim_{n\to\infty} \iint \log|x-y|(\mu_n-\mu)(dx)(\mu_n-\mu)(dy) = 0,$$

then $|F_{\mu_n} - F_\mu|_u \underset{n\to\infty}{\longrightarrow} 0$. From (3.10) and the first part, we obtain that $\mu_n$ converges weakly to $\mu$. In addition, none of the measures $\mu_n$ or $\mu$ have atoms. Thus $F_{\mu_n}$ and $F_\mu$ are continuous functions which combined with the weak convergence implies that $F_{\mu_n}$ converges pointwise to $F_\mu$. Since the functions $F_{\mu_n}$ and $F_\mu$ are distributions of probability measures, it is an easy matter to check that the convergence is actually uniform. $\square$

**Remark 3.** We do not know if the topology of convergence in $d$ is the same as the one defined by the metric $|F_\mu - F_\nu|_u$.

This result might leave one wondering if a stronger convergence takes place. In other words, is it true that $E_V(\mu_n) \underset{n\to\infty}{\longrightarrow} E_V(\mu_V)$ implies $\|\mu_n - \mu_V\|_v \underset{n\to\infty}{\longrightarrow} 0$? To this end, we can consider $V(x) = \log|\frac{|x|+\sqrt{x^2-1}}{2}|$ and notice (see [27, p. 46]) that $\mu_V$ is the arcsine law of $[-1,1]$. Thus if we consider

$$\mu_V(dx) = \mathbb{1}_{[-1,1]}(x)\frac{dx}{\pi\sqrt{1-x^2}}, \qquad \mu_n(dx) = \mathbb{1}_{[-1,1]}(x)\frac{(1-T_n(x))dx}{\pi\sqrt{1-x^2}},$$

then, using the same argument which led us to (3.4), with $h_n$ there replaced by $h_n(x) = \cos(nx)$ here, one arrives at $E(\mu_n) - E(\mu_V) = \frac{1}{n}$ while the total variation distance is $\|\mu_n - \mu_V\|_v \geqslant 1/4$.

## 4. Log-Sobolev inequality

In this section, we develop similarly the mass transportation method to prove the Log-Sobolev inequality in the free context. Note again that, as discussed in the introduction, the first assertion for strictly convex potentials was initially proved by large matrix approximation in [3].

Before we state the main result, we define inspired by Voiculescu [31], the relative free Fisher information as

$$I(\mu) = \int \big(H\mu(x) - V'(x)\big)^2\mu(dx) \quad \text{with } H\mu(x) = \text{p.v.}\int \frac{2}{x-y}\,\mu(dy). \tag{4.1}$$

for measures $\mu$ on $\mathbb{R}$ which have density $p = d\mu/dx$ in $L^3(\mathbb{R})$. In this case the principal value integral is a function in $L^3$. Otherwise we let $I(\mu)$ be equal to $+\infty$.

**Theorem 4** *(Log-Sobolev).*

1. *If $V$ is $C^2$ and $V(x) - \rho x^2$ is convex for some $\rho > 0$, then for any probability measure $\mu$ on $\mathbb{R}$,*

$$E(\mu) - E(\mu_V) \leqslant \frac{1}{4\rho} I(\mu). \tag{4.2}$$

*Equality is attained for the case $V(x) = \rho x^2$ and $\mu = \theta_{\#}\mu_V$, where $\theta(x) = x + m$. Thus the inequality* (4.2) *is sharp for translations of $\mu_V$.*

2. *If $V$ is $C^2$ and $V(x) - \rho|x|^p$ is convex for some $\rho > 0$ and $p > 1$, then for any probability measure $\mu$ on $\mathbb{R}$,*

$$E(\mu) - E(\mu_V) \leqslant \frac{k_p}{\rho^{q/p}} I_q(\mu) \quad \text{where } I_q(\mu) = \int \bigl|H\mu(x) - V'(x)\bigr|^q \mu(dx) \tag{4.3}$$

*where here $q$ is the conjugate of $p$ i.e. $1/q + 1/p = 1$ and the constant $k_p = (pc_p)^{q/p}/q$, with $c_p$ from* (2.11).

**Proof.** 1. We will assume that the measure $\mu$ has a smooth compactly supported density as the general case follows via approximation arguments discussed in details in [18]. Take the (increasing) transport map $\theta$ from $\mu_V$ into $\mu$. We write the inequality (4.2) in the following equivalent way

$$\begin{aligned}
&\frac{1}{4\rho}\int \bigl(H\mu\bigl(\theta(x)\bigr) - V'\bigl(\theta(x)\bigr)\bigr)^2 \mu_V(dx) \\
&\quad + \int \bigl(V(x) - V\bigl(\theta(x)\bigr) - V'\bigl(\theta(x)\bigr)\bigl(x - \theta(x)\bigr)\bigr)\, \mu_V(dx) \\
&\quad - \int \bigl(H\mu\bigl(\theta(x)\bigr) - V'\bigl(\theta(x)\bigr)\bigr)\bigl(x - \theta(x)\bigr)\, \mu_V(dx) \\
&\quad + \int H\mu\bigl(\theta(x)\bigr)\bigl(x - \theta(x)\bigr)\, \mu_V(dx) - \iint \log\frac{x-y}{\theta(x) - \theta(y)}\, \mu_V(dx)\, \mu_V(dy) \geqslant 0.
\end{aligned} \tag{4.4}$$

Notice now that from the convexity of $V(x) - \rho x^2$, one obtains that

$$\begin{aligned}
V(x) - V\bigl(\theta(x)\bigr) - V'\bigl(\theta(x)\bigr)\bigl(x - \theta(x)\bigr) &\geqslant \rho\bigl(x^2 - \theta(x)^2 - 2\theta(x)\bigl(x - \theta(x)\bigr)\bigr) \\
&= \rho\bigl(x - \theta(x)\bigr)^2.
\end{aligned} \tag{4.5}$$

Now,

$$\begin{aligned}
\int H\mu\bigl(\theta(x)\bigr)\bigl(x - \theta(x)\bigr)\mu_V(dx) &= \int \bigl(x - \theta(x)\bigr) \int \frac{2}{\theta(x) - \theta(y)}\, \mu_V(dy)\, \mu_V(dx) \\
&= \iint \left(\frac{x-y}{\theta(x) - \theta(y)} - 1\right) \mu_V(dx)\, \mu_V(dy)
\end{aligned} \tag{4.6}$$

where one has to interpret the second integral here in the principal value sense, however since $\theta$ is increasing, the last integral is actually taken in the Lebesgue sense.

Using these, Eq. (4.4) may be rewritten as

$$\frac{1}{4\rho}\int\big[H\mu\big(\theta(x)\big)-V'\big(\theta(x)\big)-2\rho\big(x-\theta(x)\big)\big]^2\mu_V(dx)$$
$$+\iint\bigg(\frac{x-y}{\theta(x)-\theta(y)}-1-\log\frac{x-y}{\theta(x)-\theta(y)}\bigg)\mu_V(dx)\,\mu_V(dy)\geqslant 0$$

which is seen to hold since $u-1-\log(u)\geqslant 0$ for $u\geqslant 0$.

Equality is attained for the case $V(x)=\rho x^2$ and $\theta(x)=x+c$, which corresponds to the translations of the measure $\mu_V$.

2. With the same arguments used in the above proof and the proof of Theorem 2, we use Eqs. (2.18) and (2.19) to argue that

$$\begin{aligned}&\frac{k_p}{\rho^{q/p}}\int\big|H\mu(x)-V'(x)\big|^q\mu(dx)-E(\mu)+E(\mu_V)\\&\quad\geqslant\int\bigg[\frac{k_p}{\rho^{q/p}}\big|H\mu\big(\theta(x)\big)-V'\big(\theta(x)\big)\big|^q\\&\qquad+\big(V'\big(\theta(x)\big)-H\mu\big(\theta(x)\big)\big)\big(x-\theta(x)\big)+c_p\rho\big|x-\theta(x)\big|^p\bigg]\mu_V(dx)\\&\qquad+\iint\bigg(\frac{x-y}{\theta(x)-\theta(y)}-1-\log\frac{x-y}{\theta(x)-\theta(y)}\bigg)\mu_V(dx)\,\mu_V(dy)\\&\quad\geqslant 0\end{aligned}$$

where we used Young's inequality $a^q/q+b^p/p\geqslant ab$ for $a,b\geqslant 0$ and the constant $k_p=(pc_p)^{q/p}/q$. $\square$

**Remark 4.** It was proved in [24] that a Log-Sobolev inequality always implies a transportation inequality.

## 5. HWI Inequality

This section is devoted to the free analog of the HWI inequality of Otto and Villani [25] in the classical context, connecting thus the (free) entropy, Wasserstein distance and Fisher information. As we will see, the HWI implies the Log-Sobolev inequality for strictly convex potentials. This free HWI inequality was not considered before, and in particular it is not clear whether there is a random matrix proof, delicate points involving the Wasserstein distance entering into the proof.

**Theorem 5** *(HWI inequality).*

1. *Assume that $V$ is $C^2$ such that for some $\rho\in\mathbb{R}$, $V(x)-\rho x^2$ is convex. Then, for any measure $\mu\in\mathcal{P}(\mathbb{R})$,*

$$E(\mu)-E(\mu_V)\leqslant\sqrt{I(\mu)}W_2(\mu,\mu_V)-\rho W_2^2(\mu,\mu_V).\tag{5.1}$$

*In the case $V(x) = \rho x^2$, the inequality is sharp.*

2. *If $V$ is $C^2$ and $V(x) - \rho|x|^p$ is convex for some $\rho \geqslant 0$ and $p > 1$, then for the same constant $c_p$ appearing in Theorem 2, we have that*

$$E(\mu) - E(\mu_V) \leqslant I_q^{1/q}(\mu) W_p(\mu, \mu_V) - \rho c_p W_p^p(\mu, \mu_V), \tag{5.2}$$

*where $1/p + 1/q = 1$.*

**Proof.** 1. We employ here the notations used in Theorem 4 and we will give a proof of the inequality for the case of a measure $\mu$ with smooth and compactly supported density, the general case follows through careful approximations pointed in [18]. The inequality to be proved can be restated as (5.3) + (5.4) + (5.5) $\geqslant 0$, where

$$\begin{aligned}(5.3) = {} & \left( \int \big(H\mu\big(\theta(x)\big) - V'\big(\theta(x)\big)\big)^2 \mu_V(dx) \int \big(\theta(x) - x\big)^2 \mu_V(dx) \right)^{1/2} \\ & - \int \big(H\mu\big(\theta(x)\big) - V'\big(\theta(x)\big)\big)\big(x - \theta(x)\big)\, \mu_V(dx)\end{aligned} \tag{5.3}$$

$$(5.4) = \int \big[V(x) - V\big(\theta(x)\big) - V'\big(\theta(x)\big)\big(x - \theta(x)\big) - \rho\big(\theta(x) - x\big)^2\big] \mu_V(dx) \tag{5.4}$$

$$(5.5) = \int H\mu\big(\theta(x)\big)\big(x - \theta(x)\big)\, \mu_V(dx) - \iint \log \frac{x - y}{\theta(x) - \theta(y)}\, \mu_V(dx)\, \mu_V(dy). \tag{5.5}$$

A simple application of Cauchy's inequality shows that (5.3) $\geqslant 0$. Using convexity of $V(x) - \rho x^2$ we have from Eq. (4.5), that (5.4) $\geqslant 0$. Finally, using (4.6), we have that

$$(5.5) = \iint \left( \frac{x - y}{\theta(x) - \theta(y)} - 1 - \log \frac{x - y}{\theta(x) - \theta(y)} \right) \mu_V(dx)\, \mu_V(dy) \geqslant 0,$$

which finishes the proof of (5.1). For the case $V(x) = \rho x^2$, we have equality if $\theta(x) = x + m$.

2. The inequality we want to prove is equivalent to the statement that (5.6) + (5.7) + (5.8) $\geqslant 0$, where

$$\begin{aligned}(5.6) = {} & \left| \int \big|H\mu\big(\theta(x)\big) - V'\big(\theta(x)\big)\big|^q \mu_V(dx) \right|^{1/q} \left| \int \big(\theta(x) - x\big)^p \mu_V(dx) \right|^{1/p} \\ & - \int \big(H\mu\big(\theta(x)\big) - V'\big(\theta(x)\big)\big)\big(x - \theta(x)\big)\, \mu_V(dx)\end{aligned} \tag{5.6}$$

$$(5.7) = \int \big[V(x) - V\big(\theta(x)\big) - V'\big(\theta(x)\big)\big(x - \theta(x)\big) - \rho c_p \big|\theta(x) - x\big|^p\big] \mu_V(dx) \tag{5.7}$$

$$(5.8) = \int H\mu\big(\theta(x)\big)\big(x - \theta(x)\big)\, \mu_V(dx) - \iint \log \frac{x - y}{\theta(x) - \theta(y)}\, \mu_V(dx)\, \mu_V(dy). \tag{5.8}$$

Now, (5.6) is non-negative thanks to Hölder's inequality, Eq. (5.7), follows from the convexity of $V(x) - \rho|x|^p$ and the combination of (2.18) and (2.19), while Eq. (5.8) is the same as (5.5). □

As pointed out in [25], HWI inequalities for $\rho > 0$ always implies Log-Sobolev. We give here the following formal corollary of HWI inequality.

**Corollary 1.**

1. *If $\rho > 0$, then inequality* (5.1) *implies* (4.2) *and* (5.2) *implies* (4.1).
2. *If $V(x) - \rho x^2$ is a convex for some $\rho \in \mathbb{R}$, then Talagrand's free transportation inequality with constant $C > \max\{0, -\rho\}$ implies free Log-Sobolev inequality with constant $K = \max\{\rho, \frac{(C+\rho)^2}{32C}\}$. More precisely,*

$$\forall \mu \in \mathcal{P}(\mathbb{R}), \quad CW_2^2(\mu, \mu_V) \leqslant E(\mu) - E(\mu_V) \Longrightarrow$$
$$\forall \mu \in \mathcal{P}(\mathbb{R}), \quad E(\mu) - E(\mu_V) \leqslant \frac{1}{4K} I(\mu).$$

3. *In particular, if $V$ is convex and $C^2$ such that $V''(x) \geqslant \rho > 0$ for $|x| \geqslant r$, then free Log-Sobolev inequality holds with the constant $C > 0$ from* (2.10).

**Proof.** 1. It follows as an application of Young's inequality $a^p/p + b^q/q \geqslant ab$ for $a, b \geqslant 0$.

2. For $\rho > 0$, everything is clear. In the case $\rho \leqslant 0$, then, from (5.1) and Talagrand's transportation inequality, one has for $\delta > 0$, that

$$\begin{aligned} E(\mu) - E(\mu_V) &\leqslant \sqrt{I(\mu)}\, W_2(\mu, \mu_V) - \rho W_2^2(\mu, \mu_V) \\ &\leqslant 4\delta I(\mu) + \left(\frac{1}{C\delta} - \frac{\rho}{C}\right)\big(E(\mu) - E(\mu_V)\big) \end{aligned}$$

which yields for any $\delta > \frac{1}{C+\rho}$

$$E(\mu) - E(\mu_V) \leqslant \frac{4C\delta^2}{(C+\rho)\delta - 1} I(\mu).$$

Taking minimum over $\delta > \frac{1}{C+\rho}$ gives the conclusion.

3. In the case $V$ is convex, $C^2$ and strongly convex for large values, part 2 of Theorem 2 does the rest. □

## 6. Brunn–Minkowski inequality

The (one-dimensional) free Brunn–Minkowski inequality was put forward in [24] again through random matrix approximation. We provide here a direct mass transportation proof similar to the one of its classical (one-dimensional) counterpart (see e.g. [12]). As discussed in [24], this inequality may be used to deduce in an easy way both the Log-Sobolev and transportation inequalities.

The main result of this section is the following theorem.

**Theorem 6.** *Assume that $V_1, V_2, V_3$ are some potentials satisfying* (2.1) *such that for some $a \in (0,1)$,*

$$aV_1(x) + (1-a)V_2(y) \geqslant V_3\big(ax + (1-a)y\big) \quad \text{for all } x, y \in \mathbb{R}. \tag{6.1}$$

*Then*

$$aE_{V_1}(\mu_{V_1}) + (1-a)E_{V_2}(\mu_{V_2}) \geqslant E_{V_3}(\mu_{V_3}). \tag{6.2}$$

**Proof.** Take the (increasing) transportation map $\theta$ from $\mu_{V_1}$ into $\mu_{V_2}$. This certainly exists as the measure $\mu_{V_1}$ has no atoms.

Noticing that for any measure with finite logarithmic energy, we have the obvious equality

$$\int \log|x-y|\,\mu(dx)\,\mu(dy) = 2\int_{x>y} \log(x-y)\,\mu(dx)\,\mu(dy).$$

Using this we argue that

$$\begin{aligned}
&\int aV_1(x) + (1-a)V_2\big(\theta(x)\big)\mu_{V_1}(dx) \\
&\quad - 2\iint_{x>y} \big(a\log(x-y) + (1-a)\log\big(\theta(x)-\theta(y)\big)\big)\,\mu_{V_1}(dx)\,\mu_{V_1}(dy) \\
&\geqslant \int V_3\big(ax + (1-a)\theta(x)\big)\mu_{V_1}(dx) \\
&\quad - 2\iint_{x>y} \log\big[\big(ax + (1-a)\theta(x)\big) - \big(ay + (1-a)\theta(y)\big)\big]\,\mu_{V_1}(dx)\,\mu_{V_1}(dy) \\
&= E_{V_3}(\nu) \geqslant E_{V_3}(\mu_{V_3})
\end{aligned}$$

where $\nu = (a\,\mathrm{id} + (1-a)\theta)_{\#}\mu_{V_1}$ and we used (6.1) and the concavity of the logarithm on $(0,\infty)$. The proof is complete. □

## 7. Random matrices and a first version of Poincaré inequality

In the next three sections, we investigate Poincaré type inequalities in the free (one-dimensional) context. We discuss two versions of it. The first one is suggested by large matrix approximations and the classical Poincaré inequality for strictly convex potentials, but will be proved directly. Recall first the classical Poincaré inequality (cf. e.g. [2,23,29,32]...).

**Theorem 7.** *Let $\mu(dx) = e^{-W(x)}\,dx$ be a probability measure on $\mathbb{R}^d$ such that $W(x) - r|x|^2$ is convex. Then for any compactly supported and smooth function $\phi : \mathbb{R}^d \to \mathbb{R}$, we have that*

$$\int |\nabla\phi|^2\,d\mu \geqslant r\,\mathrm{Var}_\mu(\phi). \tag{7.1}$$

Assume now that $V$ is a potential on $\mathbb{R}$ with enough growth at infinity. Consider the matrix models on $\mathcal{H}_n$, the space of Hermitian $n \times n$ matrices with the inner product $\langle A, B\rangle = \mathrm{Tr}(AB^*)$ and the probability measure given by

$$\mathbb{P}_n(dM) = \frac{1}{Z_n(V)} e^{-n\,\mathrm{Tr}(V(M))}\, dM$$

where here $dM$ is the standard Lebesgue measure on $\mathcal{H}_n$. We have that for any bounded continuous function $F : \mathbb{R} \to \mathbb{R}$,

$$\int \frac{1}{n} \mathrm{Tr}\big(F(M)\big)\mathbb{P}_n(dM) \underset{n\to\infty}{\longrightarrow} \int F(x)\, \mu_V(dx). \tag{7.2}$$

Assume in addition that $V(x) - \rho x^2$ is a convex function on $\mathbb{R}$. Then, consider $\Phi(M) = \mathrm{Tr}\,\phi(M)$, where $\phi : \mathbb{R} \to \mathbb{R}$ is a compactly supported and smooth function. Notice that $\nabla\Phi(M) = \phi'(M)$ and thus $|\nabla\Phi(M)|^2 = |\phi'(M)|^2 = \mathrm{Tr}(\phi'(M)^2)$. Since $n\,\mathrm{Tr}(V(M)) - n\rho|M|^2$ is convex as a function of $M$, we can apply Poincaré's inequality on $\mathcal{H}_n$ to obtain that

$$\int \mathrm{Tr}\big(\phi'(M)^2\big)\mathbb{P}_n(dM) \geqslant n\rho \mathrm{Var}_{\mathbb{P}_n}\big(\mathrm{Tr}\big(\phi(M)\big)\big). \tag{7.3}$$

The first term in this inequality divided by $n$ (cf. Eq. (7.2)) converges to $\int \phi'(x)^2\, \mu_V(dx)$. To understand the second term in the above equation, notice that $\mathrm{Var}(\mathrm{Tr}(\phi(M))) = \mathbb{E}[(\mathrm{Tr}(\phi(M)) - \mathbb{E}[\mathrm{Tr}(\phi(M))])^2]$. The study of the asymptotic of the linear statistics, $\mathrm{Tr}(\phi(M)) - \mathbb{E}[\mathrm{Tr}(\phi(M))]$ in the literature of random matrix is known as "fluctuations". From Johansson's paper [19], it is known that this is universal in the sense that the limit in distribution of the fluctuations is Gaussian and, at least in the case of polynomial $V$ (for which $V(x) - \rho x^2$ fulfills the conditions in there), the variance of the Gaussian limit depends only on the endpoints of the support of $\mu_V$. Moreover, in the particular case of $V(x) = 2x^2$, the variance of the distribution was computed for example in [22] and [19] as

$$\frac{1}{2\pi^2} \int_{-1}^{1} \int_{-1}^{1} \left(\frac{\phi(t) - \phi(s)}{t - s}\right)^2 \frac{1 - ts}{\sqrt{1 - t^2}\sqrt{1 - s^2}}\, dt\, ds. \tag{7.4}$$

This variance is interpreted in [8] in terms of the number operator of the arcsine law. We will come back to this aspect in Section 9.

Dividing the inequality in Eq. (7.3) by $n$ and taking the limit when $n \to \infty$, these heuristics (after a simple rescaling) suggest the following result.

**Theorem 8.** *Assume that $V(x) - \rho x^2$ is convex for some $\rho > 0$. Then for any smooth function $\phi$, one has that*

$$\int \phi'(x)^2 \mu_V(dx) \geqslant \frac{\rho}{2\pi^2} \int_a^b \int_a^b \left(\frac{\phi(x) - \phi(y)}{x - y}\right)^2 \times \frac{-2ab + (a + b)(x + y) - 2xy}{2\sqrt{(x - a)(b - x)}\sqrt{(y - a)(b - y)}}\, dx\, dy. \tag{7.5}$$

*where* $\mathrm{supp}(\mu_V)=[a,b]$. *Equality is attained for* $V(x)=\rho(x-\alpha)^2+\beta$ *and* $\phi(x)=c_1+c_2x$ *for some constants* $c_1,c_2$.

The reader may wonder if the numerator in the second fraction of (7.5) is nonnegative. This is so because

$$-2ab+(a+b)(x+y)-2xy=2\left(\left(\frac{b-a}{2}\right)^2-\left(x-\frac{a+b}{2}\right)\left(y-\frac{a+b}{2}\right)\right)\geqslant 0$$

for any $x,y\in[a,b]$.

**Proof.** Using a simple rescaling we may assume without loss of generality that $a=-1$ and $b=1$ and the inequality we have to show reduces to

$$\int \phi'(x)^2\mu_V(dx)\geqslant \frac{\rho}{2\pi^2}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\frac{1-xy}{\sqrt{1-x^2}\sqrt{1-y^2}}\,dx\,dy. \tag{7.6}$$

Then, based on Eq. (2.5), we have that

$$g(x)=\frac{\sqrt{1-x^2}}{2\pi^2}\int_{-1}^{1}\frac{V'(y)-V'(x)}{\sqrt{1-y^2}(y-x)}\,dy.$$

From the convexity of $V(x)-\rho x^2$, we learn that $\frac{V'(y)-V'(x)}{y-x}\geqslant 2\rho$ and thus that

$$g(x)\geqslant \frac{\rho}{\pi}\sqrt{1-x^2}, \tag{7.7}$$

which implies

$$\int \phi'(x)^2\mu_V(dx)\geqslant \frac{\rho}{\pi}\int_{-1}^{1}\phi'(x)^2\sqrt{1-x^2}\,dx.$$

Therefore it is enough to check that

$$\int_{-1}^{1}\phi'(x)^2\sqrt{1-x^2}\,dx\geqslant \frac{1}{2\pi}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2\frac{1-xy}{\sqrt{1-x^2}\sqrt{1-y^2}}\,dx\,dy \tag{7.8}$$

for any smooth $\phi$. Now, we make the change of variables $x=\cos t$ to justify

$$\int_{-1}^{1}\phi'(x)^2\sqrt{1-x^2}\,dx=\int_{0}^{\pi}\phi'(\cos t)^2\sin^2(t)\,dt=\int_{0}^{\pi}\psi'(t)^2\,dt$$

where $\psi(t)=\phi(\cos t)$.

On the other hand, using the change of variable $x = \cos t$, $y = \cos s$ on the right-hand side, inequality (7.8) becomes

$$\int_0^\pi \psi'(t)^2\,dt \geqslant \frac{1}{2\pi}\int_0^\pi\int_0^\pi \left(\frac{\psi(t)-\psi(s)}{\cos t - \cos s}\right)^2 (1-\cos t\cos s)\,dt\,ds. \tag{7.9}$$

To show this, we write $\psi(t) = \sum_{k=0}^\infty a_k \cos kt$ and then, because $\psi$ is a smooth function, we can differentiate term by term to get $\psi'(t) = -\sum_{k=1}^\infty k a_k \sin kt$, therefore

$$\int_0^\pi \psi'(t)^2\,dt = \frac{\pi}{2}\sum_{k=1}^\infty k^2 a_k^2$$

and

$$\begin{aligned}
&\int_0^\pi\int_0^\pi \left(\frac{\psi(t)-\psi(s)}{\cos t - \cos s}\right)^2 (1-\cos t\cos s)\,dt\,ds \\
&\quad = \sum_{k,l=1}^\infty a_k a_l \int_0^\pi\int_0^\pi \frac{(\cos kt - \cos ks)(\cos lt - \cos ls)(1-\cos t\cos s)}{(\cos t - \cos s)^2}\,dt\,ds.
\end{aligned}$$

To compute the integrals on the right-hand side of the above equation, we take the generating function of these numbers and with a little algebra one can show that

$$\begin{aligned}
&\sum_{k,l=1}^\infty u^k v^l \int_0^\pi\int_0^\pi \frac{(\cos kt - \cos ks)(\cos lt - \cos ls)(1-\cos t\cos s)}{(\cos t - \cos s)^2}\,dt\,ds \\
&= \int_0^\pi\int_0^\pi \frac{(u-u^3)(v-v^3)(1-\cos t\cos s)}{(1+u^2-2u\cos t)(1+u^2-2u\cos s)(1+v^2-2v\cos t)(1+v^2-2v\cos s)}\,dt\,ds \\
&= \frac{\pi^2 uv}{(1-uv)^2} = \pi^2\sum_{k=1}^\infty k u^k v^k
\end{aligned} \tag{7.10}$$

for all $u, v \in (-1, 1)$. The last integral can be computed as follows. First use partial fractions to justify

$$\begin{aligned}
\int_0^\pi \frac{(A + B\cos t)\,dt}{(1+u^2-2u\cos t)(1+v^2-2v\cos t)} &= \int_0^\pi \frac{C\,dt}{1+u^2-2u\cos t} + \int_0^\pi \frac{D\,dt}{1+v^2-2v\cos t} \\
&= \frac{C/2}{1-u^2} + \frac{D/2}{1-v^2}
\end{aligned}$$

where the constants $C, D$ are linear combinations of $A$ and $B$. Further, taking $A = 1$ and $B = -\cos s$ and repeating once more the partial fractions argument, one can cary out the proof of (7.10).

The main consequence of the above calculation is that

$$\int_0^\pi \int_0^\pi \frac{(\cos kt - \cos ks)(\cos lt - \cos ls)(1 - \cos t \cos s)}{(\cos t - \cos s)^2} \, dt \, ds = \pi^2 k \delta_{kl}$$

and that

$$\int_0^\pi \int_0^\pi \left( \frac{\psi(t) - \psi(s)}{\cos t - \cos s} \right)^2 (1 - \cos t \cos s) \, dt \, ds = \pi^2 \sum_{k=1}^\infty k a_k^2. \tag{7.11}$$

Therefore inequality (7.9) becomes equivalent to

$$\frac{\pi}{2} \sum_{k=1}^\infty k^2 a_k^2 \geqslant \frac{\pi}{2} \sum_{k=1}^\infty k a_k^2$$

which is obviously true. Notice that equality in this inequality is attained for the case $a_k = 0$ for all $k \geqslant 2$ and arbitrary $a_1$. This corresponds to the case $\psi(t) = c_1 + c_2 \cos t$ or $\phi(x) = c_2 x + c_1$ for some $c_1, c_2$.

Finally we point out that equality in (7.6) is attained if the equality is attained in (7.7) and (7.9). From there one can easily see from rescaling that equality in (7.5) is attained for $V(x) = \rho(x - \alpha)^2 + \beta$ and $\phi(x) = c_1 + c_2 x$. The proof of Theorem 8 is complete. □

In the above proof we showed a direct calculation for Eq. (7.11) which is natural in the course of the above proof. However, there is another way of looking at it which will appear below in Section 9 as the kernel of the number operator.

## 8. A second version of Poincaré inequality

The second version of the Poincaré inequality is motivated by the free calculus and the non-commutative derivative. It was already investigated by Biane [3] for the case of the semicircular law.

**Definition 1.** For a given probability measure $\mu$ on $\mathbb{R}$, we say that it satisfies a Poincaré inequality if there is a constant $C > 0$ such that

$$\iint \left( \frac{\phi(x) - \phi(y)}{x - y} \right)^2 \mu(dx) \, \mu(dy) \geqslant C \, \mathrm{Var}_\mu(\phi) \quad \text{for every } \phi \in C_0^1(\mathbb{R}). \tag{8.1}$$

By the best constant we mean the largest $C > 0$ for which the above inequality is satisfied and we denote it by $\mathrm{Poin}(\mu)$ or $\lambda_1(\mu)$ or $SG(\mu)$.

In the noncommutative setting for a given function $\phi$, we can think of $D\phi(x,y)=\frac{\phi(x)-\phi(y)}{x-y}$ as the noncommutative derivative of $\phi$. As pointed out by Voiculescu in [30], this is the unique map $D : C\langle x\rangle \to C\langle x\rangle \otimes C\langle x\rangle$ such that

1. $D1 = 0$.
2. $D(fg) = D(f)g + fD(g)$ for any $f, g \in C\langle x\rangle$.

First we collect a couple of obvious properties of the Poincaré constant.

**Proposition 2.**

1. *For any $a \neq 0$,*

$$\mathrm{Poin}\big((ax+b)_{\#}\mu\big) = \frac{1}{a^2}\,\mathrm{Poin}(\mu)$$

*where here and elsewhere, for a given function $f : \mathbb{R} \to \mathbb{R}$, $f_{\#}\mu$ is the push forward measure given by $(f_{\#}\mu)(A) = \mu(f^{-1}(A))$.*

2. *If $f : \mathbb{R} \to \mathbb{R}$ is a differential map such that $|f'(x)| \geqslant c > 0$ for all $x \in \mathbb{R}$, then*

$$\mathrm{Poin}(\mu) \geqslant c^2\,\mathrm{Poin}(f_{\#}\mu).$$

3. *If $\{\mu_n\}_{n\geqslant 1}$ is a sequence of probability measures which converges weakly to $\mu$, then*

$$\mathrm{Poin}(\mu) \geqslant \limsup_{n\to\infty} \mathrm{Poin}(\mu_n).$$

Next we describe some bounds for the Poincaré constant.

**Theorem 9.** *Assume that the measure $\mu$ has compact support and is not concentrated at one point. Then $\mu$ satisfies a Poincaré inequality with*

$$\frac{2}{d^2(\mu)} \leqslant \mathrm{Poin}(\mu) \leqslant \frac{1}{\mathrm{Var}(\mu)} \tag{8.2}$$

*where $d(\mu) = \mathrm{diam}(\mathrm{supp}(\mu))$ is the diameter of the support of $\mu$ and $\mathrm{Var}(\mu) = \int x^2\mu(dx) - (\int x\mu(dx))^2$. Equality on the left in* (8.2) *is attained only for the case*

$$\mu = \alpha\delta_a + (1-\alpha)\delta_b, \quad a < b,\ d0 < \alpha < 1.$$

*Equality on the right of* (8.2) *is attained only for the case of a semicircular law ($a \in \mathbb{R}$, $r > 0$)*

$$\mu(dx) = \frac{1}{2\pi r^2}\mathbb{1}_{[a-2r,a+2r]}(x)\sqrt{4r^2-(x-a)^2}\,dx.$$

*In addition, assume that $V$ is a $C^2$ potential on $\mathbb{R}$ such that for some integer $p$ and real $\rho > 0$, $V(x) - \rho x^{2p}$, is convex and $\mu$ is the minimizer of*

$$\int V(x)\,\mu(dx) - \iint \log|x-y|\,\mu(dx)\,\mu(dy)$$

*over all probability measures of* $\mathbb{R}$. *Then*

$$\frac{(p\rho\binom{2p}{p})^{\frac{1}{p}}}{8} \leqslant \mathrm{Poin}(\mu). \tag{8.3}$$

*In particular if* $p = 1$, *we get that* $\frac{\rho}{4} \leqslant \mathrm{Poin}(\mu)$.

**Proof.** For a given function $\phi \in C_0^1(\mathbb{R})$, the left-hand side of (8.2) follows from

$$\begin{aligned}\mathrm{Var}_\mu(\phi) &= \frac{1}{2}\iint \big(\phi(x) - \phi(y)\big)^2 \mu(dx)\,\mu(dy) \\ &= \frac{1}{2}\iint (x-y)^2\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \mu(dx)\,\mu(dy) \\ &\leqslant \frac{d^2(\mu)}{2}\iint \left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \mu(dx)\,\mu(dy). \end{aligned} \tag{8.4}$$

The right-hand side of (8.2) follows from (8.1) for a $\phi \in C_0^1(\mathbb{R})$ such that $\phi(x) = x$ on the support of $\mu$.

For measures $\mu = \alpha\delta_a + (1-\alpha)\delta_b$, condition (8.1) is equivalent to

$$\begin{aligned} C\alpha(1-\alpha)\big(\phi(b)-\phi(a)\big)^2 &\leqslant \alpha^2\big(\phi'(a)\big)^2 + (1-\alpha)^2\big(\phi'(b)\big)^2 \\ &\quad + 2\alpha(1-\alpha)\left(\frac{\phi(b)-\phi(a)}{b-a}\right)^2 \quad \text{for any } \phi \in C_0^1(\mathbb{R}). \end{aligned}$$

Since for any function $\phi \in C_0^\infty(\mathbb{R})$ we can find another function $\psi \in C_0^1(\mathbb{R})$ so that $\phi(a) = \psi(a)$ and $\phi(b) = \psi(b)$ and $\psi'(a) = 0$, $\psi'(b) = 0$, this is also equivalent to

$$C\alpha(1-\alpha)\big(\psi(b)-\psi(a)\big)^2 \leqslant 2\alpha(1-\alpha)\left(\frac{\psi(b)-\psi(a)}{b-a}\right)^2 \quad \text{for any } \psi \in C_0^1(\mathbb{R}).$$

This amounts to $C \leqslant 2/(b-a)^2$ and therefore, in this case, $\mathrm{Poin}(\mu) = \frac{2}{d^2(\mu)}$.

Conversely, if $\mu$ is a measure so that $\mathrm{Poin}(\mu) = \frac{2}{d^2(\mu)}$, then, for $1 > \epsilon > 0$, there is a function $\phi_\epsilon \in C_0^1(\mathbb{R})$ such that

$$\left(\frac{2}{d^2(\mu)} + \epsilon^2\right)\mathrm{Var}_\mu(\phi_\epsilon) > \iint \left(\frac{\phi_\epsilon(x)-\phi_\epsilon(y)}{x-y}\right)^2 \mu(dx)\,\mu(dy).$$

Without loss of generality we can assume that $0 = \inf \mathrm{supp}(\mu)$, $1 = \sup \mathrm{supp}(\mu)$ and $\int \phi_\epsilon\,d\mu = 0$, $\int \phi_\epsilon^2\,d\mu = 1$ where we recall that $\mathrm{supp}(\mu)$ stands for the support of $\mu$. In this case, the above inequality implies

$$
\begin{aligned}
2+\epsilon^2 \geqslant & \iint\limits_{|x-y|\geqslant 1-\epsilon} \left(\frac{\phi_\epsilon(x)-\phi_\epsilon(y)}{x-y}\right)^2 \mu(dx)\,\mu(dy) \\
& + \iint\limits_{|x-y|<1-\epsilon} \left(\frac{\phi_\epsilon(x)-\phi_\epsilon(y)}{x-y}\right)^2 \mu(dx)\,\mu(dy) \\
& \geqslant \iint\limits_{|x-y|\geqslant 1-\epsilon} \big(\phi_\epsilon(x)-\phi_\epsilon(y)\big)^2 \mu(dx)\,\mu(dy) \\
& + \frac{1}{(1-\epsilon)^2} \iint\limits_{|x-y|<1-\epsilon} \big(\phi_\epsilon(x)-\phi_\epsilon(y)\big)^2 \mu(dx)\,\mu(dy) \\
& = -\frac{\epsilon(2-\epsilon)}{(1-\epsilon)^2} \iint\limits_{|x-y|\geqslant 1-\epsilon} \Big(\phi_\epsilon(x)-\phi_\epsilon(y)\Big)^2 \mu(dx)\,\mu(dy) + \frac{2}{(1-\epsilon)^2},
\end{aligned}
$$

which results with

$$
\iint\limits_{|x-y|\geqslant 1-\epsilon} \big(\phi_\epsilon(x)-\phi_\epsilon(y)\big)^2 \mu(dx)\,\mu(dy) \geqslant 2 - \frac{\epsilon(1-\epsilon)^2}{2-\epsilon}. \tag{8.5}
$$

Now,

$$
\begin{aligned}
\iint\limits_{|x-y|\geqslant 1-\epsilon} \big(\phi_\epsilon(x)-\phi_\epsilon(y)\big)^2 \mu(dx)\,\mu(dy) &\leqslant \iint\limits_{\substack{|x-1/2|\geqslant 1/2-\epsilon \\ |y-1/2|\geqslant 1/2-\epsilon}} \big(\phi_\epsilon(x)-\phi_\epsilon(y)\big)^2 \mu(dx)\,\mu(dy) \\
&\leqslant 2\mu\big(|x-1/2|\geqslant 1/2-\epsilon\big).
\end{aligned} \tag{8.6}
$$

Thus (8.5) and (8.6) give

$$
\mu\big(|x-1/2|\geqslant 1/2-\epsilon\big) \geqslant 1 - \frac{\epsilon(1-\epsilon)^2}{4-2\epsilon} \quad \text{for any } 1>\epsilon>0.
$$

This shows that $\mu((0,1))=0$ and therefore $\mu=\alpha\delta_0+(1-\alpha)\delta_1$.

The other extreme case of inequality (8.2) is contained in Biane's paper [3] in the more general context of several noncommutative variables. For completeness we will provide here a selfcontained proof. In the first place, using Proposition 8.1, we may assume that

$$
\mu(dx) = \frac{1}{2\pi}\mathbb{1}_{[-2,2]}(x)\sqrt{4-x^2}\,dx
$$

is the semicircular law on $[-2, 2]$. Take $U_n$ to be the Chebyshev polynomials of second kind defined by $U_n(\cos(\theta)) = \frac{\sin(n+1)\theta}{\sin\theta}$. With this choice, we have that $U_n(\frac{x}{2})$ are the orthogonal polynomials with respect to $\mu$. The generating function of $U_n$ is given by

$$\sum_{n=0}^{\infty} r^n U_n(x) = \frac{1}{1 - 2rx + r^2} \quad \text{for } |x|, |r| < 1,$$

from which one gets

$$\sum_{n=0}^{\infty} r^n \frac{U_n(x) - U_n(y)}{x - y} = \frac{2r}{(1 - 2rx + r^2)(1 - 2ry + r^2)} = 2\sum_{n=0}^{\infty} r^n \sum_{k=0}^{n-1} U_k(x)U_{n-1-k}(y),$$

and then

$$\frac{U_n(x) - U_n(y)}{x - y} = 2\sum_{k=0}^{n-1} U_k(x)U_{n-1-k}(y). \tag{8.7}$$

Now, for a given $\phi \in C_0^1(\mathbb{R})$, we can write in $L^2(\mu)$ sense,

$$\phi(x) = \sum_{n=0}^{\infty} \alpha_n U_n\left(\frac{x}{2}\right),$$

yielding from orthogonality and (8.7) that

$$\mathrm{Var}_\mu(\phi) = \int \phi^2\, d\mu - \left(\int \phi\, d\mu\right)^2 = \sum_{n=1}^{\infty} \alpha_n^2 \quad \text{and}$$
$$\iint \left(\frac{\phi(x) - \phi(y)}{x - y}\right)^2 \mu(dx)\, \mu(dy) = \sum_{n=1}^{\infty} n\alpha_n^2.$$

It follows that in this case $\mathrm{Poin}(\mu) = 1 = 1/\mathrm{Var}(\mu)$ and equality is attained only for $\phi(x) = c_1 + c_2U_1(x) = c_1 + c_2x$ for some constants $c_1, c_2$.

To prove the converse, take a compactly supported measure $\mu$ and assume that $\int x\, \mu(dx) = 0$ and $\int x^2\, \mu(dx) = 1$. In order to show that $\mu$ is the semicircular distribution, it suffices to show that $\int U_n(\frac{x}{2})\mu(dx) = 0$ for all $n \geqslant 1$. We use induction to this task. Assuming true for $U_1, U_2, \ldots, U_n$, and using $U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x)$, we need to show that $xU_n(\frac{x}{2})$ integrates to 0 against $\mu$. Applying Poincaré's inequality to $U_n(\frac{x}{2}) + rU_1(\frac{x}{2})$ together with the induction hypothesis and equation (8.7), we get that for any $r \in \mathbb{R}$,

$$\int U_n^2\left(\frac{x}{2}\right)\mu(dx) + r\int xU_n\left(\frac{x}{2}\right)\mu(dx) \leqslant \iint \left(\frac{U_n(\frac{x}{2}) - U_n(\frac{y}{2})}{x - y}\right)^2 \mu(dx)\, \mu(dy),$$

which implies that $\int xU_n(\frac{x}{2})\, \mu(dx) = 0$.

In the case of the equilibrium measure of a convex potential $V$, we have the support of the measure consists of one interval $[a,b]$ and $a,b$ solve the system (cf. Eq. (2.4))

$$\frac{1}{2\pi}\int_a^b V'(x)\sqrt{\frac{x-a}{b-x}}\,dx = 1 \quad \text{and} \quad \frac{1}{2\pi}\int_a^b V'(x)\sqrt{\frac{b-x}{x-a}}\,dx = -1.$$

If we denote $c=(b-a)/2$ and $\beta=(a+b)/2$, the system above can be rewritten in terms of $\beta$ and $c$ as

$$\frac{c}{2\pi}\int_{-1}^1 V'(\beta+ct)\frac{1+t}{\sqrt{1-t^2}}\,dt = 1 \quad \text{and} \quad \frac{c}{2\pi}\int_{-1}^1 V'(\beta+ct)\frac{1-t}{\sqrt{1-t^2}}\,dt = -1$$

which is equivalent to

$$\frac{c}{2\pi}\int_{-1}^1 V'(\beta+ct)\frac{t}{\sqrt{1-t^2}}\,dt = 1 \quad \text{and} \quad \int_{-1}^1 V'(\beta+ct)\frac{1}{\sqrt{1-t^2}}\,dt = 0.$$

Since $V$ is $C^2$ the first equation can be integrated by parts to get that

$$\frac{c^2}{2\pi}\int_{-1}^1 V''(\beta+ct)\sqrt{1-t^2}\,dt = 1.$$

On the other hand we know that $V''(x)\geqslant 2p(2p-1)\rho x^{2p-2}$, hence

$$\begin{aligned}
1 &\geqslant \frac{2p(2p-1)\rho c^2}{2\pi}\int_{-1}^1 (ct+\beta)^{2p-2}\sqrt{1-t^2}\,dt \\
&\geqslant \frac{2p(2p-1)\rho c^{2p}}{2\pi}\int_{-1}^1 t^{2p-2}\sqrt{1-t^2}\,dt \\
&= \frac{p(2p-1)\rho c^{2p}\binom{2p}{p}}{4^p(2p-1)} \\
&= \frac{p\rho\binom{2p}{p}c^{2p}}{4^p}.
\end{aligned}$$

This yields

$$c \leqslant 2\left(m\rho\binom{2p}{p}\right)^{-\frac{1}{2p}}.$$

Finally, because $d(\mu)=b-a=2c$, we arrive at (8.3). $\square$

To conclude this section, we present an inequality which relates the equilibrium measure of a strong convex potential and the arcsine law.

**Theorem 10.** *Assume that $V(x) - \rho x^2$ is a convex for some $\rho > 0$ and the equilibrium measure $\mu_V$ has support $[a,b]$. Let $\mathrm{arcsine}_{a,b} = \mathbb{1}_{[a,b]}(x)\frac{1}{\pi\sqrt{(b-x)(x-a)}}\,dx$ be the arcsine law with support $[a,b]$. Then for any smooth function supported on $[a,b]$,*

$$\int \phi'(x)^2\,\mu_V(dx) \geqslant \rho\,\mathrm{Var}_{\mathrm{arcsine}_{a,b}}(\phi), \tag{8.8}$$

*where the variance is considered with respect to the* $\mathrm{arcsine}_{a,b}$ *law.*

**Proof.** It suffices to deal with the case $a=-1$, $b=1$, the rest following by simple rescaling. Recall that in the proof of Theorem 8, we use convexity to get that the density $g(x)$ of $\mu_V$ satisfies $g(x) \geqslant \frac{\rho}{\pi}\sqrt{1-x^2}$. Thus the proof reduces to

$$\frac{1}{\pi}\int_{-1}^{1} \phi'(x)^2\sqrt{1-x^2}(dx) \geqslant \mathrm{Var}_{\mathrm{arcsine}}(\phi). \tag{8.9}$$

For this, write $\phi = \sum_{n=0}^{\infty}\alpha_n T_n(x)$ the expansion of $\phi$ in terms of Chebyshev polynomials of the first kind. Now, $T_n' = nU_{n-1}$ and thus the above inequality reduces to the obvious inequality $\sum_{n=1}^{\infty} n^2\alpha_n^2 \geqslant \sum_{n=1}^{\infty}\alpha_n^2$. $\square$

We will actually see below that inequality (8.9) is simply the spectral gap for the Jacobi operator associated to the arcsine law.

## 9. Poincaré inequalities and Jacobi operators

In this section we show how the two versions of the Poincaré inequalities can be viewed as spectral gaps for some Jacobi operators. This discussion is mainly driven from the work [8] by Cabanal-Duvillard and his interpretation of the variance in (7.4) in terms of the number operator of the Jacobi operator associated to the arcsine law. This viewpoint allows for an unified perspective of the Poincaré inequalities presented in the preceding sections.

For our purpose we consider here the Jacobi operators given, for smooth functions on $(-1,1)$, by

$$L_\lambda f(x) = -\big(1-x^2\big)f''(x) + (2\lambda+1)xf'(x) \tag{9.1}$$

for $\lambda \geqslant 0$. We consider the Gegenbauer polynomials $C_n^\lambda$, $\lambda > 0$, defined by the generating function

$$\sum_{n=0}^{\infty} r^n C_n^\lambda(x) = \frac{1}{(1-rx+r^2)^\lambda}.$$

For $\lambda = 0$ we set $C_n^\lambda(x) = T_n(x)/n$, $n \geqslant 1$, where $T_n$ are the Chebyshev polynomials of the first kind.

It is known that $C_n^\lambda$ are eigenfunctions of $L_\lambda$, with eigenvalue $n(n+2\lambda)$, i.e.

$$L_\lambda C_n^\lambda = n(n+2\lambda)C_n^\lambda.$$

On the other hand the Gegenbauer polynomials are orthogonal with respect to the probability measure

$$\nu_\lambda = \frac{2^{2\lambda}\Gamma^2(\lambda+1)}{\pi\Gamma(2\lambda+1)}\mathbb{1}_{[-1,1]}(x)\big(1-x^2\big)^{\lambda-1/2}.$$

Notice that in the case of $\lambda = 0$, this becomes the arcsine law and for $\lambda = 1$, this is the semicircular law, while for $\lambda = 1/2$, this becomes the uniform measure on $[-1,1]$.

Take now the normalized Gegenbauer polynomials $\phi_n^\lambda = G_n^\lambda/\sqrt{c_n^\lambda}$, where $c_n^\lambda = \int G_n^\lambda(x)^2\nu_\lambda(dx)$. Then $\phi_n^\lambda$ form an orthonormal basis of $L^2(\nu_\lambda)$ and thus the operator $L_\lambda$ is diagonalized in this basis. Consider $N_\lambda$ to be the counting number operator with respect to the basis $\phi_n^\lambda$, i.e.

$$N_\lambda \phi_n^\lambda = n\phi_n^\lambda. \tag{9.2}$$

This implies that $L_\lambda = N_\lambda^2 + 2\lambda N_\lambda$. Therefore we have the following two inequalities

$$L_\lambda \geqslant (2\lambda+1)N_\lambda \quad \text{and} \quad N_\lambda \geqslant 1 - P_\lambda \tag{9.3}$$

where $P_\lambda$ here stands for the projection on constant functions in $L^2(\nu_\lambda)$. In other words, $P_\lambda\phi = \int \phi\nu_\lambda$.

Notice that Eq. (9.3) include two statements. The first one is the comparison of $L$ and $N$, with the spectral gap $2\lambda+1$ while the second one is the spectral gap of the counting number operator with the spectral gap 1. In the sequel we want to translate these spectral gaps in terms of Poincaré type inequality. For this matter we need to find the kernel of the operator $N$.

Then we have for any function in the domain of definition of $L_\lambda$, that $\phi = \sum_{n=0}^\infty \alpha_n\phi_n^\lambda$, and then

$$\langle L\phi,\phi\rangle_{L^2(\nu_\lambda)} = \sum_{n=0}^\infty n(n+2\lambda)\alpha_n^2.$$

On the other hand, using integration by parts, we can justify that

$$\langle L\phi,\phi\rangle_{L^2(\nu_\lambda)} = \int \phi L_\lambda\phi\, d\nu_\lambda = \int \phi'(x)^2\big(1-x^2\big)\nu_\lambda(dx).$$

For the number operator, we have that

$$\int \phi N_\lambda \phi\, d\nu_\lambda = \sum_{n=0}^\infty n\alpha_n^2 = \lim_{r\uparrow 1}\sum_{n=0}^\infty nr^{n-1}\alpha_n^2.$$

Now, for $-1 < r < 1$,

$$\sum_{n=0}^{\infty} nr^{n-1}\alpha_n^2 = \iint \phi(x)\phi(y)\sum_{n=0}^{\infty} nr^{n-1}\phi_n^{\lambda}(x)\phi_n^{\lambda}(y)\,\nu_\lambda(dx)\,\nu_\lambda(dy).$$

Furthermore, since $\int \phi_n^{\lambda} d\nu_\lambda = 0$ for $n \geqslant 1$, we also obtain that $\iint \phi^2(x)\phi_n^{\lambda}(y)\nu_\lambda(dx)\nu_\lambda(dy) = 0$ for $n \geqslant 0$ and thus, denoting $K_\lambda(r, x, y) = -\sum_{n=0}^{\infty} nr^{n-1}\phi_n^{\lambda}(x)\phi_n^{\lambda}(y)$,

$$\begin{aligned}\iint \phi(x)\phi(y)&\sum_{n=0}^{\infty} nr^{n-1}\phi_n^{\lambda}(x)\phi_n^{\lambda}(y)\,\nu_\lambda(dx)\,\nu_\lambda(dy)\\ &= \frac{1}{2}\iint \big(\phi(x)-\phi(y)\big)^2 K_\lambda(r, x, y)\,\nu_\lambda(dx)\,\nu_\lambda(dy).\end{aligned}$$

The following formula is essentially due to Watson [33] and valid for $\lambda > 0$,

$$\sum_{n=0}^{\infty} r^n\phi_n^{\lambda}(x)\phi_n^{\lambda}(y) = \frac{(1-r^2)\Gamma(2\lambda)}{2^{2\lambda-1}\Gamma^2(\lambda)}\int_{-1}^{1}\frac{(1-z^2)^{\lambda-1}}{(1-2r(xy+z\sqrt{(1-x^2)(1-y^2)})+r^2)^{1+\lambda}}\,dz.$$

For $\lambda = 0$, we have to deal with the Chebyshev polynomials of the first kind which was more or less what appeared in the proof of Theorem 8. For this case, we have that (denoting $x = \cos t$ and $y = \cos s$),

$$\sum_{n=0}^{\infty}\frac{r^n}{c_n}T_n(x)T_n(y) = \frac{1-r\cos(t+s)}{1-2r\cos(t+s)+r^2} + \frac{1-r\cos(t-s)}{1-2r\cos(t-s)+r^2}$$

where $c_n = \int T_n^2 d\nu_0 = 1$ for $n = 0$ and $1/2$ otherwise.

Thus, we obtain, after differentiation with respect to $r$ and then limit over $r \uparrow 1$, that

$$K_\lambda(x, y) = \lim_{r\uparrow 1} K_\lambda(r, x, y) = \begin{cases} \frac{\Gamma(2\lambda)}{2^{3\lambda-1}\Gamma^2(\lambda)}\int_{-1}^{1}\frac{(1-z^2)^{\lambda-1}}{(1-xy-z\sqrt{(1-x^2)(1-y^2)})^{1+\lambda}}\,dz, & \lambda > 0,\\ \frac{1-xy}{(x-y)^2}, & \lambda = 0,\\ \frac{1}{2(x-y)^2}, & \lambda = 1.\end{cases} \tag{9.4}$$

The integrand is not a rational function. In some cases, it is algebraic since $\lambda \geqslant 0$ need not be an integer.

To reveal the singularity of this kernel, we make the change of variable

$$1 - xy - z\sqrt{(1-x^2)(1-y^2)} = t\Big(1 - xy - \sqrt{(1-x^2)(1-y^2)}\Big).$$

Then, after simple algebraic manipulations, setting $f_\lambda : (0, 1) \to \mathbb{R}$,

$$f_\lambda(u) = \int_{1}^{1/u}\frac{[(t-1)(1-ut)]^{\lambda-1}}{t^{\lambda+1}}\,dt,$$

and

$$H_\lambda(x,y)=\begin{cases}\frac{\Gamma(2\lambda)(1-xy+\sqrt{(1-x^2)(1-y^2)})^\lambda}{2^{3\lambda-1}\Gamma^2(\lambda)((1-x^2)(1-y^2))^{\lambda-1/2}}f_\lambda\big(\frac{(x-y)^2}{(1-xy+\sqrt{(1-x^2)(1-y^2)})^2}\big), & \lambda>0,\\ 1-xy, & \lambda=0,\\ \frac{1}{2}, & \lambda=1,\end{cases} \tag{9.5}$$

we can rewrite Eq. (9.4) for $|x|, |y|<1$ as

$$K_\lambda(x,y)=\frac{H_\lambda(x,y)}{(x-y)^2} \tag{9.6}$$

where $H_\lambda(x,y)$ is a continuous function of $x, y\in[-1,1]$.

Now, from (9.3), we obtain the following result.

**Theorem 11.** *For any $\lambda\geqslant 0$, one has for all $\lambda\geqslant 0$ and any $\phi\in C^1([-1,1])$, that*

$$\int \phi'(x)^2\big(1-x^2\big)\,\nu_\lambda(dx)\geqslant\frac{2\lambda+1}{2}\iint\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 H_\lambda(x,y)\,\nu_\lambda(dx)\,\nu_\lambda(dy). \tag{9.7}$$

*and*

$$\iint\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 H_\lambda(x,y)\,\nu_\lambda(dx)\,\nu_\lambda(dy)\geqslant 2\operatorname{Var}_{\nu_\lambda}(\phi). \tag{9.8}$$

**Remark 5.**

1. Eq. (9.7) for $\lambda=0$ is the statement of Theorem 8 for the case $V(x)=2x^2$ and for $\lambda=1$ (more precisely, Eq. (7.8)) while Eq. (9.8) is the statement of the second Poincaré inequality contained in Theorem 9 for the semicircular law. The combination of these two inequalities is equation (8.9).
   In other words, for measures $\nu_\lambda$, the first Poincaré type inequality is driven by the comparison of the Jacobi and counting number operators defined in (9.1) and (9.2), as the second Poincaré type is the spectral gap of the counting number operator.
2. Combining Eqs. (9.7) and (9.8), we also get a Brascamp–Lieb type inequality:

$$\int \phi'(x)^2\big(1-x^2\big)\,\nu_\lambda(dx)\geqslant(2\lambda+1)\operatorname{Var}_{\nu_\lambda}(\phi). \tag{9.9}$$

   For $\lambda\geqslant 1/2$, the measure $\nu_\lambda$ is of the form $e^{-V(x)}\,dx$, where $V(x)=-c_\lambda-(\lambda-1/2)\times\log(1-x^2)$, a strictly convex function on $(-1,1)$ and according to the classical Brascamp–Lieb inequality [6],

$$\int \phi'(x)^2\frac{(1-x^2)^2}{(1+x^2)}\,\nu_\lambda(dx)\geqslant(2\lambda-1)\operatorname{Var}_{\nu_\lambda}(\phi). \tag{9.10}$$

   Notice here that neither (9.9) not (9.10) implies the other which means that they complement each other in some sense. For example if $\phi$ has support in $[-\frac{1}{2\lambda},\frac{1}{2\lambda}]$, (9.9) implies (9.10), while if $\phi$ is supported on $[-1,1]\setminus[-\frac{1}{2\lambda},\frac{1}{2\lambda}]$, (9.10) implies (9.9).

## 10. Wishart ensembles and Marcenko–Pastur distributions

In this section, we address the preceding functional inequalities for probability measures on the real positive axis in the context of the Wishart Ensembles from random matrix theory and their associated Marcenko–Pastur distributions.

We start with the random matrix heuristics although, as far as we know, it has not been used towards functional inequalities as before. The problems of large deviations principle for the distribution of the eigenvalues of Wishart ensembles is discussed in [16]. The model is as follows. Take $T(n)$ a $n \times p(n)$ random matrix with all the entries being iid $N(0,1)$ random variables. Then $T(n)T(n)^t$ for $n < p(n)$ is known as the nonsingular Wishart random ensemble. According to [17, p. 129], the distribution of the Wishart ensembles is given by

$$C_{np} e^{-\frac{p(n)}{2} \operatorname{Tr} M} (\det M)^{(p-n-1)/2}\, dM.$$

where the measure $dM = \prod_{i \leqslant j} dM_{ij}$ the restriction of the Lebegue measure on the set of $n \times n$ non-negative matrices.

It is also known (for example [17, p. 129]) that the joint distribution of eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ of $\frac{1}{p(n)} T(n)T(n)^t$ is given by

$$\frac{1}{Z_n} e^{-\frac{p(n)}{2} \sum_{i=1}^n t_i} \prod_{i=1}^n \lambda_i^{(p(n)-n-1)/2} \prod_{1 \leqslant i < j \leqslant n} |\lambda_i - \lambda_j|.$$

Our interest is in the limit distribution of $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$. The classical result states that if $n/p(n) \underset{n\to\infty}{\longrightarrow} \alpha \in (0,1]$, then the limit distribution of $\mu_n$ is the so called Marcenko–Pastur distribution given by

$$\mathbb{1}_{[(1-\sqrt{\alpha})^2, (1+\sqrt{\alpha})^2]}(x) \frac{\sqrt{4\alpha - (x-1-\alpha)^2}}{2\pi\alpha x}\, dx.$$

This is a particular model for the standard Wishart ensembles. However one can consider a more general example with potentials for which the distribution of the matrix is driven by a potential $Q : [0,\infty) \to \mathbb{R}$,

$$C_n e^{-p(n) \operatorname{Tr} Q(M)} (\det M)^{\gamma(n)} dM$$

where $dM$ stands for the Lebesgue measure on $n \times n$ positive definite matrices. The distribution of eigenvalues of $M$ is given by

$$\frac{1}{Z_n} e^{-p(n) \sum_{i=1}^n Q(t_i)} \prod_{i=1}^n t_i^{\gamma(n)} \prod_{1 \leqslant i < j \leqslant n} |t_i - t_j|.$$

The main result of [16] is that the distribution of the random measures $\mu_n = \frac{1}{p(n)} \sum_{i=1}^{p(n)} \delta_{\lambda_i}$ under the conditions $n/p(n) \underset{n\to\infty}{\longrightarrow} \alpha \in (0,1]$, $\gamma(n)/n \underset{n\to\infty}{\longrightarrow} \gamma > 0$, $\nu_n$ satisfy a large deviation principle with scale $n^{-2}$ and the rate function given by

$$R(\mu) = \tilde{E}_Q(\mu) - \inf_{\mu \in \mathcal{P}([0,\infty))} \tilde{E}_Q(\mu),$$

where

$$\tilde{E}_Q(\mu) = \int \alpha\big(Q(x) - \gamma \log(x)\big)\,\mu(dx) - \frac{\alpha^2}{2} \iint \log|x-y|\,\mu(dx)\,\mu(dy).$$

This gives the following motivation. Assume that $V : [0,\infty) \to \mathbb{R} \cup \{+\infty\}$ is a lower semi-continuous potential such that $\lim_{|x|\to\infty}(V(x) - 2\log|x|) = \infty$. Then, according to the results in [27], we know that there is a unique minimizer of

$$\inf_{\mu \in \mathcal{P}([0,\infty))} E_V(\mu).$$

In addition the equilibrium measure $\mu_V$ has compact support.

A particular case of interest is $V(x) = rx - s\log(x)$ with $r > 0, s \geqslant 0$ for which we know [27, p. 207] that the equilibrium measure is given by

$$\mu_V(dx) = \mathbb{1}_{[a,b]}(x) \frac{r\sqrt{(x-a)(b-x)}}{2\pi x}\,dx \quad \text{where } a = \frac{s+2-2\sqrt{s+1}}{r}, \qquad b = \frac{s+2+2\sqrt{s+1}}{r}. \tag{10.1}$$

One recovers the Marcenko–Pastur distribution for $V(x) = rx - s\log(x)$, $r > 0$, $s \geqslant 0$, with $r = 1/\alpha$ and $s = (1-\alpha)/\alpha$.

The natural way to deal with functional inequalities in the context of measures on the positive axis $[0,\infty)$ is to transfer measures from $[0,\infty)$ into measures on the whole $\mathbb{R}$. For a measure $\mu$ on $[0,\infty)$, consider thus the associated symmetric measure $\tilde{\mu}$ on $\mathbb{R}$ defined as

$$\mu(F) = \tilde{\mu}\big(\big\{x\colon\ x^2 \in F\big\}\big) \tag{10.2}$$

for any measurable set $F$ of $[0,\infty)$. Defining $\tilde{V}(x) = V(x^2)/2$, it is then an easy exercise to check that

$$E_V(\mu) = 2E_{\tilde{V}}(\tilde{\mu}). \tag{10.3}$$

In addition, the minimizer of $E_{\tilde{V}}$ is $\mu_{\tilde{V}} = \tilde{\mu}_V$ Further, for the non-decreasing transportation map $\theta$ of $\mu_V$ into $\mu$, define

$$\tilde{\theta}(x) = \operatorname{sign}(x)\sqrt{\theta\big(x^2\big)}, \tag{10.4}$$

which transports $\tilde{\mu}_{\tilde{V}}$ into $\tilde{\mu}$.

In addition, as it was pointed out in [18], the relative free Fisher information $I_V(\mu)$ is defined for measures $\mu$ on $[0,\infty)$ with density $p = d\mu/dx$ in $L^3([0,\infty), x\,dx)$ as

$$I_V(\mu) = \int_0^\infty x\big(H\mu(x) - V'(x)\big)^2 \mu(dx) \quad \text{with } H\mu(x) = \text{p.v.}\int \frac{2}{x-y}\,\mu(dy). \tag{10.5}$$

Otherwise we take $I_V(\mu) = +\infty$. The main reason for defining this in this way is because, cf. [18, Lemma 6.3] and the discussion following, one has

$$I_V(\mu) = 2I_{\tilde{V}}(\tilde{\mu}), \tag{10.6}$$

where $I_{\tilde{V}}$ is defined by (4.1).

To state the transportation cost result, we define the appropriate distance. For any $\mu, \nu \in \mathcal{P}([0,\infty))$, set the distance as

$$W(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \bigg(\int \big(\sqrt{x} - \sqrt{y}\big)^2 \pi(dx,dy)\bigg)^{1/2} \tag{10.7}$$

where $\Pi(\mu,\nu)$ is the set of probability measures on $\mathbb{R}^2$ with marginals $\mu$ and $\nu$.

In this context we have the following transportation cost inequality.

**Theorem 12.** *Assume that $V : (0,\infty) \to \mathbb{R}$ is $C^2((0,\infty))$ such that $V(x^2) - \rho x^2$ is convex on $(0,\infty)$ for some $\rho > 0$ and let $\mu_V$ be the equilibrium measure of $V$ on $[0,\infty)$. Then, for any probability measure $\mu$ on $[0,\infty)$, we have that*

$$\rho\, W^2(\mu,\mu_V) \leqslant E_V(\mu) - E_V(\mu_V). \tag{10.8}$$

*In the case of $V(x) = rx - s\log(x)$ with $r > 0$ and $s \geqslant 0$, this inequality with $\rho = r$ is sharp.*

**Proof.** As announced, the idea is to interpret this inequality as an inequality for potentials on the whole real line instead of $[0,\infty)$. Using the measures $\tilde{\mu}$ and $\tilde{\mu}_V$ from Eq. (10.2) together with (10.3), we have that

$$E_V(\mu) - E_V(\mu_V) = 2\big(E_{\tilde{V}}(\tilde{\mu}) - E_{\tilde{V}}(\tilde{\mu}_V)\big).$$

On the other hand, if $\theta$ is the (increasing) transportation map of $\mu_V$ into $\mu$, then it is not hard to check that

$$W^2(\mu,\nu) = \int \big(\sqrt{x} - \sqrt{\theta(x)}\big)^2 \mu_V(dx) = \int \big(x - \tilde{\theta}(x)\big)^2 \tilde{\mu}_V(dx).$$

In this framework the inequality (10.8) translates as

$$\frac{\rho}{2}\, W_2^2(\tilde{\mu},\tilde{\mu}_V) \leqslant E_{\tilde{V}}(\tilde{\mu}) - E_{\tilde{V}}(\tilde{\mu}_V). \tag{10.9}$$

From here we will use the same argument as in the proof of Theorem 2. Start with

$$E_{\tilde V}(\tilde\mu) - E_{\tilde V}(\tilde\mu_V) = \int \big(\tilde V\big(\tilde\theta(x)\big) - \tilde V(x) - \tilde V'(x)\big(\tilde\theta(x) - x\big)\big)\,\tilde\mu_V(dx)$$
$$+ \iint \bigg(\frac{\tilde\theta(x) - \tilde\theta(y)}{x-y} - 1 - \log\frac{\tilde\theta(x) - \tilde\theta(y)}{x-y}\bigg)\,\tilde\mu_V(dx)\,\tilde\mu_V(dy)$$

and notice that the second line of this is non-negative. For the first line we point out that because $\tilde V(x) - \frac{\rho}{2}x^2$ is convex and $x$ and $\tilde\theta(x)$ have the same sign, for any $x$,

$$\tilde V\big(\tilde\theta(x)\big) - \tilde V(x) - \tilde V'(x)\big(\tilde\theta(x) - x\big) \geqslant \frac{\rho}{2}\big(\tilde\theta(x) - x\big)^2,$$

which implies (10.8).

In the case $V(x) = rx - s\log(x)$, take $\theta(x) = (\sqrt{x} + m)^2$ for large $m$ and notice that $\tilde\theta(x) = x + m\,\mathrm{sign}(x)$. Therefore inequality (10.9) becomes

$$rm^2 \leqslant rm^2 + 2rm\int |x|\,\tilde\mu(dx) - 2s\int \log\bigg(\frac{|x + m\,\mathrm{sign}(x)|}{|x|}\bigg)\,\tilde\mu(dx)$$
$$- \iint \log\bigg(1 + m\frac{\mathrm{sign}(x) - \mathrm{sign}(y)}{x-y}\bigg)\,\tilde\mu(dx)\,\tilde\mu(dy)$$

which is sharp for large $m$. □

The next result is the Log-Sobolev type inequality, which was conjectured by Cabanal-Duvillard in [7, p. 140] for the case of Marcenko–Pastur distribution.

**Theorem 13.** *Let $V$ be as in the previous theorem. Then, with the definition from* (10.5) *and for any measure $\mu \in \mathcal{P}([0,\infty))$,*

$$E_V(\mu) - E_V(\mu_V) \leqslant \frac{1}{2\rho} I_V(\mu). \tag{10.10}$$

*In the case $V(x) = rx - s\log(x)$, $r > 0$ and $s \geqslant 0$ inequality* (10.10) *with $\rho = r$ is sharp.*

**Proof.** We will discuss here the proof only in the case when $\mu$ has a smooth compactly supported density, careful approximations being described in [18].

From (10.6), we have $I_V(\mu) = 2I_{\tilde V}(\tilde\mu)$, where $I_{\tilde V}(\tilde\mu) = \int (H\tilde\mu(x) - \tilde V'(x))^2\,\tilde\mu(dx)$. Rewriting everything in terms of $\tilde\mu$ and the associated quantities, the inequality to be proven can be written in the same way as we did in the proof of Theorem 4,

$$\frac{1}{2\rho}\int \big(H\tilde\mu\big(\tilde\theta(x)\big) - \tilde V'\big(\tilde\theta(x)\big)\big)^2\,\tilde\mu_{\tilde V}(dx)$$
$$+ \int \big(\tilde V(x) - \tilde V\big(\tilde\theta(x)\big) - \tilde V'\big(\tilde\theta(x)\big)\big(x - \tilde\theta(x)\big)\big)\,\tilde\mu_{\tilde V}(dx)$$
$$- \int \big(H\tilde\mu\big(\tilde\theta(x)\big) - \tilde V'\big(\tilde\theta(x)\big)\big)\big(x - \tilde\theta(x)\big)\,\tilde\mu_{\tilde V}(dx)$$
$$+ \int H\tilde\mu\big(\tilde\theta(x)\big)\big(x - \tilde\theta(x)\big)\tilde\mu_{\tilde V}(dx) - \iint \log\frac{x-y}{\tilde\theta(x) - \tilde\theta(y)}\,\tilde\mu_{\tilde V}(dx)\,\tilde\mu_{\tilde V}(dy) \geqslant 0. \tag{10.11}$$

Notice that $\tilde{V}(x) - \frac{\rho}{2}x^2$ is not convex on the whole real line but it is convex on the intervals $(0,\infty)$ and $(-\infty,0)$. The key to everything here is that $\tilde{\theta}(x)$ has the same sign as $x$ and this allows us to apply convexity of $\tilde{V}(x) - \frac{\rho}{2}x^2$ on each of the intervals $(-\infty,0)$ and $(0,\infty)$ to conclude that

$$\begin{aligned}\tilde{V}(x) - \tilde{V}\big(\tilde{\theta}(x)\big) - \tilde{V}'\big(\tilde{\theta}(x)\big)\big(x-\tilde{\theta}(x)\big) &\geqslant \frac{\rho}{2}\big(x^2 - \tilde{\theta}(x)^2 - 2\tilde{\theta}(x)\big(x-\tilde{\theta}(x)\big)\big)\\ &= \frac{\rho}{2}\big(x-\tilde{\theta}(x)\big)^2. \end{aligned}\tag{10.12}$$

From here we can follow word by word the proof of Theorem 4.

For the case $V(x) = rx$, we have equality in (10.10) if $\tilde{\theta}(x) = x + m\,\mathrm{sign}(x)$ and thus this means $\theta(x) = (\sqrt{x}+m)^2$.

In the case $V(x) = rx - s\log(x)$, we look at $\tilde{\theta}(x) = x + m$ for large $m$. In this case $\tilde{V}(x) = rx^2/2 - s\log|x|$ and then a simple calculation shows that (10.10) is equivalent to

$$\begin{aligned} &rm^2 + 2mr\int |x|\tilde{\mu}_V(dx) - 2s\int \log\bigg(\frac{|x+m\,\mathrm{sign}(x)|}{|x|}\bigg)\tilde{\mu}_V(dx)\\ &\quad - 2\iint \log\bigg(1 + m\frac{\mathrm{sign}(x)-\mathrm{sign}(y)}{x-y}\bigg)\tilde{\mu}(dx)\,\tilde{\mu}(dy)\\ &\leqslant \frac{m^2}{\rho}\int\bigg(r - \frac{s}{x(x+m\,\mathrm{sign}(x))}\bigg)^2\tilde{\mu}_V(dx). \end{aligned}$$

Dividing both sides by $m^2$ and taking the limit of $m$ to infinity implies that $\rho \leqslant r$. On the other hand $\rho = r$ validates (10.10), hence $\rho = r$ is the best constant. □

Next in line is the HWI inequality which is the content of the following statement.

**Theorem 14.** *Assume $V$ is as in Theorem* 12 *and the distance $W$ given by* (10.7). *Then for any measure $\mu \in \mathcal{P}([0,\infty))$,*

$$E_V(\mu) - E_V(\mu_V) \leqslant \sqrt{2I_V(\mu)}\,W(\mu,\mu_V) - \rho W^2(\mu,\mu_V).\tag{10.13}$$

*For the case of $V(x) = rx - s\log(x)$, $r>0$, $s\geqslant 0$, this inequality for $\rho = r$ is sharp.*

**Proof.** As it was made clear in the previous two theorems, we translate this inequality in terms of the associated symmetric measures on $\mathbb{R}$. Following upon the proofs of above theorems, we can rewrite (10.13) in the following form:

$$\begin{aligned} &\bigg(\int\big(H\tilde{\mu}\big(\tilde{\theta}(x)\big) - \tilde{V}'\big(\tilde{\theta}(x)\big)\big)^2\tilde{\mu}_V(dx)\int\big(\tilde{\theta}(x)-x\big)^2\tilde{\mu}_V(dx)\bigg)^{1/2}\\ &\quad - \int\big(H\tilde{\mu}\big(\tilde{\theta}(x)\big) - \tilde{V}'\big(\tilde{\theta}(x)\big)\big)\big(x-\tilde{\theta}(x)\big)\tilde{\mu}_V(dx)\\ &\quad + \int\big(\tilde{V}(x) - \tilde{V}\big(\tilde{\theta}(x)\big) - \tilde{V}'\big(\tilde{\theta}(x)\big)\big(x-\tilde{\theta}(x)\big) - \rho\big(\tilde{\theta}(x)-x\big)^2\big)\tilde{\mu}_V(dx)\\ &\quad + \int H\tilde{\mu}\big(\tilde{\theta}(x)\big)\big(x-\tilde{\theta}(x)\big)\tilde{\mu}_V(dx) - \iint \log\frac{x-y}{\tilde{\theta}(x)-\tilde{\theta}(y)}\tilde{\mu}_V(dx)\,\tilde{\mu}_V(dy) \geqslant 0. \end{aligned}$$

Using the fact that $\tilde{V}(x) - \frac{\rho}{2}x^2$ is convex on each interval $(-\infty, 0)$ and $(0, \infty)$ combined with the fact that $x$ and $\tilde{\theta}(x)$ have the same sign, the rest of the proof is the same as the one of Theorem 5.

For the case $V(x) = rx - s\log(x)$, using $\theta(x) = (\sqrt{x} + m)^2$, one can show that $\rho = r$ is sharp. $\square$

At last, we would like to discuss a Poincaré type inequality in this context. As in Section 7, for the heuristics, we consider the general model of random matrices with distribution

$$\begin{aligned} \mathbb{P}_n(dM) &= C_n e^{-nr\operatorname{Tr} M}(\det M)^{sn}\, dM = C_n e^{-n\operatorname{Tr}(rM - s\log(M))}\, dM \\ &= C_n e^{-n\operatorname{Tr}(V(M))}\, dM \end{aligned} \tag{10.14}$$

where $dM$ stands for the Lebesgue measure on $n \times n$ positive definite matrices and $s \geqslant 0$. For a given smooth compactly supported function $\phi : [0, \infty) \to \mathbb{R}$, we want to apply the Brascamp–Lieb inequality [6] to the function $\Phi(M) = \operatorname{Tr}\phi(M)$ on the space of positive definite matrices. Now, $\nabla\Phi(M) = \phi'(M)$.

The Hessian of $\Psi(M) := \operatorname{Tr}(V(M))$ can be interpreted as a linear map from $\mathcal{H}_n$ ($n \times n$ Hermitian matrices) into itself which is given by $\nabla^2\Psi(M)X = sM^{-1}XM^{-1}$. Hence the inverse of the Hessian is then $(\nabla^2\Psi(M))^{-1}X = \frac{1}{s}MXM$. Thus we obtain from Brascamp-Lieb that

$$\int \frac{1}{n}\operatorname{Tr}\big(\big(\nabla^2\Psi(M)\big)^{-1}\phi'(M)^2\big)\mathbb{P}_n(dM) \geqslant \operatorname{Var}_{\mathbb{P}_n}\big(\Phi(M)\big).$$

On the other hand, from [20] or [8] the variance of $\Phi(M)$ converges to $\frac{1}{4}\operatorname{Var}_{\mathrm{arcsine}_{[a,b]}}(\phi)$, where we recall that $\mathrm{arcsine}_{[a,b]} = \frac{dx}{\pi\sqrt{(x-a)(b-x)}}$ is the arcsine law on the support $[a,b]$ of $\mu_V$. Next, $\frac{1}{n}\operatorname{Tr}((\nabla^2\Psi(M))^{-1}\phi'(M)^2) = \frac{1}{sn}\operatorname{Tr}((\phi'(M)M)^2)$, whose integral against $\mathbb{P}_n$ converges to the integral of $\frac{1}{s}x^2\phi'(x)^2$ against the equilibrium measure $\mu_V$ from Eq. (10.1). These considerations suggest that

$$\int x^2\phi'(x)^2\mu_V(dx) \geqslant \frac{s}{4}\operatorname{Var}_{\mathrm{arcsine}_{[a,b]}}(\phi). \tag{10.15}$$

Notice here that one can actually make this heuristic into an actual proof of this inequality.

Motivated by these heuristics and also inspired by Theorem 8, we have the following stronger result.

**Theorem 15.** *Assume that $Q : [0, \infty) \to \mathbb{R}$ is a convex potential and let $V(x) = Q(x) - s\log(x)$ for $s > 0$ satisfy $\lim_{x\to\infty}(V(x) - 2\log(x)) = \infty$. Assume that the support of $\mu_V$ is $[a,b]$. Then for any smooth function $\phi$ on $[a,b]$, the following holds,*

$$\begin{aligned} \int x^2\phi'(x)^2\mu_V(dx) \geqslant \frac{s}{4\pi^2}\int_a^b\int_a^b \left(\frac{\phi(x) - \phi(y)}{x - y}\right)^2 \\ \times \frac{-2ab + (a+b)(x+y) - 2xy}{2\sqrt{(x-a)(b-x)}\sqrt{(y-a)(b-y)}}\, dx\, dy. \end{aligned} \tag{10.16}$$

*If $Q(x) = rx + t$, equality is attained for $\phi(x) = c_1 + \frac{c_2}{x}$, therefore* (10.16) *is sharp.*

*In particular, combining* (10.16) *with* (9.8) *for* $\lambda = 0$*, we get an improvement of* (10.15) *as*

$$\int x^2\phi'(x)^2\mu_V(dx) \geqslant \frac{s}{2}, \mathrm{Var}_{\mathrm{arcsine}_{[a,b]}}(\phi).$$

*Equality though is attained only for* $\phi$ *identically* 0.

*In the case* $V(x) = rx$*,* $r > 0$*, on* $[0, \infty)$*, there is no constant* $C > 0$ *such that inequality* (10.16) *holds with* $C$ *instead of* $s/4\pi^2$*. Nevertheless, for every smooth* $\phi$ *on* $[a, b]$*, the following holds,*

$$\int x\phi'(x)^2\mu_V(dx) \geqslant \frac{r}{4\pi^2}\int_a^b\int_a^b\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \times \frac{-2ab+(a+b)(x+y)-2xy}{2\sqrt{(x-a)(b-x)}\sqrt{(y-a)(b-y)}}\,dx\,dy, \tag{10.17}$$

*with equality for* $\phi(x) = c_1 + c_2x$*.*

As remarked after the statement of Theorem 8, the numerator in (10.17) is nonnegative.

**Proof.** The same argument as in the proof of Theorem 8, shows that the density $g(x)$ of $\mu_V$ satisfies

$$g(x) \geqslant \frac{s\sqrt{(x-a)(b-x)}}{2\pi x\sqrt{ab}},$$

therefore it suffices to show that

$$\frac{1}{\pi\sqrt{ab}}\int_a^b x\phi'(x)^2\sqrt{(x-a)(b-x)}\,dx \geqslant \frac{1}{2\pi^2}\int_a^b\int_a^b\left(\frac{\phi(x)-\phi(y)}{x-y}\right)^2 \times \frac{-2ab+(a+b)(x+y)-2xy}{2\sqrt{(x-a)(b-x)}\sqrt{(y-a)(b-y)}}\,dx\,dy.$$

Next, making the change of variable $x = (a+b)/2 + u(b-a)/2$ and denoting $\zeta(u) = \phi((a+b)/2 + u(b-a)/2)$, we reduce the problem to showing that for any smooth function $\phi$ on $[-1, 1]$, we have

$$\frac{1}{\pi\sqrt{ab}}\int_{-1}^{1}\left(\frac{a+b}{2}+\frac{b-a}{2}u\right)\zeta'(u)^2\sqrt{1-u^2}\,du \geqslant \frac{1}{2\pi^2}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\zeta(u)-\zeta(v)}{u-v}\right)^2 \times \frac{1-uv}{\sqrt{1-u^2}\sqrt{1-v^2}}\,du\,dv.$$

Denoting $\beta = \frac{b-a}{b+a}$, we have that $\frac{a+b}{2\sqrt{ab}} = \frac{1}{\sqrt{1-\beta^2}}$, and the preceding inequality reformulates as

$$\int (1+\beta u)\zeta'(u)^2\sqrt{1-u^2}\,du \geqslant \frac{\sqrt{1-\beta^2}}{2\pi}\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\zeta(u)-\zeta(v)}{u-v}\right)^2 \times \frac{1-uv}{\sqrt{1-u^2}\sqrt{1-v^2}}\,du\,dv. \tag{10.18}$$

To show this, take $\psi(t)=\zeta(\cos(t))$ and then after the change of variable $u=\cos(t)$ we need to check

$$\int_0^{\pi}\big(1+\beta\cos(t)\big)\psi'(t)^2\,dt \geqslant \frac{\sqrt{1-\beta^2}}{2\pi}\int_0^{\pi}\int_0^{\pi}\left(\frac{\psi(t)-\psi(s)}{\cos(t)-\cos(s)}\right)^2\big(1-\cos(t)\cos(s)\big)\,dt\,ds.$$

Writing $\psi(t)=\sum_{n=0}^{\infty}a_n\cos(nt)$ and using that $\psi'(t)=-\sum_{n=1}^{\infty}na_n\sin(nt)$, together with the fact that

$$\int_0^{\pi}\cos(t)\sin(nt)\sin(mt)\,dt=\begin{cases}\frac{\pi}{4} & \text{for } |m-n|=1,\\ 0 & \text{otherwise,}\end{cases}$$

and Eq. (7.11), the inequality becomes

$$\sum_{n\geqslant 1}\big(n^2a_n^2+\beta n(n+1)a_na_{n+1}\big)\geqslant\sqrt{1-\beta^2}\sum_{n\geqslant 1}na_n^2. \tag{10.19}$$

Let $\delta=\frac{1-\sqrt{1-\beta^2}}{\beta}$ be the solution $0<\delta<1$ of $\beta\delta^2-2\delta+\beta=0$. Notice that for any $n\geqslant 1$, we have

$$a_na_{n+1}\geqslant-\frac{\delta}{2}a_n^2-\frac{1}{2\delta}a_{n+1}^2$$

which implies that

$$\begin{aligned}\sum_{n\geqslant 1}\big(n^2a_n^2+\beta n(n+1)a_na_{n+1}\big)&\geqslant\sum_{n\geqslant 1}\left(n^2a_n^2-\frac{\beta n(n+1)}{2}\left(\delta a_n^2+\frac{1}{\delta}a_{n+1}^2\right)\right)\\&=\sum_{n\geqslant 1}\frac{n\beta(1-\delta^2)}{2\delta}a_n^2=\sqrt{1-\beta^2}\sum_{n\geqslant 1}na_n^2,\end{aligned}$$

what we had to prove. Notice here that equality is attained in this inequality if and only if $a_{n+1}=-\delta a_n$ for all $n\geqslant 1$, which means that $a_n=(-1)^{n-1}\delta^{n-1}a_1$. This corresponds to the function $\psi(t)=a_1\frac{\delta+\cos t}{1+\delta^2+2\delta\cos t}$, or $\zeta(u)=a_1\frac{\delta+u}{1+\delta^2+2\delta u}$ which means that $\phi(x)=a_1(r-s/x)$. Therefore equality holds also for $\phi(x)=c_1+c_2/x$.

For the second part, in the case $V(x) = rx$ with $r > 0$, notice that if there is a $C > 0$ so that (10.16) holds with $C$ instead of $s/4\pi^2$, then, following the same argument as above, we would have the equivalent of (10.19) as

$$\sum_{n\geqslant 1}\bigl(n^2a_n^2 + n(n+1)a_n a_{n+1}\bigr) \geqslant C\sum_{n\geqslant 1} n a_n^2.$$

Taking in this $a_n = \frac{(-\gamma)^n}{n}$ for $0 < \gamma < 1$, we have that $\gamma^2/(\gamma+1) \geqslant -C\log(1-\gamma^2)$, and this is certainly false for $\gamma$ close to 1.

For Eq. (10.17), notice that in this case the equilibrium measure is $\mu_V(dx) = \frac{r\sqrt{b-x}}{2\pi\sqrt{x}}$ and then after a simple rescaling this follows from Eq. (7.8). This complete the proof of the theorem. □

It is interesting to look at this inequality as a spectral gap result as in Section 9. For example in the case of the Marcenko–Pastur measure ($Q(x) = rx$), the inequality (10.16) is actually equivalent to inequality (10.18). Using the interpretation from Section 9, we can rephrase this as, for a given $\beta \in (0, 1)$,

$$\int (1+\beta x)\bigl(1-x^2\bigr)\phi'(x)^2\nu_0(dx) \geqslant \sqrt{1-\beta^2}\langle N\phi, \phi\rangle_{\nu_0}$$

where $\nu_0$ is the arcsine law on $[-1, 1]$ and $N$ is the number operator. Now we can define the operator

$$L_\beta\phi(x) = -(1+\beta x)\bigl(1-x^2\bigr)\phi''(x) - \bigl(\beta - x - 2\beta x^2\bigr)\phi'(x).$$

With this definition,

$$\langle L_\beta\phi, \phi\rangle_{\nu_0} = \frac{1}{\pi}\int (1+\beta x)\phi'(x)^2\sqrt{1-x^2}\,dx$$

and then inequality (10.18) becomes

$$\langle L_\beta\phi, \phi\rangle_{\nu_0} \geqslant \sqrt{1-\beta^2}\,\langle N\phi, \phi\rangle_{\nu_0}$$

for any smooth function $\phi$ on $[-1, 1]$. In particular this means that $L_\beta \geqslant \sqrt{1-\beta^2}N$. On the other hand it is clear that the operator $L_\beta$ can not be diagonalized by the Chebyshev polynomials of the first kind, therefore the orthogonal polynomial approach given in Section 9 does not work the same way here.

**Remark 6.** We want to point out that for the case $V(x) = rx - s\log(x)$ for $r > 0$ and $s \geqslant 0$, the parameter $r$ appears in the transportation, Log-Sobolev and HWI, while the parameter $s$ plays the dominant role in the Poincaré inequality.

## Acknowledgments

We would like to thank D. Cabanal-Duvillard for pointing to us the formula of the fluctuation for Wishart ensembles and for informing us about his Log-Sobolev conjecture in [7]. Many thanks to the anonymous referee for the pertinent and scholarly comments which pointed several shortcomings of the submitted version and led to an overall improvement of this paper.

## References

[1] G. Ben Arous, A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy, Probab. Theory Related Fields 108 (2) (1997) 183–215.
[2] D. Bakry, L'hypercontractivité et son utilisation en théorie des semigroupes, in: Lectures on Probability Theory, Saint–Flour, 1992, in: Lecture Notes in Math., vol. 1581, Springer, Berlin, 1994, pp. 1–114.
[3] P. Biane, Logarithmic Sobolev inequalities, matrix models and free entropy, Acta Math. Sin. (Engl. Ser.) 19 (3) (2003) 497–506.
[4] G. Blower, The Gaussian isoperimetric inequality and transportation, Positivity 7 (3) (2003) 203–224.
[5] S.G. Bobkov, I. Gentil, M. Ledoux, Hypercontractivity of Hamilton–Jacobi equations, J. Math. Pures Appl. (9) 80 (7) (2001) 669–696.
[6] H.J. Brascamp, E.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. Funct. Anal. 22 (4) (1976) 366–389.
[7] T. Cabanal-Duvillard, Probabilités libres et calcul stochastique, application aux grandes matrices aléatoires, Université Paris VI PhD thesis, 1999.
[8] T. Cabanal-Duvillard, Fluctuations de la loi empirique de grandes matrices aléatoires, Ann. Inst. H. Poincaré Probab. Statist. 37 (3) (2001) 373–402.
[9] D. Cordero-Erausquin, Some applications of mass transport to Gaussian-type inequalities, Arch. Ration. Mech. Anal. 161 (3) (2002) 257–269.
[10] I. Csiszár, Information-type measures of difference of probability distributions and indirect observations, Studia Sci. Math. Hungar. 2 (1967) 299–318.
[11] P.A. Deift, Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach, Courant Lecture Notes in Math., vol. 3, New York University, Courant Institute of Mathematical Sciences, New York, 1999.
[12] R.J. Gardner, The Brunn–Minkowski inequality, Bull. Amer. Math. Soc. (N.S.) 39 (3) (2002) 355–405 (in electronic).
[13] N.A. Gozlan, A characterization of dimension free concentration in terms of transportation inequalities, preprint, 2008.
[14] L. Gross, Logarithmic Sobolev inequalities, Amer. J. Math. 97 (4) (1975) 1061–1083.
[15] U. Haagerup, Seminar notes on free probability.
[16] F. Hiai, D. Petz, Eigenvalue density of the Wishart matrix and large deviations, Infin. Dimens. Anal. Quantum Probab. 1 (1998) 633–646.
[17] F. Hiai, D. Petz, The Semicircle Law, Free Random Variables and Entropy, Math. Surveys Monogr., vol. 77, Amer. Math. Soc., Providence, RI, 2000.
[18] F. Hiai, D. Petz, Y. Ueda, Free transportation cost inequalities via random matrix approximation, Probab. Theory Related Fields 130 (2004) 199–221.
[19] K. Johansson, On fluctuations of random hermitian matrices, Duke Math. J. 91 (1998) 1–24.
[20] D. Jonsson, Some limit theorems for the eigenvalues of a sample covariance matrix, J. Multivariate Anal. 12 (1) (1982) 1–38.
[21] J.H.B. Kemperman, On the optimum rate of transmitting information, Ann. Math. Statist. 40 (1969) 2156–2177.
[22] A.M. Khorunzhy, B.A. Khoruzhenko, L. Pastur, Asymptotic properties of large random matrices with independent entries, J. Math. Phys. 37 (10) (1996) 5033–5060.
[23] M. Ledoux, The Concentration of Measure Phenomenon, Math. Surveys Monogr., vol. 89, Amer. Math. Soc., Providence, RI, 2001.
[24] M. Ledoux, A (one-dimensional) free Brunn–Minkowski inequality, C. R. Acad. Sci. Paris 340 (2005) 301–304.
[25] F. Otto, C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, J. Funct. Anal. 173 (2) (2000) 361–400.

[26] I. Popescu, Talagrand inequality for the semicircular law and energy of the eigenvalues of beta ensembles, Math. Res. Lett. 14 (6) (2007) 1023–1032.
[27] E.B. Saff, V. Totik, Logarithmic Potentials with External Fields, Grundlehren Math. Wiss. (Fundamental Principles of Mathematical Sciences), vol. 316, Springer-Verlag, Berlin, 1997.
[28] M. Talagrand, Transportation cost for Gaussian and other product measures, Geom. Funct. Anal. 6 (3) (1996) 587–600.
[29] C. Villani, Topics in Optimal Transportation, Grad. Stud. Math., vol. 58, Amer. Math. Soc., Providence, RI, 2003.
[30] D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory. V. Noncommutative Hilbert transforms, Invent. Math. 132 (1) (1998) 189–227.
[31] D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory. I, Comm. Math. Phys. 155 (1) (1993) 71–92.
[32] F.-Y. Wang, Functional Inequalities, Markov Properties and Spectral Theory, Science Press, Beijing–New York, 2005.
[33] G.N. Watson, Notes on generating functions of polynomials—(3) polynomials of Legendre and Gegenbauer, J. London Math. Soc. 8 (1933) 289–292.