STABILITY FOR FUNCTIONAL AND GEOMETRIC INEQUALITIES AND A STOCHASTIC REPRESENTATION OF FRACTIONAL INTEGRALS AND NONLOCAL OPERATORS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Daesung Kim

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Rodrigo Bañuelos, Chair

Department of Mathematics

Dr. Emanuel Indrei

Department of Mathematics

Dr. Jonathon Peterson

Department of Mathematics

Dr. Samy Tindel

Department of Mathematics

Approved by:

Dr. David Goldberg

Head of the Graduate Program

To Jiwon and Saein

ACKNOWLEDGMENTS

I owe my greatest thanks to my advisor Rodrigo Bañuelos for his consistent encouragement and support. I have greatly benefitted from our weekly meeting and his insightful guidance over the last six years at Purdue. I am grateful that he introduced me to a variety of interesting problems from martingale transforms to functional inequalities, which has allowed me to have a broad perspective on analysis and probability.

I would like to thank the committee members: Professors Emanuel Indrei, Jonathon Peterson, and Samy Tindel. I am grateful to Emanuel for our discussions and collaboration on the stability problems on functional inequalities and optimal transportation. I am grateful to Professor Jonathon Peterson for organizing reading courses on discrete probability topics, which I have really enjoyed.

I would like to express my thanks to my parents and sister for their invaluable support. To my wife, Jiwon, I owe my sincere thanks for her consistent encouragement, support, and understanding. Without her, it would have been impossible to accomplish what I have now.

I give my all thanks and praises to the one who became to me "*wisdom from God, righteousness and sanctification and redemption*".

TABLE OF CONTENTS

Page

Page

ABSTRACT

Kim, Daesung PhD, Purdue University, August 2019. Stability for Functional and Geometric Inequalities and a Stochastic Representation of Fractional Integrals and Nonlocal Operators. Major Professor: Rodrigo Bañuelos.

The dissertation consists of two research topics.

The first research direction is to study stability of functional and geometric inequalities. Stability problem is to estimate the deficit of a functional or geometric inequality in terms of the distance from the class of optimizers or a functional that identifies the optimizers. In particular, we investigate the logarithmic Sobolev inequality, the Beckner–Hirschman inequality (the entropic uncertainty principle), and isoperimetric type inequalities for the expected lifetime of Brownian motion.

In Chapter 3, we derive several types of stability estimates of the logarithmic Sobolev inequality in terms of the Wasserstein distance, L^p distances, and the Kolmogorov distance. We consider the spaces of probability measures satisfying different conditions on the second moments, the lower bounds of the density, and some integrability of the density. To obtain these results, we employ the optimal transport technique, Fourier analysis, and probability theoretic approach. In Chapter 4, we construct an example to understand the conditions on the space and the distance under which stability of the logarithmic Sobolev inequality does not hold. As an application, we show that stability of the Beckner–Hirschman inequality does not hold for the normalized L^p distance with some weighted measures in Chapter 5.

In Chapter 6, we study quantitative improvements of the inequalities for the expected lifetime of Brownian motion, which state that the L^p-norms of the expected lifetime in a bounded domain for $1 \leq p \leq \infty$, are maximized when the region is a ball with the same volume. Since the inequalities also hold for a general class of Lévy

processes, it is interesting to see if the quantitative improvement can be extended to general Lévy processes. We discuss the related open problems in that direction.

The second topic of the thesis is a stochastic representation of fractional integrals and nonlocal operators. In Chapter 7, we extend the Hardy–Littlewood–Sobolev inequality to symmetric Markov semigroups. To this end, we construct a stochastic representation of the fractional integral using the background radiation process. The inequality follows from a new inequality for the fractional Littlewood–Paley square function. In Chapter 8, we prove the Hardy–Stein identity for non-symmetric pure jump Lévy processes and the L^p boundedness of a certain class of Fourier multiplier operators arising from non-symmetric pure jump Lévy processes. The proof is based on Itô's formula for general jump processes and the symmetrization of Lévy processes.

1. INTRODUCTION

The thesis consists of two parts. The first subject of this thesis is stability of functional and geometric inequalities. The second subject is the Littlewood–Paley inequality and its applications.

1.1 Stability of functional and geometric inequalities

We present some terminology regarding stability problems, introduced by Carlen [41]. Consider nonnegative functionals $\mathcal{G}$ and $\mathcal{H}$ defined on a class of admissible functions or sets X. A functional or geometric inequality can be written as

$$\mathcal{G}(u) \geq \mathcal{H}(u) \tag{1.1.1}$$

for all $u \in X$. The inequality is called *sharp* if for each $\lambda > 1$ there exists $u_\lambda \in X$ such that $\mathcal{G}(u_\lambda) < \lambda \mathcal{H}(u_\lambda)$. It is called *optimal* if there exists $u_0 \in X$ such that $\mathcal{G}(u_0) = \mathcal{H}(u_0)$. Such u_0 is called an *optimizer*. The deficit is defined by $\delta(u) = \mathcal{G}(u) - \mathcal{H}(u) \geq 0$. Once the class of optimizers X_0 is characterized, a natural question is to measure the deviation of u from the class of optimizers when $\delta(u)$ gets close to 0. Let $\mathrm{d} : X \times X \to [0, \infty)$ be a distance defined on X. We say the inequality is d-*stable* in X if for any sequence $\{u_k\}$ in X, $\delta(u_k) \to 0$ as $k \to \infty$ implies

$$\lim_{k\to\infty} \mathrm{d}(u_k, X_0) = \lim_{k\to\infty} \inf_{v\in X_0} \mathrm{d}(u_k, v) = 0. \tag{1.1.2}$$

In particular, a *stability estimate* or a *quantitative improvement* of the inequality is a lower bound of the deficit in terms of the distance

$$\delta(u) \geq \Phi(\mathrm{d}(u, X_0)) \tag{1.1.3}$$

for all u and for some modulus of continuity Φ. Sometimes, instead of a distance, we consider a nonnegative functional on X that identifies the class of optimizers.

Namely, consider a functional $\mathrm{d} : X \to [0, \infty]$ such that $\mathrm{d}(u) = 0$ if and only if $u \in X_0$. Stability with respect to this functional is defined in the same way. In contrast to (1.1.3), (1.1.2) it is also called a *non-quantitative result* or *weak stability*.

Recently, finding stability estimates has become of significant interest in the study of functional and geometric inequalities; the Sobolev inequalities [23, 44, 48], the Hardy–Littlewood–Sobolev inequality [41], the logarithmic Sobolev inequality [53, 57, 58, 82, 83, 85], the Hausdorff–Young inequality [45], the isoperimetric inequalities [59, 64, 65], and the Faber–Krahn inequalities [32, 34]. In particular, there have been great efforts to find sharp stability results. A stability estimate is sharp if the modulus of continuity Φ is best possible. That is, if Φ cannot be replaced by any other modulus of continuity Ψ which satisfies

$$\lim_{t \to 0} \frac{\Phi(t)}{\Psi(t)} = 0.$$

Another direction is to study the best possible admissible space and distance in which stability of (1.1.1) holds. Even though the inequality (1.1.1) holds for all $u \in X$, it is possible that a stability estimate holds only for $u \in \tilde{X} \subset X$. In this case, one can ask what is the largest possible subset $\tilde{X}$ of X in which stability of (1.1.1) is valid.

1.1.1 The sharp quantitative isoperimetric inequality

As an example, we review stability results for the classical isoperimetric inequality. Let D be a Borel set in $\mathbb{R}^n$, then the classical isoperimetric inequality states that

$$P(D) \geq P(B) \tag{1.1.4}$$

where B is a ball in $\mathbb{R}^n$ with $|D| = |B|$ and $P(E)$ denotes the perimeter of E. This is sharp and optimal: equality holds in (1.1.4) if and only if D is a ball. The deficit of (1.1.4) is defined by

$$\delta(D) = \frac{P(D) - P(B)}{P(B)}$$

where B is a ball with $|D| = |B|$. Fuglede [63] showed that if D is convex then there exists $\kappa(n), C_n > 0$ such that

$$\delta(D) \geq C_n a(D)^{\kappa(n)}$$

where $a(D) = \inf\{\mathrm{d}_H(D, x + B) : x \in \mathbb{R}^n\}$, d_H is the Hausdorff distance, and B is a ball with $|D| = |B|$. He constructed an one-parameter family of domains to show that $\kappa(n)$ is the sharp exponent. Note that the asymmetry $a(D)$ is not appropriate for general non-convex sets. For example, if D is a ball in $\mathbb{R}^n$ $(n \geq 3)$ with a long and thin tail, then $a(D)$ could be large whereas the deficit is close to 0. Thus it is natural to deal with the Fraenkel asymmetry

$$A(D) = \inf\left\{\frac{|D \triangle (x + B)|}{|D|} : x \in \mathbb{R}^n, B \text{ is a ball with } |B| = |D|\right\}$$

for a general stability estimate. Hall [74] proved that if D has an axis of symmetry, then

$$\delta(D) \geq C_n A(D)^2 \tag{1.1.5}$$

with an explicit dimensional constant C_n. For a general class of sets, he used the Steiner symmetrization and the estimate from [75] to deduce (1.1.5) with the exponent 4. It was conjectured that the sharp exponent is 2. One can see this by considering an ellipse which is very close to a ball; see [75, pp. 88–89]. Fusco, Maggi, and Pratelli [64] gave an affirmative answer to the conjecture. They proved (1.1.5) for a Borel set with finite volume.

1.1.2 The logarithmic Sobolev inequality

In Chapter 3, we study stability of the logarithmic Sobolev inequality. In Chapter 4, we investigate conditions on probability measure spaces and metrics under which the LSI is not stable. As an application, we discuss instability of the Beckner–Hirschman inequality in Chapter 5. Chapter 3 is based on joint work with Emanuel Indrei [82], and Chapter 4 and 5 are based on my work [85].

Let $d\gamma$ be the standard Gaussian measure on $\mathbb{R}^n$. The classical logarithmic Sobolev inequality (the LSI) states that for a probability measure $f d\gamma$

$$\frac{1}{2}\,\mathrm{I}(f) = \frac{1}{2}\int \frac{|\nabla f|^2}{f}\,d\gamma \geq \int f \log f \, d\gamma = \mathrm{H}(f) \tag{1.1.6}$$

where $\mathrm{I}(f)$ and $\mathrm{H}(f)$ are the Fisher information and the relative entropy respectively. Note that I and H are nonnegative functionals and well-defined on the space of probability measures $f d\gamma$ with $\sqrt{f} \in W^{1,2}(\mathbb{R}^n, d\gamma)$. The constant $\frac{1}{2}$ is sharp and equality holds if and only if $f(x) = e^{b\cdot x - |b|^2/2}$ for some $b \in \mathbb{R}^n$.

In Chapter 3, we explore various probability measure spaces and metrics in which stability of the LSI holds. To be specific, we find several types of lower bounds of the deficit $\delta(f) := \frac{1}{2}\,\mathrm{I}(f) - \mathrm{H}(f)$ in terms of the Wasserstein distances, the Kolmogorov distance, and L^p distances for $p \geq 1$, under different conditions on the function f. To obtain these results, we employ several different techniques: optimal transport theory, Fourier analysis, and probability.

We considered the space of probability measures on $\mathbb{R}^n$ whose second moments are bounded by $M > 0$, denoted by $\mathcal{P}_2^M(\mathbb{R}^n)$. The first main result (Theorem 3.2.1) is to show that if $f d\gamma$ is a centered probability measure in $\mathcal{P}_2^M(\mathbb{R})$ then

$$\delta(f) \geq C_M \|f - 1\|^4_{L^1(d\gamma)}. \tag{1.1.7}$$

The proof is mainly based on the optimal transport technique, which was introduced by Cordero-Erausquin [50] and adapted to the context of stability of the LSI by [83] and thereafter [57]. We consider the Brenier map between $f d\gamma$ and $d\gamma$, which is the solution to the optimal transportation problem. First, we derive W_1 stability of the LSI (Theorem 3.2.6) from that of Talagrand's transportation inequality, which was obtained by [18] in dimension 1, [51] for higher dimensions (see also [57] for $W_{1,1}$-stability). By a lower bound of the deficit (2.2.7) which follows from the Monge–Ampère equation, we derive L^1-stability (1.1.7). Under different assumptions on f (see (3.2.2) and (3.2.3)), we exploit the deficit bound (2.2.7) of Cordero-Erausquin to show that the deficit is bounded below by the L^1 distance of $\log f$ from some

affine function (Theorem 3.2.9). Combining W_1–stability and a compactness argument (the Rellich–Kondrachov theorem), we also derive non-quantitative L^1–stability in $\mathcal{P}_2^M(\mathbb{R}^n)$, for $n \geq 2$ (Theorem 3.2.13).

Compared to the previous results [57, 83], our stability results in $\mathcal{P}_2^M$ can be thought of as an extension in terms of probability measure spaces. Indrei and Marcon [83] showed W_2–stability in a class of probability measures $f d\gamma$ such that $(-1+\varepsilon) \leq D^2(-\log f) \leq M$ for $\varepsilon, M > 0$. The proof is based on the optimal transport technique (2.2.7). In [57], a strict improvement of the LSI for the class of probability measures that satisfy a $(2,2)$-Poincaré inequality was proved, which yields stability bounds with respect to W_2 and L^1. One can see that these spaces are contained in $\mathcal{P}_2^M$ for some M. Note that the authors in [57] also considered stability estimates in a general probability measure space via Talagrand's transportation inequality.

The second approach is concerned with the deficit bound (2.2.4) derived by Carlen [40] (see Theorem 2.2.1). To characterize the case of equality in (1.1.6), Carlen [40] derived a lower bound of the deficit in terms of the relative entropy of the Fourier–Wiener transform from the entropic uncertainty principle, which was conjectured by Hirschman [80] and proven by Beckner [20]. By investigating the behavior of the relative entropy of the Fourier–Wiener transform when the deficit gets close to 0, we obtain non-quantitative L^1-stability (Theorem 3.2.14). Applying the optimal transport technique to the Fourier–Wiener transform, we obtain a lower bound of the deficit which holds for a wide class of functions (Theorem 3.2.16). As a corollary of this bound, we prove non-quantitative L^1–stability under some integrability assumptions (Corollary 3.2.18 and 3.2.19) .

From the probabilistic point of view, we derive stability estimates in terms of the Kolmogorov distance. The proof is mainly based on the quantitative versions of Cramér's theorem of [25, 68, 103]. Cramér's theorem says that if the sum of two independent random variables has a normal distribution, then both random variables are normal. Combining quantitative versions of Cramér's theorem (Theorem 2.3.1 and 2.3.2) with a convolution type deficit bound of the LSI in [58] (see Theorem

3.4.1), we derive stability estimates in terms of the Kolmogorov distance under some moment assumptions (Theorem 3.2.21 and 3.2.22).

In the process of finding the best possible function spaces and metrics, a natural question is whether the previous stability results can be improved. In Chapter 4, we give a partial answer by showing that there exists a sequence of centered probability measures in $\mathcal{P}_2^M(\mathbb{R})$ such that the deficit converges to 0 but the distance from the optimizer does not converge to 0 in terms of W_2 and L^p for $p > 1$ (Theorem 4.1.1). Furthermore, we construct a sequence of centered probability measures in $\mathcal{P}_2(\mathbb{R})$ such that the deficit converges to 0 and the W_1-distance from the Gaussian measure goes to ∞ (Theorem 4.1.2). The implication of these results is that W_2–stability of [26] and W_1–stability of Theorem 3.2.6 are sharp in terms of $\mathcal{P}_2^M(\mathbb{R}^n)$ and that L^1–stability in $\mathcal{P}_2^M(\mathbb{R})$ of Theorem 3.2.1 is sharp in terms of the L^p distances.

To construct a sequence of probability measures whose deficit converges to 0, we start with the Gaussian measure with a small perturbation in the tails. It turns out that this perturbation controls the second moment and the relative entropy, which leads to the desired result.

The problem of finding the best possible function space and metric remains open. The most general space in the setting of the LSI is the space of probability measures with finite second moments, since if the relative entropy is finite then so is the second moment.

In Chapter 5, we prove that stability of the entropic uncertainty principle, which is also called the Beckner–Hirschman inequality (BHI), fails with respect to the normalized L^p distances some weighted measures. For a nonnegative function h in $L^2(\mathbb{R})$ with $\|h\|_2 = 1$, the entropic uncertainty principle states that

$$\delta_{\mathrm{BH}}(h) = S(|h|^2) + S(|\widehat{h}|^2) - (1 - \log 2) \geq 0$$

where $S(\cdot)$ denotes Shannon's entropy, $\widehat{h}$ is the Fourier transform of h, and $\delta_{\mathrm{BH}}(h)$ is the deficit of the Beckner–Hirschman inequality. Carlen [40] showed that the deficit of the LSI is bounded below by that of the BHI, which implies that the example constructed in Chapter 4 has a small deficit of the BHI. With careful computation,

we show that there exists a sequence of L^2 normalized functions such that the deficit converges to 0 but the distance from the class of optimizers does not. In these results, we consider the L^p distances with polynomial (Theorem 5.1.1) and exponential weights (Theorem 5.1.2).

The entropic uncertainty principle was first proposed by Hirschman [80], and proved by Beckner [20] by differentiating the sharp Hausdorff–Young inequality with respect to the exponent. Inspired by the quantitative Hausdorff–Young inequality of Christ [45], it is natural to ask if there is a stability estimate for the BHI. The heuristic consideration in Chapter 5 suggests that the BHI could be stable in terms of the L^2 distance, which is an interesting open problem.

1.1.3 The expected lifetime of Brownian motion

In Chapter 6, we investigate stability of isoperimetric type inequalities arising from stochastic analysis and their relation to geometric inequalities. This is based on my work [86].

Let $\alpha \in (0, 2]$ and D a bounded domain in $\mathbb{R}^n$. Let X_t^α be the symmetric α–stable process with generator $-(-\Delta)^{\alpha/2}$. The first exit time of X_t^α from D is defined by

$$\tau_D^\alpha = \inf\{t \geq 0 : X_t^\alpha \notin D\}$$

and the expected lifetime by $u_D^\alpha(x) = \mathbb{E}^x[\tau_D^\alpha]$, where $\mathbb{E}^x$ is the expectation associated with X_t^α starting at $x \in \mathbb{R}^n$. For $\alpha = 2$, X_t^α is Brownian motion with generator Δ.

Bañuelos and Méndez-Hernández [16] showed that several isoperimetric type inequalities for Brownian motion continue to hold for a wide class of Lévy processes using the symmetrization of Lévy processes and the multiple integral rearrangement inequalities of Brascamp–Lieb–Luttinger [31]. Indeed, they proved that if Y_t is a Lévy process, its Lévy measure is absolutely continuous with respect to the Lebesgue

measure, and f and V are nonnegative continuous functions, then for any $x \in D$ and $t > 0$,

$$\mathbb{E}^0[f^*(Y_t^*) \exp\left(\int_0^t V^*(Y_s^*)\, ds\right); \tau_B^{Y_t^*} > t] \geq \mathbb{E}^x[f(Y_t) \exp\left(\int_0^t V(Y_s)\, ds\right); \tau_D^{Y_t} > t]$$

where f^* and V^* are the symmetric decreasing rearrangements of f and V, Y_t^* is the symmetrization of Y_t, and B is a ball centered at 0 with $|D| = |B|$; see [16, Theorem 1.4]. A particular case of this is that for all $t \geq 0$ and $x \in \mathbb{R}^n$,

$$\mathbb{P}^0(\tau_B^\alpha > t) \geq \mathbb{P}^x(\tau_D^\alpha > t), \tag{1.1.8}$$

which yields

$$u_B^\alpha(0) \geq u_D^\alpha(x), \tag{1.1.9}$$

where B is a ball centered at 0 with $|B| = |D|$. In fact, (1.1.8) gives

$$\mathbb{E}^0(\tau_B^\alpha)^p \geq \mathbb{E}^x(\tau_D^\alpha)^p \tag{1.1.10}$$

for all $p > 0$. Talenti [110] proved that the L^p norm of a solution of a second-order elliptic equation is maximized when the elliptic operator and the domain are symmetrically rearranged. In particular, the result yields that for $p > 0$, $\alpha = 2$, and a bounded domain D,

$$\|u_B\|_p \geq \|u_D\|_p \tag{1.1.11}$$

where B is a ball with $|B| = |D|$.

Given the above isoperimetric type inequalities for the first exit times of the α-stable processes and their connection to the classical torsion function, there are many questions that arise concerning quantitative versions of these inequalities. The goal of Chapter 6 is to study quantitative versions of the expected lifetime inequalities (1.1.9) for $\alpha = 2$ and (1.1.11) for $p \geq 1$.

The first main result is a lower bound of the deficit of (1.1.9) in terms of the deviations of x and D from the optimizers. Note that equality holds in (6.2.1) if D is a ball and $u_D(x) = \max_{y \in D} u_D(y)$. The deviation of x is represented by $|\{y \in D :$

$u_D(y) > u_D(x)\}|$, and the deviation of D by the Fraenkel asymmetry, which is defined by

$$A(D) = \inf\left\{\frac{|D\triangle B|}{|D|} : B \text{ is a ball with } |B| = |D|\right\}. \tag{1.1.12}$$

The proof is based on the proof of (1.1.9) for $\alpha = 2$ in [6, 110], and the sharp quantitative isoperimetric inequality [64]. In order to estimate the asymmetry of the level set, we use the idea of Hansen and Nadirashvili [76] as in the proof of the boosted Pólya–Szegö inequality [33, Lemma 2.9].

The second result is a quantitative inequality for the L^p norm of the expected lifetime (1.1.11), $1 \le p \le \infty$. We define the L^p deficit of the expected lifetime inequality for $1 \le p \le \infty$ by

$$\delta_p(D) = 1 - \left(\frac{\|u_D\|_p}{\|u_B\|_p}\right)^{\kappa(p)}$$

where $\kappa(p) = p$ for $1 \le p < \infty$, $\kappa(\infty) = 1$, and B is a ball centered at 0 with $|B| = |D|$. Let $n \ge 2$ and D be a bounded domain in $\mathbb{R}^n$. For $1 \le p \le \infty$, we have

$$\delta_p(D) \ge C_{n,p} A(D)^{2+\kappa(p)}. \tag{1.1.13}$$

The torsional rigidity of D is defined by $T(D) = \|u_D\|_1$. In this context, we call u_D the torsion function of D. The Saint-Venant inequality states that the torsional rigidity is maximized when the region is a ball. If $p = 1$, the result produces the non-sharp quantitative Saint-Venant inequality

$$T(B) - T(D) \ge C_{n,1} T(B) A(D)^3, \tag{1.1.14}$$

which was proven in [33]. Thus the result can be thought of as an extension of (1.1.14) to the case $1 < p \le \infty$. Note that Brasco, De Philippis, and Velichkov [34] showed that the sharp exponent of (1.1.14) is 2 in the sense that the power cannot be replaced by any smaller number. Their method, however, does not give an explicit dimensional constant because the proof relies on the selection principle of Cicalese and Leonardi [49].

The key step in the proof is the removal of t_* defined in (6.2.3). In [33], the authors proved a non-sharp quantitative Saint-Venant inequality using the boosted Pólya–Szegö inequality. In the proof, they used the variational representation for $T(D)$ to replace the term t_* by $A(D)$ (up to dimensional constant). In our case, however, the L^p norm of the expected lifetime does not have an appropriate variational formula for $1 < p \leq \infty$. To overcome this difficulty, we find a critical level t_0 which is comparable to $A(D)$ and use the layer cake representation of the L^p norm for $p \in (1, \infty)$ and the strong Markov property for $p = \infty$.

The fractional analogue of (1.1.14) is proven in [32]. They showed that if $n \geq 2$, $\alpha \in (0, 2)$, and D is an open set with $|D| = 1$, then

$$T_\alpha(B) - T_\alpha(D) \geq C_{n,\alpha} A(D)^{\frac{6}{\alpha}}$$

where $C_{n,\alpha}$ is explicit and B is a ball with $|B| = 1$. Here $T_\alpha(D)$ is the fractional torsional rigidity defined in (6.4.3). Furthermore, they proved that if D has Lipschitz boundary and satisfies the exterior ball condition, then the exponent can be lowered to $2 + \frac{2}{\alpha}$. It turns out that our method for removing t_* yields the same exponent without any additional geometric assumptions on D.

1.2 Littlewood–Paley inequality

Littlewood–Paley square (quadratic) functions have been of interest for many years with many applications in harmonic analysis and probability. On the analysis side, these include the classical square functions obtained from the Poisson semigroup as in [106] and more general heat semigroups as in [107]. On the probability side, these correspond to the celebrated Burkholder–Gundy inequalities which are of fundamental importance in modern stochastic analysis.

Littlewood–Paley L^p inequalities have played an important role in a broad area of analysis and probability. These inequalities give a nice way of understanding the qualitative and quantitative properties of functions and operators. In the classical case,

the L^p inequalities for square functions are obtained from the Calderón–Zygmund theory, which relies on the property of harmonic functions.

In Chapter 7, we introduce a fractional analogue of the Littlewood–Paley square function and derive an L^p–L^q inequality for the square function. It turns out that the square function and its inequality hold for a general setting. As an application, we prove a Hardy–Littlewood–Sobolev inequality for symmetric Markov semigroups. This is based on my work [84].

In [106], Stein provided an alternative approach to obtaining the L^p bound for the square function using the so-called Hardy–Stein identity. In Chapter 8, we extend this to non-symmetric pure jump Lévy processes and derive the L^p inequalities for the corresponding square functions. As an application, we define a certain class of the Fourier multipliers and prove the L^p boundedness of the multipliers. This is based on joint work with Rodrigo Bañuelos [11].

1.2.1 The Hardy–Littlewood–Sobolev inequality

The Hardy–Littlewood–Sobolev (HLS) inequality, first derived by [77, 78, 105], states that

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{f(x)g(y)}{|x-y|^{n-\alpha}}dxdy \le C_{n,\alpha,p}\|f\|_p\|g\|_r$$

where $1 < p < q < \infty$, $1 = \frac{1}{p} + \frac{1}{r} - \frac{\alpha}{n}$, $0 < \alpha < n$, $f \in L^p(\mathbb{R}^n)$, and $g \in L^r(\mathbb{R}^n)$. Lieb [90] showed the existence of optimizers and obtained the explicit formulas for optimizers in special cases. In light of its geometric implications, a subsequent problem is to extend the sharp HLS inequality to a more general setting than $\mathbb{R}^n$. Because Lieb's result is based on rearrangement techniques which do not apply to outside of $\mathbb{R}^n$, it is necessary to find a new way of proving Lieb's inequality. There have been several attempts along this line, for instance [42, 60]. Frank and Lieb [61] extended the sharp HLS inequality to the Heisenberg group using a radically new method.

In Chapter 7, we study an extension of the HLS inequality to symmetric Markov semigroups. We give a stochastic representation for the fractional integrals for sym-

metric Markov semigroups and derived an analogue of the HLS inequality for the semigroups. The stochastic representation is based on the techniques of Gundy and Varopoulos [70–72] where the background radiation processes and time reversal were used to obtain the probabilistic representations for the Riesz transforms. The representation is a variation of the one used by Applebaum and Bañuelos [3], which is based on the space-time Brownian motion and martingale inequalities. Unlike the space-time Brownian motion representation which requires a gradient in the space variable (or a carré du champ), the representation in Theorem 7.2.1 only requires the time derivative which is well defined for general semigroups.

To prove the (non-sharp) HLS inequality for symmetric Markov semigroups, we introduce a fractional Littlewood–Paley square function for symmetric Markov semigroups and derive a new L^p–L^q inequality for the square function. The proof is based on the ergodic inequality for maximal functions, the optimal splitting technique of [79, 106], and an estimate for the classical Littlewood–Paley square functions in [107].

The basic question, in connection with the problem of finding the sharp inequality, is how to bypass the Littlewood–Paley square function method and the optimal splitting argument. This optimal splitting is also a key step in the proof of Applebaum and Bañuelos [3], although it is done in combination with the Burkholder–Davis–Gundy inequalities.

The stochastic representation of the fractional integral can be thought of as a martingale transform where the predictable process is not bounded. Martingale transform techniques have been used quite effectively in the study of singular integral operators, particularly in obtaining optimal, or nearly optimal inequalities. Given the powerful martingale and Bellman function methods pioneered by Burkholder [37] to obtain sharp inequalities for martingale transforms and their many subsequent uses in various problems in analysis and probability, it is natural to ask if these techniques can be extended to martingale transforms with unbounded predictable processes and provide a different proof of the sharp HLS inequalities which could be extended to

other settings. At this moment, it is unclear how to obtain the sharp results with the Bellman function methods. This remains an interesting challenging problem.

1.2.2 Hardy–Stein identity for nonlocal operators and Fourier multipliers

Littlewood–Paley square functions and their L^p inequalities have been extensively studied with applications in the study of function spaces, PDEs, and Fourier multiplier operators. From the probabilistic point of view, square functions and the L^p inequalities correspond to the quadratic variations of martingales and the Burkholder–Davis–Gundy inequalities. In the classical case, the L^p inequalities for square functions are obtained from the Calderón–Zygmund theory, which relies on the property of harmonic functions. In [106], Stein provided an alternate approach to obtaining the L^p bound for the square function when $1 < p < 2$. Using the chain rule and Green's theorem, he derived the so-called Hardy–Stein identity [106, Equation (16), p.88], which states that for $f \in L^p(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} |f|^p dx = \int_0^\infty \int_{\mathbb{R}^d} y \Delta u^p dx dy$$

where u is the harmonic extension of f to the upper half-space. This approach can be adapted to more general diffusion operators for which the chain rule holds.

In [10], the authors extended the Littlewood–Paley L^p inequalities for $1 < p < \infty$ to nonlocal operators arising from symmetric pure jump Lévy processes. Their proof is based on the Burkholder–Gundy inequalities and the Hardy–Stein type identity for symmetric pure jump Lévy processes. As an application, they introduced a certain family of Fourier multiplier operators and proved the L^p boundedness.

In Chapter 8, we extend the Hardy–Stein identity of [10] to non-symmetric pure jump Lévy processes. For $a, b \in \mathbb{R}$ and $p \in (1, \infty)$, let $F(a, b; p)$ be the second-order Taylor remainder of the maps $x \mapsto |x|^p$ given by $F(a, b; p) = |b|^p - |a|^p - pa|a|^{p-2}(b-a)$. Let P_t be the semigroup corresponding to a non-symmetric pure jump Lévy process

and ν the Lévy measure. In this setting, we prove the following Hardy–Stein identity (Theorem 8.3.1): for $f \in L^p(\mathbb{R}^d)$ and $1 < p < \infty$,

$$\int_{\mathbb{R}^d} |f|^p dx = \int_{\mathbb{R}^d} \int_0^\infty \int_{\mathbb{R}^d} F(P_t f(x), P_t f(x+y); p)\, \nu(dy) dt dx.$$

Compared to the result of [10] where the authors used properties of the semigroups, our proof relies on Itô's formula for general jump processes, which allows us to extend the identity to non-symmetric cases. Furthermore, it gives a Hardy–Stein type identity for uniformly integrable martingales in $L^2 \cap L^p$ (Theorem 8.3.5).

We also prove the L^p–boundedness of a certain class of Fourier multiplier operators for non-symmetric pure jump Lévy processes (Theorem 8.4.1). Since the two-sided L^p–inequalities for square functions rely heavily on the symmetry of the Lévy measures, the application to Fourier multiplies also requires it. To bypass this difficulty, we employ the symmetrization technique as in [16].

2. PRELIMINARIES

2.1 Probability metrics

2.1.1 The Wasserstein distances

For $p \geq 1$ and a probability measure μ, the p-th moment of μ is given by $m_p(\mu) = \int_{\mathbb{R}^n} |x|^p d\mu$. We say that μ has *finite p-th moment* if $m_p(\mu) < \infty$. The space of probability measures with finite p-th moments is denoted by $\mathcal{P}_p(\mathbb{R}^n)$. The *Wasserstein distance of order p* between two probability measures $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^n)$ is

$$W_p(\mu,\nu) = \inf_{\pi} \left(\iint_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^p d\pi(x,y) \right)^{\frac{1}{p}}$$

where the infimum is taken over all probability measures π on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ and ν. In general, one can define the optimal transportation cost with a cost function $c(x,y)$ on $\mathbb{R}^n \times \mathbb{R}^n$ by

$$\mathcal{T}_c(\mu,\nu) = \inf_{\pi} \left(\iint_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y) d\pi(x,y) \right). \tag{2.1.1}$$

In particular, W_1 is called the Kantorovich–Rubinstein distance and W_2 is called the quadratic Wasserstein distance.

For $p \geq 1$, W_p defines a metric on $\mathcal{P}_p(\mathbb{R}^n)$. For $p_1 < p_2$ and probability measures $\mu, \nu \in \mathcal{P}_{p_2}(\mathbb{R}^n)$, it follows from Jensen's inequality that $W_{p_1}(\mu,\nu) \leq W_{p_2}(\mu,\nu)$ and $\mathcal{P}_{p_2}(\mathbb{R}^n) \subseteq \mathcal{P}_{p_1}(\mathbb{R}^n)$. The Wasserstein distance of order p is stronger than the weak convergence: let ν_k be a sequence of probability measures in $\mathcal{P}_p(\mathbb{R}^n)$, then ν_k converges to μ in W_p if and only if $\nu_k \rightharpoonup \mu$ weakly and $m_p(\nu_k) \to m_p(\mu)$ as $k \to \infty$.

Let μ and ν be probability measures with finite second moments. Then there exists a map $T : \mathbb{R}^n \to \mathbb{R}^n$ such that $\nu(A) = \mu(T^{-1}(A))$ for all Borel sets A in $\mathbb{R}^n$ and

$$W_2^2(\mu,\nu) = \int_{\mathbb{R}^n} |T(x) - x|^2 d\mu.$$

It is well-known that the map T is uniquely determined μ-almost everywhere and is the gradient of a convex function φ such that $T = \nabla\varphi$. The map is called the *Brenier map*.

We say a function φ is 1-Lipschitz if $|\varphi(x) - \varphi(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}^n$. The Kantorovich–Rubinstein distance W_1 has a dual form

$$W_1(\mu, \nu) = \sup\left\{\int_{\mathbb{R}^n} \varphi(d\mu - d\nu) : \varphi \in L^1(d|\mu - \nu|), \varphi \text{ is 1-Lipschitz.}\right\}.$$

On the real line, we have explicit formulas for W_1. For probability measures μ and ν on $\mathbb{R}$, let F and G be the distribution functions of μ and ν. Then the W_1 distance between μ and ν can be written as

$$W_1(\mu, \nu) = \int_0^1 |F^{-1}(t) - G^{-1}(t)|dt = \int_{\mathbb{R}} |F(x) - G(x)|dx.$$

Let γ be the Gaussian measure and $d\nu = f d\gamma$. The relative entropy functional $\nu \mapsto \mathrm{H}(f)$ is stronger than the total variation distance but weaker than the L^p-norm for $p > 1$ in a sense that

$$2\|f - 1\|^2_{L^1(d\gamma)} \leq \mathrm{H}(f) \leq \frac{2}{p-1}\|f - 1\|^p_{L^p(d\gamma)} + 2\|f - 1\|_{L^p(d\gamma)}. \tag{2.1.2}$$

The first inequality is called Pinsker's inequality and the second inequality follows from Hölder's inequality and the fact that $t \log t \leq \frac{2}{p-1}|t - 1|^p + 2|t - 1|$, for all $t \geq 0$ (see [56, p.93]). In particular, the second inequality tells us that if the relative entropy does not converge to zero then $f d\gamma$ does not converge to $d\gamma$ in L^p for $p > 1$, which is a key ingredient in the proof of Theorem 4.1.1.

Talagrand [108] introduced the inequality

$$\delta_{\mathrm{Tal}}(f) = 2\mathrm{H}(f) - W_2^2(f d\gamma, d\gamma) \geq 0 \tag{2.1.3}$$

where $\delta_{\mathrm{Tal}}(f)$ is the deficit of Talagrand's transportation inequality. This implies that the relative entropy is stronger than the quadratic Wasserstein distance. Otto and Villani [99] proved that the LSI implies Talagrand's transportation inequality. If $\nu \in \mathcal{P}_2$ is centered, then Cordero-Erausquin [51] showed

$$\delta_{\mathrm{Tal}}(f) \geq C \min\left\{W_1^2(f d\gamma, d\gamma), W_1(f d\gamma, d\gamma)\right\}. \tag{2.1.4}$$

Note that a comparable stability result was also shown in [57] . The quantitative Talagrand's transportation inequality is one of the main ingredients in the proof of Theorem 3.2.6. Otto and Villani proved the HWI inequality which is an "interpolation" inequality between the relative entropy, the Wasserstein distance, and the Fisher information

$$\mathrm{H}(f) \leq W_2(d\nu, d\gamma)\sqrt{\mathrm{I}(f)} - \frac{1}{2}W_2^2(d\nu, d\gamma). \tag{2.1.5}$$

We refer the reader to [1, 113] for further details.

2.1.2 The total variation distance

Let μ and ν be probability measures. The total variation distance between μ and ν is defined by

$$d_{\mathrm{TV}}(\mu, \nu) = \sup_A |\mu(A) - \nu(A)|$$

where the supremum is taken over all Borel sets in $\mathbb{R}^n$ and yields a stronger topology than the weak topology. That is, if $d_{\mathrm{TV}}(\mu, \nu_k) \to 0$ as $k \to \infty$, then ν_k converges weakly to μ (however, the converse does not hold). The total variation distance can be thought of as the optimal transportation distance (2.1.1) with $c(x, y) = \mathbb{1}_{\{x \neq y\}}$. It has a dual form

$$d_{\mathrm{TV}}(\mu, \nu) = \sup_{0 \leq |\varphi| \leq 1} \int_{\mathbb{R}^n} \varphi(d\mu - d\nu).$$

If $d\nu = f d\mu$, then the total variation distance $d_{\mathrm{TV}}(\mu, \nu)$ can be written in terms of the L^1–norm

$$d_{\mathrm{TV}}(\mu, \nu) = \frac{1}{2}\|f - 1\|_{L^1(d\mu)}.$$

It is well-known that the total variation distance is comparable to the Hellinger distance

$$\|\sqrt{f} - 1\|_{L^2(d\mu)}^2 \leq \|f - 1\|_{L^1(d\mu)} \leq 2\|\sqrt{f} - 1\|_{L^2(d\mu)}. \tag{2.1.6}$$

2.1.3 Comparison between probability metrics

This subsection is devoted to introduce probability metrics and investigate their relations. The following is based on [25, 67, 113].

Definition 2.1.1. Let $(\Omega, \mathscr{F}, \lambda)$ be a measure space. For probability measures $d\mu = f d\lambda$ and $d\nu = g d\lambda$, the Hellinger distance is defined by

$$d_{\mathrm{H}}(\mu, \nu) = \Big(\int_{\Omega} |\sqrt{f} - \sqrt{g}|^2 d\lambda \Big)^{\frac{1}{2}}.$$

Note that d_{H} is a metric and $0 \leq d_{\mathrm{H}}(\mu, \nu) \leq \sqrt{2}$.

Definition 2.1.2. Let $(\Omega, \mathscr{F})$ be a measurable space. Let μ and ν be probability measures on $(\Omega, \mathscr{F})$. The total variation distance is

$$d_{\mathrm{TV}}(\mu, \nu) = \sup_{h} \left| \int_{\Omega} h d\mu - \int_{\Omega} h d\nu \right|$$

where the supremum is taken over all measurable functions $h : \Omega \to \mathbb{R}$ with $|h(x)| \leq 1$.

Definition 2.1.3. Let (Ω, d) be a Polish space. Let μ and ν be probability measures on Ω. For a Borel set B and $\varepsilon > 0$, $B^{\varepsilon} = \{x \in \Omega : \inf_{y \in B} d(x, y) \leq \varepsilon\}$. The Prokhorov metric is defined by

$$d_{\mathrm{P}}(\mu, \nu) = \inf\{\varepsilon > 0 : \mu(B) \leq \nu(B^{\varepsilon}) + \varepsilon \text{ for all Borel sets } B\}.$$

If X and Y random variables with the laws μ and ν, then it follows from Strassen's theorem that

$$d_{\mathrm{P}}(\mu, \nu) = \inf_{\mathbb{P}} \{\varepsilon > 0 : \mathbb{P}(d(X, Y) > \varepsilon) < \varepsilon\}$$

where the infimum is taken over all joint distributions of X and Y. Similarly, we have

$$d_{\mathrm{TV}}(\mu, \nu) = \inf \mathbb{E}[\mathbb{1}_{\{X \neq Y\}}] = \sup\{\mu(F) - \nu(F) : F \text{ closed}\}.$$

Definition 2.1.4. The Kolmogorov distance between two probability measures μ and ν on $\mathbb{R}$ is given by

$$d_{\mathrm{K}}(\mu, \nu) = \sup_{x \in \mathbb{R}} |\mu((-\infty, x]) - \nu((-\infty, x])|.$$

If F and G are distribution functions of μ and ν, then we denote by $d_{\mathrm{K}}(F,G) = d_{\mathrm{K}}(\mu,\nu)$. One can see that $0 \le d_{\mathrm{K}}(\mu,\nu) \le 1$.

Definition 2.1.5. Let μ and ν be probability measures on $\mathbb{R}$ with distribution functions F and G. The Lévy metric is defined by

$$d_{\mathrm{L}}(\mu,\nu) = d_{\mathrm{L}}(F,G) = \inf\{\varepsilon > 0 : G(x-\varepsilon) - \varepsilon \le F(x) \le G(x+\varepsilon) + \varepsilon, \forall x \in \mathbb{R}\}.$$

Proposition 2.1.6. *Let μ and ν be probability measures on $\mathbb{R}$, then we have*

$$\begin{aligned} d_{\mathrm{L}}(\mu,\nu) &\le \min\{d_{\mathrm{K}}(\mu,\nu), d_{\mathrm{P}}(\mu,\nu)\} \\ &\le \max\{d_{\mathrm{K}}(\mu,\nu), d_{\mathrm{P}}(\mu,\nu)\} \\ &\le \min\{d_{\mathrm{TV}}(\mu,\nu), \sqrt{W_1(\mu,\nu)}\}. \end{aligned}$$

Proposition 2.1.7. *Let μ be a probability measure on $\mathbb{R}$ and γ the standard Gaussian measure on $\mathbb{R}$, then*

$$d_{\mathrm{K}}(\mu,\gamma) \le 2 d_{\mathrm{P}}(\mu,\gamma).$$

Proposition 2.1.8 ([25, Proposition A.1.2])**.** *Let $\mu,\nu \in \mathcal{P}_2^M(\mathbb{R})$, then*

$$\begin{aligned} W_1(\mu,\nu) &\le 2 d_{\mathrm{L}}(\mu,\nu) + 2\sqrt{M} d_{\mathrm{L}}(\mu,\nu)^{1/2}, \\ W_1(\mu,\nu) &\le 4\sqrt{M} d_{\mathrm{K}}(\mu,\nu)^{1/2}. \end{aligned}$$

Proposition 2.1.9. *Let Ω be a measurable space. Let μ and ν be probability measures on Ω, then*

$$d_{\mathrm{H}}(\mu,\nu)^2 \le d_{\mathrm{TV}}(\mu,\nu) \le 2 d_{\mathrm{H}}(\mu,\nu).$$

2.2 The LSI deficit bounds

Let $d\gamma = (2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}} dx$ be the standard Gaussian measure on $\mathbb{R}^n$ and f a non-negative function in $L^1(d\gamma)$ such that $d\nu = f d\gamma$ is a probability measure. We define the Fisher information and the relative entropy of f with respect to γ by

$$\begin{aligned} \mathrm{I}(\nu) = \mathrm{I}(f) &= \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\gamma, \\ \mathrm{H}(\nu) = \mathrm{H}(f) &= \int_{\mathbb{R}^n} f \log f d\gamma. \end{aligned}$$

The classical logarithmic Sobolev inequality (the LSI) states that

$$\delta(f) = \frac{1}{2}\,\mathrm{I}(f) - \mathrm{H}(f) \geq 0. \tag{2.2.1}$$

We call $\delta(f)$ the deficit of the LSI. In this section, we discuss some estimates on the LSI deficit that we will call upon later.

2.2.1 Carlen's deficit estimate

Carlen [40] characterized the equality cases in two ways: if $f \in L^p(\mathbb{R}^{2n})$ is a product function in (x, y) and $\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right)$, then f and its factors are Gaussian functions. Thereafter, he proved a Minkowski-type inequality and derived the strict superadditivity of the Fisher information. Combining this with the factorization theorem, he deduced that equality holds in (2.2.1) only if $e^{b\cdot x - \frac{b^2}{2}}$, $b \in \mathbb{R}^n$.

The second proof is based on the Beckner–Hirschman entropic uncertainty principle. Indeed, he derived a lower bound of the LSI deficit in terms of the relative entropy of the Fourier–Wiener transform, which leads to the characterization of the equality cases.

Let $g(x) := 2^{\frac{n}{4}} e^{-\pi|x|^2}$ and $dm = g(x)^2 dx$. The Fourier transform of f in $L^2(\mathbb{R}^n)$ is

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} f(x)\, dx.$$

Let $U : L^2(dx) \to L^2(dm)$ be defined by $f \mapsto f/g$ and $\mathcal{W} := U\mathcal{F}U^*$ on $L^2(dm)$ where U^* is the adjoint operator of U. The operator $\mathcal{W}$ is called the *Fourier–Wiener transform*. Let $f \in L^2(dm)$ with $\|f\|_{L^2(dm)} = 1$. By the Plancherel theorem, we have $\|\mathcal{W}f\|_{L^2(dm)} = \|f\|_{L^2(dm)} = 1$. The LSI deficit with respect to dm is defined by

$$\delta_c(f) = \frac{1}{2\pi}\int_{\mathbb{R}^n} |\nabla f|^2\, dm - \int_{\mathbb{R}^n} |f|^2 \log|f|^2\, dm.$$

For a probability measure $f d\gamma$, let $u_f(x) = (f(2\sqrt{\pi}x))^{1/2}$. Then $u_f^2 dm$ is a probability measure and $\delta(f) = \delta_c(u_f)$.

For a nonnegative function ρ on $\mathbb{R}^n$ with $\int \rho\, dx = 1$, the entropy of ρ is given by

$$S(\rho) = -\int \rho \log \rho\, dx. \tag{2.2.2}$$

The Beckner–Hirschman inequality [20] states that for a function h with $\int |h|^2 dx = 1$,

$$S(|h|^2) + S(|\mathcal{F}(h)|^2) \geq n(1 - \log 2). \tag{2.2.3}$$

Let μ be a probability measure and f a nonnegative function such that $f\,d\mu$ is a probability measure. The relative entropy of f with respect to μ is denoted by

$$\mathrm{Ent}_\mu(f) = \int f \log f \, d\mu.$$

Theorem 2.2.1 ([40, Theorem 6]). *Let $f \in L^2(dm)$ be normalized, then*

$$\delta_c(f) \geq \mathrm{Ent}_{dm}(|\mathcal{W}f|^2). \tag{2.2.4}$$

Proof. Let $h = U^* f$, then (2.2.3) yields

$$\begin{aligned} S(|h|^2) + S(|\mathcal{F}(h)|^2) &= S(|fg|^2) + S(|\mathcal{W}(f)g|^2) \\ &= \int_{\mathbb{R}} (|h|^2 + |\mathcal{F}(h)|^2)(2\pi|x|^2 - \frac{n}{2}\log 2)\, dx \\ &\quad - (\mathrm{Ent}_{dm}(|f|^2) + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2)) \\ &\geq n(1 - \log 2). \end{aligned}$$

Since $\|h\|_2 = \|\mathcal{F}(h)\|_2 = 1$, it suffices to show that

$$\frac{1}{2\pi}\int (|2\pi x h|^2 + |2\pi x \mathcal{F}(h)|^2)\, dx = \frac{1}{2\pi}\int |\nabla f|^2\, dm + n.$$

Using $\nabla(g^{-1}) = 2\pi x g^{-1}$, $-2\pi x \mathcal{F}(h) = \mathcal{F}(\nabla h)$, and Parseval's formula, we have

$$\begin{aligned} \int (|2\pi x h|^2 + |2\pi x \mathcal{F}(h)|^2) dx &= \int (|\nabla(g^{-1})h|^2 + |g^{-1}\nabla h|^2)\, dm \\ &= \int |\nabla(g^{-1})h + g^{-1}\nabla h|^2\, dm - 2\pi \int x \cdot \nabla(|h|^2)\, dx \\ &= \int |\nabla f|^2\, dm + 2\pi n, \end{aligned}$$

which finishes the proof. □

Remark 2.2.2. For $h \in L^2(dx)$ with $\int |h|^2\, dx = 1$, we define the deficit of the Beckner–Hirschman inequality by

$$\delta_{\mathrm{BH}}(h) = S(|h|^2) + S(|\widehat{h}|^2) - n(1 - \log 2).$$

In fact, the proof of Theorem 2.2.1 yields

$$\delta_c(f) - \int_{\mathbb{R}^n} |W(f)|^2 \log |W(f)|^2 dm = \delta_{\mathrm{BH}}(fg). \tag{2.2.5}$$

Remark 2.2.3. Suppose $f \geq 0$ and $\delta_c(f) = 0$, then Theorem 2.2.1 yields $\mathrm{Ent}_{dm}(|\mathcal{W}f|^2) = 0$. By Cramér's theorem, one obtains $f(x) = e^{2\pi(b\cdot x - \frac{|b|^2}{2})}$ for some $b \in \mathbb{R}^n$, which is equivalent to the cases of equality in (2.2.1). Indeed, since $|\mathcal{W}f|^2 = 1$ a.e., we have

$$|\mathcal{F}(h)(x)|^2 = \mathcal{F}(h(x))\mathcal{F}(h(-x)) = g^2(x)$$

where $h(x) = g(x)f(x)$. By the Fourier inversion theorem,

$$\int h(x)h(x+y)\,dy = 2^{-\frac{n}{2}} e^{-\frac{\pi|x|^2}{2}},$$

which yields in turn that h is Gaussian by Cramér's theorem. Since $|\mathcal{W}f| = 1$, we get $\mathcal{W}f = e^{ib\cdot x}$ as desired.

In §3.2.2, we investigate the lower bound $\mathrm{Ent}_{dm}(|\mathcal{W}f|^2)$ to obtain weak stability of the LSI. Combining Carlen's estimate with the optimal transport method, we also get several types of deficit bounds which hold for a wide class of probability measures.

2.2.2 Optimal transport method

Let μ and ν be Borel probability measures on $\mathbb{R}^n$. We say that a map $T : \mathbb{R}^n \to \mathbb{R}^n$ pushes μ forward to ν if $\nu(B) = \mu(T^{-1}(B))$ for every Borel set $B \subset \mathbb{R}^n$. Brenier [36] and McCann [94] showed that if μ is absolutely continuous with respect to the Lebesgue measure, then there exists a convex function φ such that $T = \nabla\varphi$ pushes μ forward to ν and $\nabla\varphi$ is uniquely determined μ-a.s. If μ and ν have finite second moments, then $\pi_0 = (Id \times \nabla\varphi)$ is the optimal plan for

$$W_2(\mu,\nu)^2 = \inf_{\pi} \iint |x-y|^2\, d\pi(x,y) = \int |x - T(x)|^2\, d\mu(x),$$

where the infimum is taken over all probability measures π on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ and ν.

Cordero-Erausquin [50] used the Brenier map to derive the following inequality that holds for a wide class of probability measures, which entails the logarithmic Sobolev inequalities, Talagrand's transport inequalities, and the HWI inequalities.

Theorem 2.2.4 ([50, Theorem 1]). *Let μ be a probability measure on $\mathbb{R}^n$ of the form $d\mu(x) = e^{-V(x)}dx$, where V is a twice differentiable function satisfying* $\operatorname{Hess} V \geq c$ *for some $c \in \mathbb{R}$. Let $f, g : \mathbb{R}^n \to [0, \infty)$ be non-negative compactly supported functions. Assume that $f \in C^1$ and $\int f d\mu = \int g d\mu$. If $T(x) = x + \nabla\theta$ is the Brenier map pushing $f d\mu$ forward to $g d\mu$, then*

$$\operatorname{Ent}_\mu(g) \geq \operatorname{Ent}_\mu(f) + \int \nabla f \cdot \nabla\theta d\mu + \frac{c}{2}\int f|\nabla\theta|^2 d\mu + \int \left(\Delta_A\theta - \log\det(I + \operatorname{Hess}\theta)\right) f d\mu \quad (2.2.6)$$

where Δ_A denotes the Aleksandrov Laplacian.

Remark 2.2.5. For $\mu = \gamma$ and $g \equiv 1$ (using an approximation argument), (2.2.6) yields

$$\delta(f) \geq \frac{1}{2}\int |\nabla(\log f) + \nabla\theta|^2 f d\gamma + \int \left(\Delta_A\theta - \log\det(I + \operatorname{Hess}\theta)\right) f d\mu. \quad (2.2.7)$$

Remark 2.2.6. If we apply this theorem to $\mu = dm$ (i.e. $V(x) = 2\pi|x|^2$ and $c = 4\pi$), then one can see

$$\frac{1}{2\pi}\int |\nabla f|^2 \, dm + \operatorname{Ent}_{dm}(|g|^2) \geq \operatorname{Ent}_{dm}(|f|^2) + \frac{1}{2\pi}\int |2\pi f\nabla\theta + \nabla f|^2 \, dm$$

where $T(x) = x + \nabla\theta$ is the Brenier map pushing $|f|^2 dm$ forward to $|g|^2 dm$. In particular, if $g = \mathcal{W}f$, then

$$\delta_c(f) + \operatorname{Ent}_{dm}(|\mathcal{W}f|^2) \geq \frac{1}{2\pi}\int |2\pi f\nabla\theta + \nabla f|^2 \, dm. \quad (2.2.8)$$

2.2.3 Scaling asymmetry of the logarithmic Sobolev inequality

Following the proof of [53, Proposition 1], we obtain a lower bound of the deficit in terms of the second moment and the relative entropy.

Proposition 2.2.7. *If $d\nu = f d\gamma \in \mathcal{P}_2(\mathbb{R}^n)$ and $\sqrt{f} \in W^{1,2}(\mathbb{R}^n, d\gamma)$, then*

$$\delta(f) \geq \frac{1}{4n}\left((m_2(\gamma) - m_2(\nu)) + 2\mathrm{H}(f)\right)^2. \tag{2.2.9}$$

Proof. Let $u \in H^1(\mathbb{R}^n, d\gamma)$ be such that $\int |u|^2 d\gamma = 1$ and $\int |x|^2|u|^2 d\gamma = s < \infty$. Let $d\gamma = \gamma(x)dx$. We define $v = u\sqrt{\gamma}$, then $\int |v|^2 dx = 1$ and $\int |x|^2|v|^2 dx = s$. Direct computations show that

$$|\nabla v|^2 = |\sqrt{\gamma}\nabla u - \frac{1}{2}u\sqrt{\gamma}x|^2 = |\nabla u|^2\gamma + \frac{1}{4}|x|^2|u|^2\gamma + \frac{1}{2}\nabla(u^2)\cdot\nabla\gamma,$$

and

$$\begin{aligned}\int |\nabla v|^2 dx &= \int |\nabla u|^2 d\gamma + \frac{1}{4}\int |x|^2|u|^2 d\gamma + \frac{1}{2}\int \nabla(u^2)\cdot\nabla\gamma dx \\ &= \int |\nabla u|^2 d\gamma + \frac{1}{4}\int |x|^2|u|^2 d\gamma - \frac{1}{2}\int u^2(-n + |x|^2)d\gamma \\ &= \int |\nabla u|^2 d\gamma - \frac{1}{4}\int |x|^2|u|^2 d\gamma + \frac{n}{2}\int |u|^2 d\gamma.\end{aligned}$$

Similarly, we have

$$\int |v|^2 \log|v|^2 dx = \int |u|^2 \log|u|^2 d\gamma - \frac{1}{2}\int |x|^2|u|^2 d\gamma - \frac{n}{2}\log(2\pi)\int |u|^2 d\gamma.$$

It then follows from the LSI with respect to γ that

$$\int |\nabla u|^2 d\gamma - \frac{1}{2}\int |u|^2 \log|u|^2 d\gamma = \int |\nabla v|^2 dx - \frac{1}{2}\int |v|^2 \log|v|^2 dx - \frac{n}{4}\log(2\pi e^2).$$

Let $w(x) := \lambda^{-\frac{n}{2}} v(x/\lambda)$ for $\lambda > 0$, then $\int_{\mathbb{R}^n} |w|^2 dx = 1$, $\int_{\mathbb{R}^n} |x|^2|w|^2 dx = \lambda^2 s$,

$$\int |\nabla v|^2 dx = \lambda^2 \int |\nabla w|^2 dx,$$

$$\int |v|^2 \log|v|^2 dx = \int |w|^2 \log|w|^2 dx + n\log\lambda.$$

The LSI with respect to the Lebesgue measure yields

$$\lambda^2 \int |\nabla w|^2 dx - \frac{n}{2}\log\lambda \geq \frac{1}{2}\int |w|^2 \log|w|^2 dx + \frac{n}{4}\log(2\pi e^2).$$

Optimizing the LHS in λ, we have

$$\int |\nabla w|^2 dx \geq \frac{n\pi e}{2}\exp\Big(\frac{2}{n}\int |w|^2 \log|w|^2 dx\Big).$$

Let $w = \sqrt{f\gamma}$, $d\nu = f d\gamma$, and $A = \frac{1}{n}(2\mathrm{H}(f) + (m_2(\gamma) - m_2(\nu))$, then

$$\delta(f) \geq \frac{n}{2}(e^A - 1 - A) \geq \frac{1}{4n}(2\mathrm{H}(f) + (m_2(\gamma) - m_2(\nu))^2.$$

□

2.3 Stability for Cramér's theorem

Cramér's theorem says that if the sum of two independent random variables has a normal distribution, then both random variables are normal. Let X and Y be independent random variables with distribution functions F and G respectively, then the Kolmogorov distance between X and Y is given by

$$d_{\mathrm{K}}(F,G) = \sup_{x\in\mathbb{R}} |F(x) - G(x)|.$$

Let $F * G$ be the distribution of the sum $X + Y$ so that it is defined by

$$F * G(x) = \int_{\mathbb{R}} F(x-y)dG(y).$$

If p_1 and p_2 are density functions of X and Y, one can write it as

$$F * G(x) = \int_{-\infty}^{x} p_1 * p_2(t)dt.$$

Let $\gamma_{b,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|x-b|^2}{2\sigma^2}}$ be the Gaussian density with mean b, variance σ^2, and $\Phi_{b,\sigma}$ its distribution function. For simplicity, let $\Phi_\sigma := \Phi_{0,\sigma}$ and $\Phi := \Phi_{0,1}$. We recall the following stability result of Cramér's theorm from [68, 103].

Theorem 2.3.1 ([25, Theorem 2.2]). *Let $\varepsilon > 0$ and $N = N(\varepsilon) = 1 + \sqrt{2\log(1/\varepsilon)}$. Let X_1, X_2 be random variables with distribution functions F_1, F_2. We also put*

$$a_i = \int_{-N}^{N} x dF_i(x), \quad \sigma_i^2 = \int_{-N}^{N} x^2 dF_i(x) - a_i^2$$

for $i = 1, 2$. Suppose that F_1 and F_2 have median zero and $\sigma_1, \sigma_2 > 0$. If $d_{\mathrm{K}}(F_1 * F_2, \Phi) \le \varepsilon < 1$, then there exist absolute constants $C_1, C_2 > 0$ such that for $i = 1, 2$,

$$d_{\mathrm{K}}(F_i, \Phi_{a_i,\sigma_i}) \le \frac{C_i}{\sigma_i\sqrt{\log(1/\varepsilon)}} \min\left\{\frac{1}{\sqrt{\sigma_i}}, \log\log\frac{e^e}{\varepsilon}\right\}.$$

A general version of the stability result can be found in [25].

Theorem 2.3.2 ([25, Theorem 2.3]). *Let X_1, X_2 be independent random variables with $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$ and $\mathrm{Var}[X_1 + X_2] = 1$. For $i = 1, 2$, let F_i be the distribution*

function of X_i *and* $v_i^2 = \mathrm{Var}(X_i)$. *If* $d_{\mathrm{K}}(F_1 * F_2, \Phi_1) \le \varepsilon < 1$, *then there exists* $C > 0$ *such that*

$$d_{\mathrm{K}}(F_i, \Phi_{v_i}) \le \frac{C}{v_i \sqrt{\log \frac{1}{\varepsilon}}} \min\left\{ \frac{1}{\sqrt{v_i}}, \log\log \frac{e^e}{\varepsilon} \right\}$$

for $i = 1, 2$.

2.4 Entropic uncertainty principle

For a nonnegative function h on $\mathbb{R}^n$, the entropy of h is given by

$$S(h) = -\int_{\mathbb{R}^n} h \log h \, dx.$$

Let $h \in L^2(\mathbb{R}^n)$ with $\|h\|_2 = 1$. The Beckner–Hirschman inequality (the BHI in short) states that

$$S(|h|^2) + S(|\widehat{h}|^2) \ge n(1 - \log 2) \tag{2.4.1}$$

where $\widehat{h}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} h(x) dx$. It is also called the entropic uncertainty principle. By differentiating the (non-sharp) Hausdorff–Young inequality in p at $p = 2$, Hirschman obtained $S(|h|^2) + S(|\widehat{h}|^2) \ge 0$. He conjectured in [80] that the Gaussian functions are extremal for the inequality and the best constant in the right hand side of (2.4.1) is $n(1-\log 2)$. Beckner [20] found the best constant in the Hausdorff–Young inequality for all $p \in [1, 2]$, which gave an affirmative answer to the conjecture.

Even though the Gaussian functions satisfy the equality, it was an open problem to show that the Gaussians are the only optimizers. Lieb [91] characterized the class of optimizers for the Hausdorff–Young inequality and the BHI. Indeed, he proved that every optimizer for a convolution operator with a Gaussian kernel is Gaussian. Equality holds in (2.4.1) if and only if h is of the form

$$h(x) = ce^{-\langle x, Jx\rangle + x\cdot v}$$

where $c \in \mathbb{C}$, $v \in \mathbb{C}^n$, and J is an $n \times n$ real positive definite matrix (see [40, Remarks in p.207]).

2.5 Torsional rigidity

Let $\alpha \in (0,2]$ and D a bounded domain in $\mathbb{R}^n$. Let X_t^α be the rotationally symmetric α–stable process with generator $-(-\Delta)^{\alpha/2}$. The first exit time of X_t^α from D is defined by

$$\tau_D^\alpha = \inf\{t > 0 : X_t^\alpha \notin D\}.$$

The expected lifetime is defined by $u_D^\alpha(x) = \mathbb{E}^x[\tau_D^\alpha]$ where $\mathbb{E}^x$ is the expectation associated with X_t^α starting at $x \in \mathbb{R}^n$. Note that $u_D^\alpha(x)$ is a solution to the equation

$$\begin{cases} (-\Delta)^{\frac{\alpha}{2}} u(x) = 1, & x \in D, \\ u(x) = 0, & x \notin D \end{cases} \tag{2.5.1}$$

in the weak sense. If B is a ball of radius R and centered at the origin, then $u_B^\alpha(x)$ is explicitly given by

$$u_B^\alpha(x) = C_{n,\alpha}(R^2 - |x|^2)^{\frac{\alpha}{2}}.$$

For $\alpha = 2$, X_t^α is Brownian motion with generator Δ. In this case, we drop the superscript α.

The semigroup P_t^α associated with X_t^α killed upon exiting D is given by

$$P_t^\alpha f(x) = \mathbb{E}^x[f(X_t^\alpha); t < \tau_D^\alpha]$$

on $L^2(D)$. The general semigroup theory yields (see [52]) that there exists an orthonormal basis $\{\varphi_n\}$ of $L^2(D)$ and the corresponding eigenvalues $0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots$ such that $P_t^\alpha \varphi_n = e^{-t\lambda_n}\varphi_n$ and $(-\Delta)^{\alpha/2}\varphi_n = \lambda_n \varphi_n$. Using the representation of the transition density of X_t^α

$$p_t(x,y) = \sum_{n=1}^{\infty} e^{-\lambda_n t}\varphi_n(x)\varphi_n(y),$$

one obtains

$$\mathbb{P}^x(\tau_D^\alpha > t) = \int_D p_t(x,y)dy = \sum_{n=1}^{\infty} e^{-\lambda_n t}\|\varphi_n\|_1 \varphi_n(x)$$

and

$$u_D^\alpha(x) = \int_0^\infty \mathbb{P}^x(\tau_D^\alpha > t)dt = \sum_{n=1}^{\infty} \frac{\|\varphi_n\|_1}{\lambda_n}\varphi_n(x).$$

In addition (see [29, Theorem 4.4]), there exist constants c_1, c_2 depending on D and α such that $c_1 u_D^\alpha(x) \le \varphi_1(x) \le c_2 u_D^\alpha(x)$ for all $x \in D$. For further information, we refer the reader to [29] and the references therein.

The classical torsional rigidity of D is defined by $T(D) = \|u_D\|_1$ for $\alpha = 2$. We say that $u_D(x)$ is the torsion function of D. Let $W_0^{1,2}(D)$ be the completion of $C_0^\infty(D)$ with respect to the norm $u \mapsto \|\nabla u\|_2$. We have variational representations of the torsional rigidity

$$\begin{aligned} T(D) &= \max\left\{ \frac{\|u\|_1^2}{\|\nabla u\|_2^2} : u \in W_0^{1,2}(D), u \neq 0 \right\} \\ &= \max\left\{ 2\|u\|_1 - \|\nabla u\|_2^2 : u \in W_0^{1,2}(D), u \neq 0 \right\}. \end{aligned} \tag{2.5.2}$$

Since u_D is an optimizer for the maximization problems, we have $T(D) = \|u_D\|_1 = \|\nabla u_D\|_2^2$. There are two important inequalities for $T(D)$. The Saint-Venant inequality, an isoperimetric type inequality for $T(D)$, states that if D is a set of finite measure in $\mathbb{R}^n$ then

$$|B|^{-\frac{n+2}{n}} T(B) \ge |D|^{-\frac{n+2}{n}} T(D)$$

where B is a ball. The Kohler-Jobin inequality states that for a ball B,

$$\lambda_1(D) T(D)^{\frac{2}{n+2}} \ge \lambda_1(B) T(B)^{\frac{2}{n+2}}.$$

Note that the classical Faber–Krahn inequality for the first eigenvalue λ_1 follows from these two inequalities for $T(D)$:

$$\frac{\lambda_1(D)}{\lambda_1(B)} \ge \left(\frac{T(B)}{T(D)}\right)^{\frac{2}{n+2}} \ge \left(\frac{|B|}{|D|}\right)^{\frac{2}{n}}. \tag{2.5.3}$$

Furthermore, it is well-known [34] that stability of Saint-Venant inequality can be transferred to that of Faber–Krahn inequalities for the first eigenvalues. To see this, suppose that there is a modulus of continuity $\Phi : [0,\infty) \to [0,\infty)$ such that $\Phi(t) = 0$ if and only if $t = 0$, and

$$|B|^{-\frac{n+2}{n}} T(B) - |D|^{-\frac{n+2}{n}} T(D) \ge \Phi(A(D))$$

where $A(D)$ is the Fraenkel asymmetry defined in (1.1.12). Without loss of generality, we assume that $|D| = 1$ and B is a ball with $|B| = 1$. If $T(B) \leq 2T(D)$, it follows from (2.5.3) that

$$\frac{\lambda_1(D)}{\lambda_1(B)} - 1 \geq \left(\frac{T(B)}{T(D)}\right)^{\frac{2}{n+2}} - 1 \geq C_n \left(\frac{T(B)}{T(D)} - 1\right) \geq C_n \Phi(A(D)).$$

If $T(B) > 2T(D)$, then $\lambda_1(D) - \lambda_1(B) \geq c_n$ for some universal constant c_n. Since $0 \leq A(D) < 2$, if there exists $M > 0$ such that $\Phi(x) \leq M$ for all $x \in [0, 2)$, then one can choose $C_{n,M}$ small enough that

$$\lambda_1(D) - \lambda_1(B) \geq C_{n,M} \Phi(A(D)).$$

Thus we obtain

$$|D|^{\frac{2}{n}} \lambda_1(D) - |B|^{\frac{2}{n}} \lambda_1(B) \geq C_{n,M} \Phi(A(D)).$$

This is called the Faber–Krahn hierarchy (see [33, Proposition A.1]).

The fractional torsional rigidity for $0 < \alpha < 2$ is defined by

$$T_\alpha(D) = \int_D u_D^\alpha(x)\, dx = \int_D \int_0^\infty \mathbb{P}^x(\tau_D^\alpha > t)\, dt dx.$$

There has been recent progress in the study of the fractional torsional rigidity. The isoperimetric inequality for $T_\alpha(D)$, a fractional analogue of the Saint-Venant inequality, follows from [16, Corollary 5.4] where the isoperimetric inequality was proven for a general class of Lévy processes. For the stable processes, it also follows from the sharp rearrangement inequality of [62, Theorem A.1]. Recently, Brasco, Cinti, and Vita [32] proved a quantitative improvement of the fractional Saint-Venant inequality. Their method is based on the extension of [39] and the symmetrization argument of [65].

3. STABILITY OF THE LOGARITHMIC SOBOLEV INEQUALITY

We investigate different probability measure spaces and metrics under which the logarithmic Sobolev inequality is stable. We consider the Wasserstein distances, the Kolmogorov distance, and the L^p distances for $p \geq 1$. To obtain these results, we use optimal transport theory, Fourier analysis, and probability. This chapter is based on joint work with Emanuel Indrei [82].

3.1 Setting

Let $d\gamma = (2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}} dx$ be the standard Gaussian measure on $\mathbb{R}^n$ and f a non-negative function in $L^1(d\gamma)$ such that $d\nu = f d\gamma$ is a probability measure. We define the Fisher information and the relative entropy of f with respect to γ by

$$\mathrm{I}(\nu) = \mathrm{I}(f) = \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\gamma,$$
$$\mathrm{H}(\nu) = \mathrm{H}(f) = \int_{\mathbb{R}^n} f \log f d\gamma.$$

The classical logarithmic Sobolev inequality (the LSI) states that

$$\delta(f) = \frac{1}{2}\mathrm{I}(f) - \mathrm{H}(f) \geq 0. \tag{3.1.1}$$

We call $\delta(f)$ the deficit of the LSI. Note that the constant $\frac{1}{2}$ is sharp and $\mathrm{I}(f), \mathrm{H}(f)$ are well-defined if $\sqrt{f} \in W^{1,2}(\mathbb{R}^n, d\gamma)$. Equality holds in (3.1.1) if and only if $e^{b\cdot x - \frac{b^2}{2}}$ for some $b \in \mathbb{R}^n$. Note that the Gaussian measure (that is, $f = 1$) is the only centered optimizer. There are several proofs based on the central limit theorem [69], the Ornstein–Uhlenbeck semigroup [88], the Prékopa–Leindler inequality [27], optimal transport theory [50], and harmonic analysis [21, 40].

We are interested in measuring the deviation of a centered probability measure $d\nu = f d\gamma$ from the Gaussian measure γ, which is the only centered optimizer. Let $\mathcal{A}$ be a family of centered probability measures and d a metric or a functional that identifies the equality cases. We say that the LSI is *weakly stable* under $(d, \mathcal{A})$ if $\{f_k d\gamma\} \subset \mathcal{A}$ and $\delta(f_k) \to 0$ implies $d(f_k d\gamma, d\gamma) \to 0$ as $k \to \infty$. The LSI is *stable* if a modulus of continuity is explicit: for a function $\Phi : [0, \infty) \to [0, \infty)$ such that $\Phi(t) = 0$ if and only if $t = 0$,

$$\delta(f) \geq \Phi(d(f d\gamma, d\gamma))$$

for all $f d\gamma \in \mathcal{A}$. Let $\mathcal{P}_2(\mathbb{R}^n)$ be the class of probability measures with finite second moments, and $\mathcal{P}_2^M(\mathbb{R}^n)$ the class of probability measures whose second moments are bounded by $M > 0$.

3.2 Statements of stability results

3.2.1 Optimal transport method

We present stability estimates obtained by the optimal transport technique (Theorem 2.2.4 and Remark 2.2.5).

Theorem 3.2.1. *Let $f d\gamma$ be a centered probability measure in $\mathcal{P}_2^M(\mathbb{R})$. Then there exists $C = C(M) > 0$ such that*

$$\delta(f) \geq C \|f - 1\|^4_{L^1(d\gamma)}. \tag{3.2.1}$$

In the next chapter, we will see that (3.2.1) is false in L^p if $p > 1$ (Theorem 4.1.1) : there exists a sequence of centered probability measures $f_k d\gamma \in \mathcal{P}_2^M(\mathbb{R})$ (also on $\mathbb{R}^n$) for which $\delta(f_k) \to 0$ and

$$\liminf_{k \to \infty} \|f_k - 1\|_{L^p(d\gamma)} > 0.$$

A sufficient additional condition for L^p–stability is higher integrability.

Corollary 3.2.2. *Let $fd\gamma$ be a centered probability measure in $\mathcal{P}_2^M(\mathbb{R})$ such that $\int |f|^{2p-1}d\gamma \leq N$ for some $p > 1$ and $N > 0$. Then there exists $C = C(M, N, p) > 0$ such that*

$$\delta(f) \geq C\|f - 1\|^{8p}_{L^p(d\gamma)}.$$

We extend Theorem 3.2.1 to higher dimension, under some tensorization assumptions.

Corollary 3.2.3. *Let $fd\gamma$ be a probability measure such that $\int x_i f d\gamma = 0$ and*

$$\int |x_i|^2 f d\gamma(x_i) \leq M \quad a.e.\ x' = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

for some $M > 0$ and some $i = 1, 2, \cdots, n$. Then there exists $C = C(M) > 0$ such that

$$\delta(f) \geq C\|f - 1\|^4_{L^1(d\gamma)}.$$

Remark 3.2.4. Consider a class of probability measures $fd\gamma$ such that $\|f\|_{L^\infty(\mathbb{R}^n)} < R$ for some $R > 1$. If $f(x_1, \ldots, x_n) = \Pi_{i=1}^n f_i(x_i)$ and f_j is centered for some j, then f satisfies the above condition. Therefore, the constant C is independent of the dimension for this function space.

Corollary 3.2.5. *Suppose $f(x_1, \ldots, x_n) = \Pi_{i=1}^n f_i(x_i)$, where $f_i \in \mathcal{P}_2^M(\mathbb{R})$ and $fd\gamma$ is a centered probability measure. Then there exists $C = C(n, M) > 0$ such that*

$$\delta(f) \geq C\|f - 1\|^4_{L^1(d\gamma)}.$$

To prove Theorem 3.2.1, we apply the optimal transport technique and deduce that the total variation distance is bounded above by W_1 and the LSI deficit. Then we employ the following W_1–stability result.

Theorem 3.2.6. *Let $fd\gamma$ be a centered probability measure in $\mathcal{P}_2^M(\mathbb{R}^n)$. There exists a constant $C = C(n, M) > 0$ such that*

$$\delta(f) \geq C \min\{W_1(fd\gamma, d\gamma), W_1^4(fd\gamma, d\gamma)\}.$$

Remark 3.2.7. The proof is based on the stability estimates of Talagrand's transportation inequality in terms of W_1 by [18, 51]. In the same way, one can obtain $W_{1,1}$-stability from [57, Theorem 5].

Remark 3.2.8. Since W_1–stability is not true in $\mathcal{P}_2(\mathbb{R}^n)$, the constant $C = C(n, M)$ in Theorem 3.2.6 cannot be taken independent of M (see Theorem 4.1.2). Furthermore, the constant C necessarily depends on the dimension for the following reason: there exists a sequence of centered probability measures $\{f_k d\gamma\}$ in $\mathcal{P}_2^M(\mathbb{R})$ such that $\delta(f_k) \to 0$ by Example 4.3.2. Then Theorem 3.2.16 implies

$$n \leq n + \limsup_{k\to\infty} \mathrm{I}(f_k) \leq M.$$

The proof of Theorem 3.2.6 is based on the observation that the relative entropy is bounded by the deficit and the second moment via the HWI inequality. Then we combine this with a stability estimate for Talagrand's transportation inequality [51, 57].

The following theorem does not impose additional regularity assumptions or bounds on the second moment and yields L^1-stability in case that there is an L^1 bound on the densities. For $g \in L^1(d\gamma)$ and $\alpha > 0$, we define

$$\mathcal{B}(\alpha) = \{f d\gamma \in \mathcal{P} : f(x) \geq \alpha \text{ a.e. } x\}, \tag{3.2.2}$$

$$\mathcal{B}(\alpha, g) = \{f d\gamma \in \mathcal{P} : \alpha \leq f(x) \leq g(x) \text{ a.e. } x\} \tag{3.2.3}$$

where $\mathcal{P}$ is the space of probability measures.

Theorem 3.2.9. *Let $\alpha \in (0, 1]$ and $f d\gamma \in \mathcal{B}(\alpha)$ be a centered probability measure. Then there exists $C(\alpha, n) > 0$ and a linear function $L_f = a_f \cdot x + b_f$ such that*

$$\delta(f) \geq C(\alpha, n) \| \log f - L_f \|^2_{L^1(d\gamma)}, \tag{3.2.4}$$

where $a_f \in \mathbb{R}^n$, $b_f \in \mathbb{R}$, and $|a_f| + |b_f| \leq c$ for some $c = c(n, \alpha) > 0$.

Corollary 3.2.10. *Let $\alpha \in (0, 1]$ and $\{f_k d\gamma\} \subset \mathcal{B}(\alpha)$ be a sequence of centered probability measures such that $\delta(f_k) \to 0$ as $k \to \infty$. Then there exist a subsequence $\{f_{k_j}\} \subset \{f_k\}$ and a constant $c \in [\alpha, 1]$ such that $f_{k_j} \to c$ a.e. as $j \to \infty$.*

Corollary 3.2.11. *Let $\alpha \in (0,1]$, $g \in L^1(d\gamma)$, and $\{f_k d\gamma\} \subset \mathcal{B}(\alpha, g)$ be a sequence of centered probability measures. If $\delta(f_k) \to 0$ as $k \to \infty$, then $f_k \to 1$ in $L^1(d\gamma)$.*

Remark 3.2.12. For any $M, \alpha > 0$ and $g \in L^1(d\gamma)$, we have $\mathcal{B}(\alpha, g) \not\subset \mathcal{P}_2^M(\mathbb{R}^n)$ and $\mathcal{P}_2^M(\mathbb{R}^n) \not\subset \mathcal{B}(\alpha, g)$. To see this, it suffices to consider the case $n = 1$. Let $M > 0$ be fixed and $f_k d\gamma$ be a sequence of probability measures constructed as in Example 4.3.2 with $w = 2$. Then we can choose v so that $\{f_k d\gamma\}$ is included in $\mathcal{P}_2^M$ as we have seen in the end of Section 4.3. Since the minimum of f_k converges to 0, we get $\mathcal{P}_2^M \not\subset \mathcal{B}(\alpha, g)$. We define a sequence of functions f_k such that $f_k(x) = f_k(-x)$ and

$$f_k(x) = \begin{cases} \dfrac{e^{\frac{x^2}{2}}}{C_k \pi (x^2+1)}, & x \in [0,k], \\ \dfrac{e^{\frac{k^2}{2}}}{C_k \pi (k^2+1)}, & x \in (k, \infty) \end{cases}$$

where

$$C_k = \frac{2}{\pi}\Big(\arctan(k) + \frac{e^{\frac{k^2}{2}}(1 - \Phi(k))}{k^2+1} \Big).$$

Note that $f_k d\gamma$ is a probability measure and $C_k \to 1$ as $k \to \infty$. Furthermore, there exist $C, \alpha > 0$ such that $f_k \geq \alpha$ for all k and

$$f_k(x) \leq \frac{C e^{\frac{x^2}{2}}}{\pi(x^2+1)} \in L^1(d\gamma)$$

for all x and k. Since the second moment of $f_k d\gamma$ diverges, we conclude that $\mathcal{B}(\alpha, g) \not\subset \mathcal{P}_2^M(\mathbb{R}^n)$.

Combining Theorem 3.2.6 with the standard compactness argument, we obtain weak L^1-stability in $\mathcal{P}_2^M(\mathbb{R}^n)$.

Theorem 3.2.13. *Let $M \geq n$ and $\{f_k d\gamma\}$ be a sequence of centered probability measures in $\mathcal{P}_2^M(\mathbb{R}^n)$. If $\delta(f_k) \to 0$ as $k \to \infty$, then $f_k \to 1$ in $L^1(d\gamma)$.*

3.2.2 Fourier analytic method

Let $g(x) = 2^{\frac{n}{4}} e^{-\pi|x|^2}$ and $dm = g^2(x)dx$. The LSI deficit with respect to dm is defined by

$$\delta_c(f) = \frac{1}{2\pi} \int_{\mathbb{R}^n} |\nabla f|^2 \, dm - \int_{\mathbb{R}^n} |f|^2 \log |f|^2 \, dm$$

for a normalized function $f \in L^2(dm)$. Note that if $u_f(x) = f(2\sqrt{\pi}x)^{\frac{1}{2}}$, then we have $\delta_c(u_f) = \delta(f)$ by change of variable. Since

$$\|u_f - 1\|^2_{L^2(dm)} \leq \|f - 1\|_{L^2(d\gamma)} \leq 2\|u_f - 1\|_{L^2(dm)},$$

L^2–stability with respect to $\delta_c(f)$ is equivalent to L^1–stability with respect to $\delta(f)$.

Recall that Carlen [40] derived the lower bound of $\delta_c(f)$ in terms of the relative entropy of the Fourier–Wiener transform (see Theorem 2.2.1). We investigate the case where $\mathrm{Ent}_{dm}(|\mathcal{W}f|^2)$ converges to 0 and use a compactness argument to obtain the following weak L^2–stability result.

Theorem 3.2.14. *Let $M > 0$, $\epsilon \in (0, 2\pi)$, and $\{f_k\}$ be a sequence of normalized and centered functions in $L^2(dm)$. Suppose*

$$\int |f_k|^2 e^{-(2\pi-\epsilon)|x|^2} dx \leq M$$

for all k. If $\delta_c(f_k) \to 0$ as $k \to \infty$, $f_k \to 1$ in $L^2(dm)$.

Remark 3.2.15. As we have seen in Corollary 3.2.2, higher integrability assumption yields weak L^p-stability for $p > 2$.

The optimal transport method and Carlen's deficit estimate (2.2.4) yields the following inequality which in particular implies weak L^2–stability for $\mathcal{P}_2^M(\mathbb{R}^n)$ with respect to δ_c.

Theorem 3.2.16. *Let f be normalized in $L^2(dm)$. Then*

$$2\sqrt{\pi n} W_2(dm, |\mathcal{W}f|^2 dm) + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2) \geq 2\pi \int |x|^2 dm - 2\pi \int |x|^2 |f|^2 dm + \frac{1}{2\pi} \int |\nabla f|^2 dm. \quad (3.2.5)$$

As a consequence,

$$\sqrt{2n}\delta_c^{\frac{1}{2}}(f) + \delta_c(f) \geq 2\pi \int |x|^2 dm - 2\pi \int |x|^2 |f|^2 dm + \frac{1}{2\pi} \int |\nabla f|^2 dm. \quad (3.2.6)$$

Remark 3.2.17. Stability for Talagrand's transportation inequality (2.1.4) yields an extra remainder term while passing from (3.2.5) to (3.2.6).

Corollary 3.2.18. *Let $M \geq 1$ and $\{f_k\}$ be a sequence of normalized and centered functions in $L^2(dm)$. Suppose $\int |f_k|^4 dm \leq M$ for all k. If $\delta_c(f_k) \to 0$ as $k \to \infty$, then*

$$\lim_{k\to\infty} \int |\nabla f_k|^2 dm = 0.$$

In particular, we have $f_k \to 1$ in $L^2(dm)$.

Corollary 3.2.19. *Let $M \geq \frac{n}{4\pi}$ and $\{f_k\}$ be a sequence of normalized and centered functions in $L^2(dm)$. Suppose $\int |x|^2 |f_k|^2 dm \leq M$ for all k. If $\delta_c(f_k) \to 0$ as $k \to \infty$, then $f_k \to 1$ in $L^2(dm)$.*

Remark 3.2.20. Suppose there is a modulus of continuity ω and $C = C(M) > 0$ such that $\|f - 1\|_{L^p(d\gamma)} \leq C\omega(\delta(f))$ as $\delta(f) \to 0$ and $f \in \mathcal{P}_k^M(\mathbb{R}^n)$ for $k \in \mathbb{N}$ and $p \geq 1$. Then C necessarily depends on the dimension since $\int |x|^k d\gamma \leq M$.

3.2.3 Probabilistic method

Another approach to proving stability estimates for the LSI is to investigate quantitative versions of Cramér's theorem [24, 25] and combine them with a convolution type deficit estimate of the LSI in [58]. We consider the space of probability measures in $\mathcal{P}_2^M$ satisfying further integrability and assumptions on the second moment. For probability measures μ and ν on $\mathbb{R}$, the Kolmogorov distance is given by

$$d_{\mathrm{K}}(\mu, \nu) = \sup_{x\in\mathbb{R}} |\mu((-\infty, x]) - \nu((-\infty, x])|.$$

Theorem 3.2.21. *Let f be a symmetric nonnegative function on $\mathbb{R}$ and $d\mu = f d\gamma \in \mathcal{P}_2^M(\mathbb{R})$ with $m_2(\mu) = k$. Let $v(x) = f(\frac{x}{\sqrt{2}})^2$ and assume that $d\nu := v d\gamma$ is a probability measure. Then there exists $\varepsilon_0 > 0$ such that if $\delta(v) \leq \varepsilon \leq \varepsilon_0$, then*

$$d_{\mathrm{K}}(\mu, \gamma_\varepsilon) \leq \frac{C_k}{\sqrt{\log \frac{1}{\varepsilon}}} \tag{3.2.7}$$

where C_k depends on k and γ_ε is a Gaussian measure given by

$$d\gamma_\varepsilon = \frac{1}{\sqrt{4\pi\sigma_\varepsilon^2}} e^{-\frac{|x|^2}{4\sigma_\varepsilon^2}} dx,$$

for some $\sigma_\varepsilon^2 > 0$ depending on ε.

Theorem 3.2.22. *Let f be a symmetric nonnegative function on $\mathbb{R}$, $d\mu = f d\gamma$, and $m_2(\mu) = 1$. Let $v(x) = f(\frac{x}{\sqrt{2}})^2$ and assume that $d\nu := v d\gamma$ is a probability measure. We have*

$$\delta(v) \geq \Psi(d_{\mathrm{K}}(\mu, \gamma))$$

where $\Psi(t) = e^{-\frac{c}{t^2}}$ for some $c > 0$.

Remark 3.2.23. Note that for $d\gamma_{\frac{1}{\sqrt{2}}} = \frac{1}{\sqrt{\pi}} e^{-|x|^2} dx$ and $v(x) = f(\frac{x}{\sqrt{2}})^2$,

$$\delta(v) = \int |\nabla f|^2 d\gamma_{\frac{1}{\sqrt{2}}} - \int |f|^2 \log |f|^2 d\gamma_{\frac{1}{\sqrt{2}}}.$$

Remark 3.2.24. By Proposition 2.1.6, Theorem 3.2.6 implies that

$$\delta(f) \geq C_M \min\{d_{\mathrm{K}}(\mu, \gamma)^2, d_{\mathrm{K}}(\mu, \gamma)^8\}.$$

On the other hand, it follows from Proposition 2.1.8 that Theorem 3.2.22 implies

$$\delta(v) \geq c_1 \Psi(c_2 W_1(\mu, \gamma)^2).$$

Note that if t is small then $\Psi(t)$ is bounded by t^8, which implies that Theorem 3.2.6 is stronger than Theorem 3.2.22. Notice also that Theorem 3.2.22 has a scaled version of the deficit $\delta(v)$.

3.3 Dimension-free stability estimates

One of the most important features of the logarithmic Sobolev inequality is that the sharp constant $\frac{1}{2}$ is dimension-independent, which leads to many interesting applications. It is natural to ask if there is a dimension-free quantitative improvement of the LSI.

We observe that Carlen's deficit estimate (2.2.1) is dimension-free. He showed that the LSI deficit is bounded below by the relative entropy of the Fourier–Wiener transform, which yields the characterization of the equality cases. This estimate is, however, not metric-involved. The first result on dimension-free stability estimates

in terms of a metric is found in [83], where W_2–stability was considered in the space of probability measures satisfying the differential inequalities

$$-1+\varepsilon \leq D^2(-\log f) \leq M$$

for $\varepsilon, M > 0$. The estimate only depends on the choice of ε and M. In [57], the authors considered the space of probability measures satisfying Poincaré inequalities. Indeed, they proved a strict improvement of the LSI in within the class of probability measures satisfying a $(2,2)$-Poincaré inequality with a constant $\lambda > 0$

$$\lambda \int g^2 \, f d\gamma \leq \int |\nabla g|^2 \, f d\gamma.$$

for every smooth function g with $\int g \, f d\gamma = 0$. The improvement yields stability estimates in terms of W_2 and L^1, which depend only on the Poincaré constant λ. In [58], it was shown that if $f d\gamma$ is a probability measure satisfying

$$\mathcal{F}(e^{-\pi|x|^2} f(2\sqrt{\pi}x)) \geq 0,$$

then $\delta(f) \geq \frac{1}{2}\|f-1\|_2^4$, which is dimension-free. We note that this estimate does not have any parameters while the above estimates have the parameters that define the probability measure spaces.

The stability estimate in Corollory 3.2.3 is dimension-free. We show L^1–stability in the space of probability measures such that for some $M > 0$ and some $i = 1, 2, \cdots, n$, $\int x_i f d\gamma = 0$ and

$$\int |x_i|^2 f d\gamma(x_i) \leq M \quad \text{a.e. } x' = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

Note that the constant depends only on M. In particular, as we have seen in Remark 3.2.4, we have a dimension-free L^1–stability estimate in the space

$$\{f d\gamma \in \mathcal{P} : f(x_1, \cdots, x_n) = \prod_{i=1}^{n} f_i(x_i), \|f\|_\infty < R\}$$

for $R > 0$.

Recently, it was shown in [55] that the LSI can be self-improved with a dimension-free estimate. Previously, a self-improvement of the LSI with a dimensional constant

was proven in [26, 53]. The authors in [55] derived a dimension independent estimate in terms of the Fisher information matrix. As a consequence, they proved that if the covariance matrix of a measure is dominated by the identity matrix, then the deficit of the LSI is bounded below by some functionals in term of the eigenvalues of the Fisher information matrix.

However, the logarithmic Sobolev inequality turns out to be so delicate that such dimension-free stability estimates require strong restrictions on probability measures and distance functionals. As we have seen in Remark 3.2.8, any W_1–stability estimates in $\mathcal{P}_2^M$ should depend on M by Example 4.3.2 and the constant M necessarily depends on n by Theorem 3.2.16. In Theorem 4.1.1, we show that W_2–stability fails in $\mathcal{P}_2^M(\mathbb{R}^n)$ for $M > n$, which implies that dimension-dependency is necessary for W_2–stability. We remark that it was shown in [55] that there exist a sequence of dimensions $n_k \in \mathbb{N}$ and a sequence of probability measures μ_k on $\mathbb{R}^{n_k}$ such that the deficit converges to 0 but the W_2 distance from the class of optimizers diverges.

3.4 Proofs of the main results

3.4.1 Proofs of Theorem 3.2.1 and its corollaries

Proof of Theorem 3.2.1. Let T be the Brenier map between $f d\gamma$ and $d\gamma$. Recall that the Gaussian measure $d\gamma$ satisfies the $(1,1)$-Poincaré inequality

$$\int |f-1| d\gamma \le 2 \int |\nabla f| \, d\gamma = 2 \int |\nabla (\log f)| \, f d\gamma.$$

Combining this with (2.2.7), we have

$$\begin{aligned}
\int |f-1| \, d\gamma &\le 2 \int |\nabla \log f| \, f d\gamma \\
&\le 2 \int |\nabla \log f - T(x) + x| \, f d\gamma + 2 \int |T(x) - x| \, f d\gamma \\
&\le 2 \Big(\int |\nabla \log f - T(x) + x|^2 f d\gamma \Big)^{\frac{1}{2}} + 2 \int |T(x) - x| f d\gamma \\
&\le 2\sqrt{2} \delta^{\frac{1}{2}}(f) + 2 \int |T(x) - x| f d\gamma.
\end{aligned}$$

Note that in one-dimensional case, the Brenier map between $fd\gamma$ and $d\gamma$ gives an optimal transport plan for W_1 as well as W_2. In other words, we have

$$\int |T(x) - x| f d\gamma = W_1(fd\gamma, d\gamma).$$

Applying Theorem 3.2.6, we get

$$\int |f - 1| d\gamma \leq 2\sqrt{2}\delta^{\frac{1}{2}}(f) + C \max(\delta(f), \delta^{\frac{1}{4}}(f))$$

as desired. □

Proof of Corollary 3.2.2. The result follows from Cauchy–Schwarz inequality

$$\begin{aligned}\int |f-1|^p d\gamma &= \int |f-1|^{p-\frac{1}{2}} |f-1|^{\frac{1}{2}} d\gamma \\ &\leq (2^{2p-2}(N+1))^{\frac{1}{2}} \|f-1\|_{L^1(d\gamma)}^{\frac{1}{2}} \\ &\leq C_M (2^{2p-2}(N+1))^{\frac{1}{2}} \delta^{\frac{1}{8}}(f).\end{aligned}$$

□

Proof of Corollary 3.2.3. For fixed $x' = (x_1, \ldots, x_{i-1}, x_{i+1}, \cdots, x_n)$, let $g_{x'}(x_i) = f(x')$. Theorem 3.2.1 implies

$$\frac{1}{2}\int \frac{(\partial_{x_i} g_{x'}(x_i))^2}{g_{x'}(x_i)} d\gamma(x_i) \geq \int g_{x'}(x_i) \log g_{x'}(x_i) d\gamma(x_i) + c\Big(\int |g_{x'}(x_i) - 1| d\gamma(x_i)\Big)^4.$$

Since we have

$$\int\int \frac{|\nabla f|^2}{f} d\gamma(x_i) d\gamma(x') \geq \int\int \frac{(\partial_{x_i} g_{x'}(x_i))^2}{g_{x'}(x_i)} d\gamma(x_i) d\gamma(x'),$$

it follows from Jensen's inequality that

$$\begin{aligned}\delta(f) &\geq C \int \Big(\int |g_{x'}(x_i) - 1| d\gamma(x_i)\Big)^4 d\gamma(x') \\ &\geq C\Big(\int\int |g_{x'}(x_i) - 1| d\gamma(x_i) d\gamma(x')\Big)^4 \\ &= C\|f-1\|_{L^1(d\gamma)}^4.\end{aligned}$$

□

Proof of Corollary 3.2.5. By applying Theorem 3.2.1 to f_i, it follows that

$$\sum_{i=1}^{n} \delta(f_i) \geq c \sum_{i=1}^{n} \Big(\int |f_i(x_i) - 1| d\gamma(x_i) \Big)^4.$$

Since the Fisher information and the relative entropy of $f = f_1 f_2 \cdots f_n$ are

$$\begin{aligned} \int \frac{|\nabla f|^2}{f} \, d\gamma &= \int \frac{|\nabla (f_1 f_2 \cdots f_n)|^2}{f} \, d\gamma \\ &= \sum_{i=1}^{n} \int \frac{(\partial_{x_i} f_i(x_i))^2 \Pi_{j \neq i} (f_j(x_j))^2}{f} \, d\gamma \\ &= \sum_{i=1}^{n} \int \frac{(\partial_{x_i} f_i)^2}{f_i} \, d\gamma(x_i) \end{aligned}$$

and

$$\int f \log f \, d\gamma = \sum_{i=1}^{n} \int f_i(x_i) \log f_i(x_i) \, d\gamma(x_i),$$

we have $\delta(f) = \sum_{i=1}^{n} \delta(f_i)$. The result follows from

$$\begin{aligned} \sum_{i=1}^{n} \Big(\int |f_i(x_i) - 1| \, d\gamma(x_i) \Big)^4 &\geq n^{-3} \Big(\sum_{i=1}^{n} \int |f_i(x_i) - 1| \, d\gamma(x_i) \Big)^4 \\ &\geq n^{-3} \Big(\int |f - 1| \, d\gamma \Big)^4. \end{aligned}$$

□

3.4.2 Proofs of Theorem 3.2.6

By the HWI inequality (2.1.5), we obtain

$$\begin{aligned} \mathrm{H}(f) &\leq W_2(f d\gamma, d\gamma) \sqrt{\mathrm{I}(f)} - \frac{1}{2} W_2^2(f d\gamma, d\gamma) \\ &\leq \frac{1}{2t} \mathrm{I}(f) + \frac{t-1}{2} W_2^2(f d\gamma, d\gamma) \end{aligned}$$

for any $t > 1$. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Brenier map from $d\gamma$ to $f d\gamma$, then

$$\begin{aligned} \mathrm{H}(f) &\leq \frac{t}{2} \int_{\mathbb{R}^n} |T(x) - x|^2 d\gamma + \frac{t}{t-1} \delta(f) \\ &\leq t \Big(\int_{\mathbb{R}^n} |x|^2 f d\gamma + \int_{\mathbb{R}^n} |x|^2 d\gamma \Big) + \frac{t}{t-1} \delta(f) \\ &\leq t(n + M) + \frac{t}{t-1} \delta(f). \end{aligned}$$

Note that it is well-known (for example [57]) that $16\delta(f)\mathrm{H}(f) \geq \delta^2_{\mathrm{Tal}}(f)$ where $\delta_{\mathrm{Tal}}(f)$ is defined in (2.1.3). Thus we obtain

$$\delta^2(f) + (t-1)(n+M)\delta(f) - \frac{(t-1)}{16t}\delta^2_{\mathrm{Tal}}(f) \geq 0.$$

Solving this inequality for $\delta(f)$ and applying (2.1.4), we get

$$\begin{aligned}
\delta(f) &\geq \frac{t-1}{2}(n+M)\Big(\big(1 + \frac{\delta^2_{\mathrm{Tal}}(f)}{4t(t-1)(n+M)}\big)^{\frac{1}{2}} - 1\Big) \\
&\geq \frac{t-1}{6}(n+M)G\left(\frac{\delta_{\mathrm{Tal}}(f)}{2\sqrt{t(t-1)(n+M)}}\right) \\
&\geq \frac{t-1}{6}(n+M)G\left(\frac{C_{\mathrm{CE}}}{2\sqrt{t(t-1)(n+M)}}G(W_1(fd\gamma, d\gamma))\right)
\end{aligned}$$

where $G(x) = \min\{x, x^2\}$. We finish the proof by choosing $t > 1$ such that $C_{\mathrm{CE}} = 2\sqrt{t(t-1)(n+M)}$. □

3.4.3 Proofs of Theorem 3.2.9 and its corollaries

Proof of Theorem 3.2.9. Let $T = \nabla\Phi = (T^1, T^2, \ldots, T^n)$ be the Brenier map from $fd\gamma$ to $d\gamma$ and $\{\lambda_i\}$ the eigenvalues of $DT - Id$. By (2.2.7), we have

$$\delta(f) \geq \frac{1}{2}\int |T(x) - x + \nabla \log f|^2 \, fd\gamma + \int \sum_{i=1}^n (\lambda_i - \log(1+\lambda_i))\, fd\gamma.$$

Since $f(x) \geq \alpha$ for all x, there exists a constant C such that $|\log f| \leq C(|f|+1)$. This implies that $\|\log f - L_f\|_{L^1(d\gamma)}$ is bounded so that it suffices to assume that $\delta(f) \leq 1$. Since $t - \log(1+t) \geq (1-\log 2)\min\{t, t^2\}$ for $t \geq 0$ and $f(x) \geq \alpha$ for all x, we have

$$\begin{aligned}
\int \sum_{i=1}^n (\lambda_i - \log(1+\lambda_i))\, fd\gamma &\geq C_\alpha \sum_{i=1}^n \Big(\int_{\{|\lambda_i|\leq 1\}} |\lambda_i|^2 d\gamma + \int_{\{|\lambda_i|>1\}} |\lambda_i| d\gamma\Big) \\
&\geq C_\alpha \sum_{i=1}^n \Big(\Big(\int_{\{|\lambda_i|\leq 1\}} |\lambda_i| d\gamma\Big)^2 + \int_{\{|\lambda_i|>1\}} |\lambda_i| d\gamma\Big) \\
&\geq C_\alpha \sum_{i=1}^n \Big(\Big(\int_{\{|\lambda_i|\leq 1\}} |\lambda_i| d\gamma\Big)^2 + \Big(\int_{\{|\lambda_i|>1\}} |\lambda_i| d\gamma\Big)^2\Big) \\
&\geq C_{\alpha,n}\Big(\sum_{i=1}^n \int |\lambda_i| d\gamma\Big)^2
\end{aligned}$$

where $C_{\alpha,n} = (1 - \log 2)\alpha/n$. Let $a = \int T d\gamma$ and $a = (a_1, \dots, a_n)$, then

$$|a| \le \int |T| d\gamma \le \frac{1}{\alpha} \int |x| d\gamma.$$

By the $(1,1)$-Poincaré inequality for $d\gamma$, we have

$$\begin{aligned}
\sum_{i=1}^{n} \int |\lambda_i| \, d\gamma &\ge C \sum_{i,j} \int |\partial_j T^i - \delta_{ij}| \, d\gamma \\
&\ge C \sum_{i} \int |\nabla (T^i - x_i)| \, d\gamma \\
&\ge C \sum_{i} \int |T^i - x_i - a_i| \, d\gamma
\end{aligned}$$

for some universal constant C. Thus we have

$$\begin{aligned}
\int |\nabla \log f + a| \, d\gamma &\le \int |\nabla \log f + (T - x)| \, d\gamma + \int |T - x - a| \, d\gamma \\
&\le \frac{1}{\sqrt{\alpha}} \Big(\int |\nabla \log f + (T - x)|^2 \, f d\gamma \Big)^{\frac{1}{2}} + \int |T - x - a| \, d\gamma \\
&\le C(\alpha, n) \delta^{\frac{1}{2}}(f).
\end{aligned}$$

Let $b = \int \log f d\gamma \in (\log \alpha, 0)$. It then follows from the $(1,1)$-Poincaré inequality for $d\gamma$ that

$$\int |\nabla \log f + a| \, d\gamma \ge c \int |\log f + a \cdot x - b| \, d\gamma,$$

which finishes the proof. □

Proof of Corollary 3.2.10. Theorem 3.2.9 implies that

$$\lim_{k \to \infty} \int |\log \left(f_k e^{-(a_{f_k} \cdot x + b_{f_k})} \right)| \, d\gamma = 0$$

Since $|a_{f_k}| + |b_{f_k}|$ is uniformly bounded in k, there exists $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that $a_{f_k} \to a$ and $b_{f_k} \to b$ as $k \to \infty$ along a subsequence. There exists a further subsequence such that $\log(f_k e^{-(a_{f_k} \cdot x + b_{f_k})}) \to 0$ and in turn that $f_k \to e^{a \cdot x + b}$ a.e. as $k \to \infty$. Since $f_k(x) \ge \alpha$ for all k and x, we have $a = 0$ as desired. □

Proof of Corollary 3.2.11. Suppose that there exists a subsequence $\{f_k\}$ such that $\|f_k - 1\|_{L^1(d\gamma)} \ge C > 0$ for all k. By Corollary 3.2.10, f_k converges to a constant

a.e. along a subsequence as $k \to \infty$. It then follows from the dominated convergence theorem and the mass constraint that f_k converges to 1 a.e., which is a contradiction. □

3.4.4 Proof of Theorem 3.2.13

Let $\{f_j\}$ be any subsequence of the original sequence. We will show that there exists a further subsequence $\{f_{j(k)}\}$ that converges to 1 in $L^1(\mathbb{R}^n, d\gamma)$. By (2.1.6), it suffices to show that $\sqrt{f_{j(k)}} \to 1$ in $L^2(\mathbb{R}^n, d\gamma)$. Since the deficit converges to zero, it follows from Theorem 3.2.16 that $\{\mathrm{I}(f_j)\}_{j\geq 1}$ is uniformly bounded in j. Let $h_j = f_j\gamma$ where $\gamma(x) = (2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}}$, then

$$\begin{aligned}
\mathrm{I}(f_j) &= 4\int_{\mathbb{R}^n} |\nabla(\sqrt{f_j})|^2 d\gamma \\
&= 4\int_{\mathbb{R}^n} |\nabla(\sqrt{h_j}) - \sqrt{f_j}\nabla(\sqrt{\gamma})|^2 dx \\
&= 4\int_{\mathbb{R}^n} |\nabla(\sqrt{h_j})|^2 dx - 2n + \int_{\mathbb{R}^n} |x|^2 d\nu_j.
\end{aligned}$$

So $\{\sqrt{h_j}\}_{j\geq 1}$ is bounded in $W^{1,2}(\mathbb{R}^n)$.

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. The Rellich–Kondrashov theorem says that there exists a subsequence $\{h_{j(k)}\}_{k\geq 1}$ such that $\sqrt{h_{j(k)}}$ converges to a function g in $L^2(\Omega)$. Since $\sqrt{h_j}$ is nonnegative for all j, we let $g = \sqrt{f\gamma}$.

We claim that $f = 1$ a.e. in Ω. Let $d\nu_j = f_j d\gamma$. Since $\delta(f_{j(k)})$ converges to 0, we have $W_1(\nu_{j(k)}, \gamma) \to 0$ by Theorem 3.2.6. This implies that $\nu_{j(k)} \rightharpoonup \gamma$ weakly, that is,

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi d\nu_{j(k)} = \int_{\mathbb{R}^n} \varphi d\gamma$$

for all $\varphi \in C_b^0(\mathbb{R}^n)$. Let $\varepsilon > 0$ and φ be a bounded continuous function such that $|\varphi| \leq K$ for some $K > 0$. We pick $N \in \mathbb{N}$ such that

$$\Big|\int_\Omega \varphi d\nu_{j(k)} - \int_\Omega \varphi d\gamma\Big| \leq \frac{\varepsilon}{2}, \quad \Big|\int_\Omega |\sqrt{f_{j(k)}} - \sqrt{f}|^2 d\gamma\Big| \leq \frac{\varepsilon^2}{16K^2}$$

for any $k \geq N$. Since $\int_\Omega f_{j(k)} d\gamma \leq 1$ for all k and

$$\int_\Omega f d\gamma \leq \int_\Omega |\sqrt{f_{j(k)}} - \sqrt{f}|^2 d\gamma + \int_\Omega f_{j(k)} d\gamma,$$

we obtain $\int_\Omega f d\gamma \leq 1$. One can see that

$$\begin{aligned}\Big|\int_\Omega (f-1)\varphi d\gamma\Big| &\leq \Big|\int_\Omega (f_{j(k)}-1)\varphi d\gamma\Big| + \Big|\int_\Omega (f-f_{j(k)})\varphi d\gamma\Big| \\ &\leq \frac{\varepsilon}{2} + K\int_\Omega |f-f_{j(k)}|d\gamma \\ &\leq \frac{\varepsilon}{2} + K\Big(\int_\Omega |\sqrt{f}-\sqrt{f_{j(k)}}|^2 d\gamma\Big)^{\frac{1}{2}}\Big(\int_\Omega |\sqrt{f}+\sqrt{f_{j(k)}}|^2 d\gamma\Big)^{\frac{1}{2}} \\ &\leq \varepsilon.\end{aligned}$$

This holds for all $\varepsilon > 0$ and all $\varphi \in C_b^0(\Omega)$. Thus, we conclude that $f = 1$ a.e. in Ω.

Let $B_k := \{x \in \mathbb{R}^n : |x| < k\}$ for each $k \in \mathbb{N}$. Choose a subsequence $\{f_{1,j}\}_{j\geq 1}$ such that $\sqrt{f_{1,j}} \to 1$ in $L^2(B_1, d\gamma)$ as $j \to \infty$. On B_2, we can find a further subsequence $\{f_{2,j}\}_{j\geq 1} \subseteq \{f_{1,j}\}_{j\geq 1}$ such that $\sqrt{f_{2,j}} \to 1$ in $L^2(B_2, d\gamma)$ as $j \to \infty$. Iterating this procedure, we have $\{f_{k,j}\}_{j,k\geq 1}$ such that $\sqrt{f_{k,j}} \to 1$ in $L^2(B_k, d\gamma)$ as $j \to \infty$. Define $f^{(k)} := f_{k,k}$ and let $\varepsilon > 0$.

Since ν_j converges weakly to γ, the family $\{\nu_j\}$ is tight by Prokhorov's theorem. Thus, we can choose $N_1 \in \mathbb{N}$ be such that $\int_{\mathbb{R}^n \setminus B_k} d\nu_j < \frac{\varepsilon}{8}$ and $\int_{\mathbb{R}^n \setminus B_k} d\gamma < \frac{\varepsilon}{8}$ for all $k \geq N_1$. By definition, there exists $N_2 \in \mathbb{N}$ such that $\int_{B_k} |\sqrt{f^{(k)}} - 1|^2 d\gamma < \frac{\varepsilon}{2}$ for all $k \geq N_2$. Combining our observation, we have

$$\Big|\int_{\mathbb{R}^n} |\sqrt{f^{(k)}} - 1|^2 d\gamma\Big| \leq \Big|\int_{B_k} |\sqrt{f^{(k)}} - 1|^2 d\gamma\Big| + 2\Big|\int_{\mathbb{R}^n \setminus B_k} (f^{(k)} + 1) d\gamma\Big| < \varepsilon$$

for any $k \geq \max\{N_1, N_2\}$. Therefore, we conclude that $\sqrt{f^{(k)}} \to 1$ in $L^2(\mathbb{R}^n, d\gamma)$ as desired. □

3.4.5 Proof of Theorem 3.2.14

Suppose that there exists a subsequence $\{f_k\}$ such that $\|f_k - 1\|_{L^2(dm)} \geq C > 0$ for all k. Let $dm_\epsilon = e^{-(2\pi-\epsilon)|x|^2} dx$. Since

$$\int |f_k|^2 dm_\epsilon \leq M,$$

f_k converges weakly to f in $L^2(dm_\epsilon)$ along a subsequence. Since $\varphi(x) := e^{-2\pi i\xi\cdot x}e^{(\pi-\epsilon)|x|^2}$ is in $L^2(dm_\epsilon)$ for each ξ, we have

$$\lim_{k\to\infty}\int e^{-2\pi i\xi\cdot x}U^*f_k(x)dx = \int e^{-2\pi i\xi\cdot x}U^*f(x)dx.$$

Therefore, $\mathcal{W}f_k(\xi) \to \mathcal{W}f(\xi)$ for every $\xi \in \mathbb{R}^n$. On the other hand, it follows from $\delta_c(f_k) \to 0$ and Carlen's deficit estimate (2.2.4) that

$$\lim_{k\to\infty} \mathrm{Ent}_{dm}(|\mathcal{W}f_k|^2) = 0.$$

By Pinsker's inequality (2.1.2), we have $|\mathcal{W}f_k|^2 \to 1$ in $L^1(dm)$ as $k \to \infty$ and $|\mathcal{W}f|^2 = 1$ a.e. This implies $f = 1$ by Cramér's theorem. Since f_k is normalized in $L^2(dm)$, therefore $f_k \rightharpoonup g$ weakly in $L^2(dm)$ along a further subsequence. By uniqueness of weak limits, we have $g = 1$. This yields $f_k \to 1$ in $L^2(dm)$, which is a contradiction. □

3.4.6 Proof of Theorem 3.2.16

Proof of Theorem 3.2.16. Let $f \in L^2(dm)$ be normalized and T the Brenier map between $dm = |U^*|^2dx$ and $|\mathcal{W}f|^2dm = |\widehat{U^*f}|^2dx$. Note that there exists a convex function ϕ such that $T = \nabla\phi$ and it satisfies the Monge–Ampère equation

$$\begin{aligned}\log\det D^2\phi &= \log\frac{|U^*|^2}{|\widehat{U^*f}(T)|^2}\\ &= \frac{n}{2}\log 2 - 2\pi|x|^2 - \log|\widehat{U^*f}(T)|^2.\end{aligned}$$

Integrating of both sides with respect to $dm = |U^*|^2dx$, we get

$$\int |U^*|^2\log|\widehat{U^*f}(T)|^2dx + \int |U^*|^2\log\det D^2\phi dx = \frac{n}{2}\log 2 - 2\pi\int |x|^2|U^*|^2dx.$$

Let $\psi(x) = \phi(x) - \frac{1}{2}|x|^2$ and λ_i be the eigenvalues of $D^2\psi$, then

$$\begin{aligned}
2\pi \int |x|^2 |\widehat{U^* f}|^2 dx - 2\pi \int |x|^2 \, dm &= \int \log\det D^2\phi \, dm + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2) \\
&= \int \sum_{i=1}^{n} \log(1+\lambda_i) \, dm + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2) \\
&\leq \int \Delta\psi \, dm + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2) \\
&= 4\pi \int (T(x) - x) \cdot x \, dm + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2) \\
&\leq 4\pi W_2(dm, |\mathcal{W}f|^2 dm)(m_2(dm))^{\frac{1}{2}} + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2)
\end{aligned}$$

where $m_2(dm)$ is the second moment of dm. By the Plancherel theorem, we have

$$\begin{aligned}
4\pi^2 \int |x|^2 |\widehat{U^* f}|^2 dx &= \int |\widehat{\nabla U^* f}|^2 dx \\
&= \int |\nabla U^* f|^2 dx \\
&= 4\pi^2 \int |x|^2 |U^* f|^2 dx - 4\pi \int x \cdot \nabla f f dm + \int |\nabla f|^2 dm.
\end{aligned}$$

Using

$$-2 \int x \cdot \nabla f f dm = n - 4\pi \int |x|^2 |f|^2 dm,$$

we obtain

$$\begin{aligned}
2\pi \int |x|^2 dm - 2\pi \int |x|^2 |f|^2 dm + \frac{1}{2\pi} \int |\nabla f|^2 dm & \\
\leq 2\sqrt{\pi n} W_2(dm, |\mathcal{W}f|^2 dm) + \mathrm{Ent}_{dm}(|\mathcal{W}f|^2). &
\end{aligned}$$

Therefore, it follows from Talagrand's transportation inequality (2.1.3) for the measure dm and the entropic uncertainty principle (2.2.3) that

$$\begin{aligned}
2\pi \int |x|^2 dm - 2\pi \int |x|^2 |f|^2 dm + \frac{1}{2\pi} \int |\nabla f|^2 dm &\leq 2\sqrt{\pi n} W_2(dm, |\mathcal{W}f|^2 dm) + \delta_c(f) \\
&\leq \sqrt{2n}\sqrt{\delta_c(f)} + \delta_c(f).
\end{aligned}$$

□

Proof of Corollary 3.2.18. Suppose that there exists a subsequence $\{f_k\}$ such that $\delta_c(f_k) \to 0$ as $k \to \infty$ and $\int |\nabla f_k|^2 dm \geq c > 0$ for all k. Since $\int |f_k|^4 \, dm \leq C$ for all k, along a further subsequence, f_k^2 converges weakly to f^2 in $L^2(dm)$ for some $f \in L^4(dm)$. In particular, we have

$$\int |x|^2 |f_k|^2 dm \to \int |x|^2 |f|^2 dm$$

as $k \to \infty$ because $|x|^2 \in L^2(dm)$. By Theorem 3.2.6 and Theorem 3.2.16, $f = 1$ and

$$\int |\nabla f_k|^2 dm \to 0$$

as $k \to \infty$, which is a contradiction. □

Proof of Corollary 3.2.19. Suppose that there exists a subsequence $\{f_k\}$ such that $\delta_c(f_k) \to 0$ and $\|f_k - 1\|_{L^1(d\gamma)} \geq c > 0$. Since Theorem 3.2.16 implies $\int |\nabla f_k|^2 dm \leq M$ for all k, we have $f_k \to f$ in $L^2(dm)$ as $k \to \infty$ along a further subsequence, for some $f \in L^2(dm)$. Theorem 3.2.6 yields $f = 1$, which is a contradiction. □

3.4.7 Proofs of Theorem 3.2.21 and Theorem 3.2.22

We recall a convolution type deficit estimate for the LSI of [58].

Theorem 3.4.1 ([58, Theorem 4.1]). *Let $f \in L^2(dm)$, $f(x) = f(-x)$, $\|f\|_2 = 1$, and $h = fg$. Then there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}} |h * h - g * g|^2 dx \leq C \delta_c(f)^{\frac{1}{4}} (\|h - g\|_{\frac{6}{5}}^2 + \|h - g\|_2)^{\frac{3}{2}}.$$

The following lemma is an L^1–L^2 estimate under a second moment assumption, which allows connecting stability of Cramér's theorem in [68, 103] (see Theorem 2.3.1 and 2.3.2) with Theorem 3.4.1.

Lemma 3.4.2. *Let u be a nonnegative function in $L^1(dx) \cap L^2(dx)$ such that*

$$\int_{\mathbb{R}} x^2 u(x) dx = k \|u\|_1 < \infty$$

for some $k > 0$. Then we have $\|u\|_1 \leq e^{\frac{k+1}{2}} \|u\|_2$.

Proof. Let $p(x) = u(x)/\|u\|_1$ and $q(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$. Since $\varphi = x \log x$ for $x \geq 0$ is convex ($\varphi(0) = 0$), one can see by Jensen's inequality that

$$\int_{\mathbb{R}} p(x) \log \frac{p(x)}{q(x)} dx = \int_{\mathbb{R}} \varphi(\frac{p(x)}{q(x)}) q(x) dx \geq \varphi(\int_{\mathbb{R}} p(x) dx) = 0.$$

So, we have

$$\begin{aligned} \int_{\mathbb{R}} p(x) \log p(x) dx &\geq \int_{\mathbb{R}} p(x) \log q(x) dx \\ &= -\int_{\mathbb{R}} x^2 p(x) dx - \frac{1}{2} \log \pi \\ &\geq -(k+1). \end{aligned} \tag{3.4.1}$$

Let $1 \leq p_0, p_1 \leq 2$, $\theta \in (0,1)$, and $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. It follows from Hölder's inequality that

$$\|u\|_{p_\theta} \leq \|u\|_{p_0}^{1-\theta} \|u\|_{p_1}^{\theta}. \tag{3.4.2}$$

This implies that the map $p \mapsto J(p) := \log \|f\|_p^p$ is convex on $[1,2]$. On the other hand, the derivative of $J(p)$ is given by

$$\frac{d}{dp} J(p) = \frac{1}{\|u\|_p^p} \int_{\mathbb{R}} |u|^p \log |u| dx.$$

By the convexity of $J(p)$, we have $J(2) - J(1) \geq J'(1)$. So, we apply (3.4.1) to obtain

$$\begin{aligned} \log \|u\|_2^2 - \log \|u\|_1 &\geq \frac{1}{\|u\|_1} \int_{\mathbb{R}} |u| \log |u| dx \\ &= \int_{\mathbb{R}} p(x) \log p(x) dx + \log \|u\|_1 \\ &> -(k+1) + \log \|u\|_1, \end{aligned}$$

which yields the desired result. □

Proof of Theorem 3.2.21 . Let $h(x) = \widetilde{f}(x) g(x)$ and $\widetilde{f}(x) = f(\sqrt{2\pi} x)$ then one can easily see that

$$\int_{\mathbb{R}} |h|^2 dx = 1, \quad \int_{\mathbb{R}} h dx = 2^{\frac{1}{4}}, \quad \int_{\mathbb{R}} x^2 h dx = \frac{2^{-\frac{3}{4}} k}{\pi}. \tag{3.4.3}$$

Let X_1, X_2 be i.i.d. random variables with the density $p(x) = \frac{2^{-1/4}}{\sqrt{\pi}} h(\frac{x}{\sqrt{\pi}})$ and the distribution function F. Note that F has median zero and $\mathrm{Var}[X_1] = \frac{k}{2}$. Since the Kolmogorov distance is bounded by the total variation, one can see that

$$d_{\mathrm{K}}(F * F, \Phi_1) \leq \frac{1}{2} \int_{\mathbb{R}} |p * p(x) - \gamma(x)| dx.$$

Since we have $h * h(x) = \sqrt{2\pi} p * p(\sqrt{\pi} x)$ and $g * g(x) = \sqrt{2\pi}\gamma(\sqrt{\pi}x)$, we obtain

$$d_{\mathrm{K}}(F * F, \Phi_1) \leq \frac{1}{2\sqrt{2}} \int_{\mathbb{R}} |h * h(x) - g * g(x)| dx.$$

Let $u := h * h - g * g$, then we have $\|u\|_1 \leq 2\sqrt{2}$ and

$$\begin{aligned} \int_{\mathbb{R}} x^2 |u| dx &\leq \int_{\mathbb{R}} x^2 (h * h)(x) dx + \int_{\mathbb{R}} x^2 (g * g)(x) dx \\ &\leq 2^{\frac{5}{4}} \Big(\int_{\mathbb{R}} x^2 h(x) dx + \int_{\mathbb{R}} x^2 g(x) dx \Big) \\ &\leq C(k+1). \end{aligned}$$

By Lemma 3.4.2, we have $\|u\|_1 \leq C_k \|u\|_2$ where $C_k > 0$ depends only on k. Combining our observation with Theorem 3.4.1, we obtain

$$d_{\mathrm{K}}(F * F, \Phi_1) \leq C_k (\|h - g\|_{\frac{6}{5}}^2 + \|h - g\|_2)^{\frac{3}{4}} \delta_c(\widetilde{f})^{\frac{1}{8}}$$

where $\widetilde{f} = f(\sqrt{2\pi}x)$. Note that $\delta_c(\widetilde{f}) = \delta(v)$. It follows from (3.4.2) and (3.4.3) that $(\|h - g\|_{\frac{6}{5}}^2 + \|h - g\|_2)^{\frac{3}{4}}$ is bounded by a universal constant and that

$$d_{\mathrm{K}}(F * F, \Phi_1) \leq C_k \delta(v)^{\frac{1}{8}}.$$

Choose $\varepsilon_0 > 0$ such that $C_k \varepsilon_0^{\frac{1}{8}} < 1$. Let $\varepsilon > 0$ be such that $\delta(v) < \varepsilon < \varepsilon_0$, and put $\eta = C_k \varepsilon^{\frac{1}{8}}$, $N = N(\eta) = 1 + \sqrt{2 \log(1/\eta)}$ and

$$\sigma(\eta)^2 = \int_{-N(\eta)}^{N(\eta)} x^2 p(x) dx.$$

Note that $\sigma(\eta)^2 \nearrow \mathrm{Var}[X_1] = \frac{1}{2}m_2(\mu)$ as $\eta \to 0$. So, we choose ε_0 small enough so that $\frac{1}{4}m_2(\mu) < \sigma(\eta)^2$ for all $\varepsilon < \varepsilon_0$. It then follows from Theorem 2.3.1 that

$$\begin{aligned} d_{\mathrm{K}}(F, \Phi_{\sigma(\eta)}) &\leq \frac{C}{\sigma(\eta)\sqrt{\log(1/\eta)}} \min\left\{ \frac{1}{\sqrt{\sigma(\eta)}}, \log\log \frac{e^e}{\eta} \right\} \\ &\leq \frac{C}{\sigma(\eta)^{\frac{3}{2}}\sqrt{\log(1/\eta)}} \\ &\leq \frac{C}{m_2(\mu)^{\frac{3}{4}}\sqrt{\frac{1}{8}\log(\frac{1}{\varepsilon}) - \log C_k}} \\ &\leq \frac{C_k}{\sqrt{\log\frac{1}{\varepsilon}}}. \end{aligned}$$

By change of variables, we have $d_{\mathrm{K}}(F, \Phi_{\sigma(\eta)}) = d_{\mathrm{K}}(\mu, \gamma_\varepsilon)$ where

$$d\gamma_\varepsilon = \frac{1}{\sqrt{4\pi\sigma(\eta)^2}} e^{-\frac{|x|^2}{4\sigma(\eta)^2}} dx,$$

which yields (3.2.7). □

Proof of Theorem 3.2.22 . Let $h(x) = \widetilde{f}(x)g(x)$ and $\widetilde{f}(x) = f(\sqrt{2\pi}x)$. Let X_1, X_2 be i.i.d. random variables with the density $p(x) = \frac{2^{-1/4}}{\sqrt{\pi}} h(\frac{x}{\sqrt{\pi}})$ and the distribution function F. Since $m_2(\mu) = 1$, we have $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \frac{1}{2}$. The same argument then leads to

$$d_{\mathrm{K}}(F * F, \Phi_1) \leq c_1 \delta_c(\widetilde{f})^{\frac{1}{8}} = c_1 \delta(v)^{\frac{1}{8}}$$

for some universal constant c_1. So, we choose $\varepsilon_0 > 0$ such that $c_1 \varepsilon_0^{\frac{1}{8}} < 1$. Assume $\delta(v) < \varepsilon < \varepsilon_0$. We apply Theorem 2.3.2 to obtain

$$d_{\mathrm{K}}(F, \Phi_{\frac{1}{\sqrt{2}}}) < \frac{c_2}{\sqrt{\log\frac{1}{\varepsilon}}}.$$

Note that $d_{\mathrm{K}}(F, \Phi_{\frac{1}{\sqrt{2}}}) = d_{\mathrm{K}}(\mu, \gamma)$ by change of variables. Let $\Psi(s)$ be the inverse of the map $t \mapsto \frac{c_2}{\sqrt{\log\frac{1}{t}}}$, then $\delta(v) \geq \Psi(d_{\mathrm{K}}(\mu, \gamma))$ as desired. □

4. INSTABILITY OF THE LOGARITHMIC SOBOLEV INEQUALITY

We have seen that there are different types of stability estimates for the LSI according to the choice of probability measure spaces and distances. A natural question is to find the best possible probability measure space and distance in which the LSI is stable. In this chapter, we investigate this question. To be specific, we show that there are no stability in $\mathcal{P}_2^M(\mathbb{R})$ (resp. $\mathcal{P}_2(\mathbb{R})$) with respect to W_2 and $L^p(d\gamma)$ for $p > 1$ (resp. W_1). This chapter is based on my work [85].

4.1 Main results

The first result shows that the W_2-stability estimate obtained in [26, Corollary 1.2] cannot be improved in terms of the probability measure space $\mathcal{P}_2^1(\mathbb{R})$. It also implies that the L^1-stability estimate in Theorem 3.2.1 is best possible in terms of the L^p distances. Note that there is an L^p-stability estimate in $\mathcal{P}_2^M(\mathbb{R})$ $(p > 1)$ with a higher integrability assumption (see Corollary 3.2.2).

Theorem 4.1.1. *Let $M > 1$ and $p > 1$. There exists a sequence of centered probability measures $d\nu_k = f_k d\gamma$ in $\mathcal{P}_2^M(\mathbb{R})$ such that $\lim_{k\to\infty} \delta(f_k) = 0$,*

$$\liminf_{k\to\infty} W_2(\nu_k, \gamma) \geq C_1,$$

and

$$\liminf_{k\to\infty} \|f_k - 1\|_{L^p(d\gamma)} \geq C_2,$$

for some $C_1, C_2 > 0$.

The next result is W_1-instability in $\mathcal{P}_2(\mathbb{R})$, which implies that the W_1-stability estimate in Theorem 3.2.6 cannot be improved in terms of the space $\mathcal{P}_2^M(\mathbb{R})$.

Theorem 4.1.2. *There exists a sequence of centered probability measures* $d\nu_k = f_k d\gamma$ *in* $\mathcal{P}_2(\mathbb{R})$ *such that* $\lim_{k\to\infty} \delta(f_k) = 0$ *and* $\lim_{k\to\infty} W_1(\nu_k, \gamma) = \infty$.

The key idea of the proofs is as follows. Using the class of the LSI optimizers, we construct a sequence of centered probability measures with a small deficit. We then control the second moments and the relative entropies so as to conclude that the distances from the standard Gaussian measure, which is the only centered optimizer, do not converge to zero.

Theorem 4.1.1 and 4.1.2 deal with probability measures on the real line. These results, however, can be directly generalized to the higher dimensional case. Let ν_k be the sequence of probability measures on $\mathbb{R}$ constructed in Example 4.3.2 and γ_{n-1} the standard Gaussian measure on $\mathbb{R}^{n-1}$. If we define a probability measure $\widetilde{\nu_k}$ on $\mathbb{R}^n$ by $\widetilde{\nu_k} = \nu_k \otimes \gamma_{n-1}$, then we have $\mathrm{I}(\widetilde{\nu_k}) = \mathrm{I}(\nu_k)$, $\mathrm{H}(\widetilde{\nu_k}) = \mathrm{H}(\nu_k)$, and $\delta(\widetilde{\nu_k}) = \delta(\nu_k)$. Furthermore, we have $m_2(\widetilde{\nu_k}) = (n-1) + m_2(\nu_k)$ and $m_1(\widetilde{\nu_k}) \geq m_1(\nu_k) - m_1(\gamma_{n-1})$. Controlling the second moment and the relative entropy of ν_k as in the proofs of Theorem 4.1.1 and 4.1.2, we extend the results to $\mathbb{R}^n$ for $n \geq 2$.

In Proposition 4.3.4, we show that the sequence ν_k in $\mathcal{P}_2(\mathbb{R})$ constructed in Theorem 4.1.2 converges to γ in $L^1(d\gamma)$. Thus it is still open to show L^1-stability in $\mathcal{P}_2$. Note that if $\mathrm{H}(f)$ is finite, then it follows from Jensen's inequality that the second moment is finite. So $\mathcal{P}_2$ is the most general probability measure space in the setting of the LSI.

4.2 Literature review

We review previous stability results and compare the probability spaces and the conditions used in this literature.

4.2.1 Wasserstein distance

Indrei and Marcon [83] showed that if $fd\gamma$ is a centered probability measure with the log-concavity condition on the density

$$-1+\varepsilon \leq D^2(\log(\frac{1}{f})) \leq M \tag{4.2.1}$$

for $\varepsilon, M > 0$, then

$$\delta(f) \geq C_{\varepsilon,M} W_2^2(fd\gamma, d\gamma). \tag{4.2.2}$$

Their method relies on the optimal transport method (Theorem 2.2.4) and Caffarelli's contraction theorem [38]. Note that W_2–stability cannot hold for all probability measures since it would improve the constant in the sharp LSI (see [83, Remark 4.3]).

Let $\lambda > 0$ and $\mathcal{P}(\lambda)$ be the space of probability measure $fd\gamma$ satisfying a $(2,2)$-Poincaré inequality with a constant λ: for every smooth function g with $\int g\, fd\gamma = 0$,

$$\lambda \int g^2\, fd\gamma \leq \int |\nabla g|^2\, fd\gamma. \tag{4.2.3}$$

It was shown in [57] that every probability measure $fd\gamma \in \mathcal{P}(\lambda)$ for $\lambda > 0$ satisfies the following improvement of the LSI

$$\mathrm{H}(f) \leq \frac{c(\lambda)}{2}\, \mathrm{I}(f) \tag{4.2.4}$$

where $c(\lambda) = \frac{1-\lambda+\lambda\log\lambda}{(1-\lambda)^2} < 1$. The proof is based on an interpolation along the Ornstein–Uhlenbeck semigroup. In particular, this yields W_2–stability

$$\delta(f) \geq c_1(\lambda) W_2^2(fd\gamma, d\gamma). \tag{4.2.5}$$

where $c_1(\lambda) = \frac{1}{2}(\frac{1}{c(\lambda)} - 1)$. Note that every probability measure $fd\gamma$ with (4.2.1) satisfies a $(2,2)$-Poincaré inequality. Thus the W_2–stability bound (4.2.5) of [57] is an improvement of (4.2.2). We note that if $fd\gamma \in \mathcal{P}(\lambda)$ then

$$m_2(fd\gamma) = \int |x|^2\, fd\gamma \leq \frac{n}{\lambda}$$

by (4.2.3) with $g(x) = x_i$ for $i = 1, 2, \cdots, n$. Thus we have $\mathcal{P}(\lambda) \subseteq \mathcal{P}_2^{n/\lambda}(\mathbb{R}^n)$.

The Fisher information and the relative entropy with respect to $d\gamma$ have different scaling. From this observation, W_2–stability was derived in [26] (see also [53, Theorem 1] and Proposition 2.2.7), which states that if $fd\gamma \in \mathcal{P}_2^n(\mathbb{R}^n)$ is centered, then

$$\delta(f) \geq CW_2(fd\gamma, d\gamma)^4.$$

4.2.2 Total variation distance

One of the consequences of (4.2.4) in [57] is an L^1-stability estimate, which states that if $fd\gamma$ satisfies $(2,2)$-Poincaré inequality with constant λ then

$$\delta(f) \geq c_2(\lambda)\|f-1\|^2_{L^1(d\gamma)}$$

where $c_2(\lambda) = \frac{1}{4}(1-c(\lambda))$.

In [58, Proposition 4.7], the authors proved that if $fd\gamma$ is a probability measure satisfying

$$\mathcal{F}(e^{-\pi|x|^2} f(2\sqrt{\pi}x)) \geq 0, \tag{4.2.6}$$

then

$$\delta(f) \geq \frac{1}{2}\|f-1\|_2^4, \tag{4.2.7}$$

which also implies an L^1–stability estimate. The proof is based on Carlen's deficit bound (2.2.4) and Pinsker's inequality (2.1.2). It is remarkable that the positivity of the Fourier transform is quite different from $\mathcal{P}_2^M$. Indeed, the spaces of probability measures $fd\gamma$ satisfying (4.2.6) is not included in $\mathcal{P}_2^M$ for any M, and vice versa.

Proposition 4.2.1. *Let $\mathcal{S}$ be the space of probability measures $fd\gamma$ satisfying (4.2.6). For any $M > 0$, we have $\mathcal{S} \not\subset \mathcal{P}_2^M(\mathbb{R}^n)$ and $\mathcal{P}_2^M(\mathbb{R}^n) \not\subset \mathcal{S}$.*

Proof. Since the LSI is L^2-stable in $\mathcal{S}$ by (4.2.7), Theorem 4.1.1 implies that $\mathcal{P}_2^M \not\subset \mathcal{S}$ for all $M > 0$. Let $f_k d\gamma$ be the centered Gaussian with variance k, then $\{f_k d\gamma\}$ is not included in $\mathcal{P}_2^M$ for any $M > 0$. Since $e^{-\pi|x|^2}\sqrt{f_k(2\pi x)}$ is also Gaussian, its Fourier transform is positive. Thus we get $\mathcal{S} \not\subset \mathcal{P}_2^M$. □

We note that the positivity condition for the Fourier transform can be relaxed in a sense that $\mathcal{F}(e^{-\pi|x|^2} f(2\pi x)^{\frac{1}{2}})$ belongs to some region in the complex plane. See [58].

4.3 Examples

In this section, we construct a sequence of centered probability measures to prove Theorem 4.1.1 and Theorem 4.1.2. First, we find a sequence of centered probability measures such that the deficit of the LSI goes to 0. By Lemma 4.4.1 and (2.1.2), it is enough to control the second moments and the relative entropies of the sequence to show that it does not converge to γ in the Wasserstein distances and the $L^p(d\gamma)$ distances for $p > 1$.

Recall that $\delta(f) = 0$ if and only if $f(x) = \exp(b \cdot x - \frac{1}{2}|b|^2)$, for $b \in \mathbb{R}^n$. We start with a trivial example.

Example 4.3.1. Let $b \in \mathbb{R}^n$, $g_b(x) = e^{b\cdot x - \frac{|b|^2}{2}}$, and $d\nu_b = g_b d\gamma$. Since g_b are the optimizers of the LSI, we have $\delta(g_b) = 0$ for all $b \in \mathbb{R}^n$. Indeed, a direct calculation yields that

$$\begin{aligned}
\mathrm{I}(\nu_b) &= \int_{\mathbb{R}^n} \frac{|\nabla g_b|^2}{g_b} d\gamma = |b|^2 \int_{\mathbb{R}^n} g_b d\gamma = |b|^2, \\
\mathrm{H}(\nu_b) &= \int_{\mathbb{R}^n} g_b \log g_b d\gamma = \int_{\mathbb{R}^n} \left(b \cdot (x+b) - \frac{1}{2}|b|^2 \right) d\gamma = \frac{1}{2}|b|^2, \\
m_2(\nu_b) &= \int_{\mathbb{R}^n} |x|^2 g_b d\gamma = \int_{\mathbb{R}^n} |x+b|^2 d\gamma = n + |b|^2.
\end{aligned}$$

Note that $\mathrm{I}(\nu_b)$, $\mathrm{H}(\nu_b)$, and $m_2(\nu_b)$ all tend to ∞, as $|b| \to \infty$. Notice also that the measure $g_b d\gamma$ is not centered provided $b \neq 0$.

Now we present the main example.

Example 4.3.2. Let $g_b(x) = e^{bx - \frac{b^2}{2}}$ and $\gamma(x) = (2\pi)^{-\frac{1}{2}} e^{-x^2/2}$ for $x, b \in \mathbb{R}$. We denote by $d\gamma = \gamma(x)dx$ and set $\Phi(x) = \int_{-\infty}^{x} d\gamma$. For each $k \in \mathbb{N}$, let $\widetilde{f}_k$ be a function in $C^\infty(\mathbb{R})$ such that

$$\widetilde{f}_k(x) = \begin{cases} 1, & x \in [0, k] \\ l_k(x), & x \in (k, k + \frac{1}{k}] \\ \alpha g_b(x), & x \in (k + \frac{1}{k}, \infty) \end{cases}$$

and $\widetilde{f}_k(x) = \widetilde{f}_k(-x)$ where

(i) $b = b_k = 2(k + \frac{1}{k}) + \sqrt{k}$,

(ii) $\alpha = \alpha_k = v b_k^{-w} \in (0, \frac{1}{2})$ for $v, w > 0$,

(iii) $l_k \in C^\infty(\mathbb{R})$ satisfies $l_k(k) = 1$, $l_k(k + \frac{1}{k}) = \alpha g_b(k + \frac{1}{k})$, $|l_k'(x)| \leq 2k$, and

$$\alpha g_b(k + \frac{1}{k}) \leq l_k(x) \leq 1$$

for all $x \in (k, k + \frac{1}{k}]$.

We observe that $\alpha g_b(k + \frac{1}{k}) < \frac{1}{2}$ for all $k \in \mathbb{N}$. Note also that $\widetilde{f}_k \in L^1(d\gamma)$ and

$$\begin{aligned}\int_{\mathbb{R}} \widetilde{f}_k d\gamma &= 2\int_0^k d\gamma + 2\int_k^{k+\frac{1}{k}} l_k(x) d\gamma + 2\alpha \int_{k+\frac{1}{k}}^{\infty} g_b(x) d\gamma \\ &= (2\Phi(k) - 1) + 2\int_k^{k+\frac{1}{k}} l_k(x) d\gamma + 2\alpha\Phi(b - k - \frac{1}{k}).\end{aligned}$$

Since $l_k(x) \leq 1$ and

$$\left| \int_k^{k+\frac{1}{k}} l_k(x) d\gamma \right| \leq \int_k^{k+\frac{1}{k}} \gamma(x) dx \leq \frac{1}{k}\gamma(k) = o(1),$$

we have $\int_{\mathbb{R}} \widetilde{f}_k d\gamma \to 1$, as $k \to \infty$. Let $c_k = (\int_{\mathbb{R}} \widetilde{f}_k d\gamma)^{-1}$ and define $f_k = c_k \widetilde{f}_k$ and $d\nu_k = f_k d\gamma$. The constants v and w in $\alpha = \alpha_k$ will be determined later. They play a role in controlling the second moment and the relative entropy of ν_k. Note that the following lemma and proposition do not depend on the choices of v and w.

Lemma 4.3.3. *Let f_k and ν_k be defined as in Example 4.3.2. Then we have*

$$\lim_{k\to\infty} \delta(f_k) = 0.$$

Proof. Direct computations give

$$\begin{aligned}\mathrm{I}(f_k) &= 2c_k \int_k^{k+\frac{1}{k}} \frac{|l_k'(x)|^2}{l_k(x)} d\gamma + 2c_k\alpha \int_{k+\frac{1}{k}}^{\infty} \frac{|g_b'(x)|^2}{g_b(x)} d\gamma \\ &= 2c_k \int_k^{k+\frac{1}{k}} \frac{|l_k'(x)|^2}{l_k(x)} d\gamma + 2c_k\alpha b^2 \Phi(b - k - \frac{1}{k})\end{aligned}$$

and

$$\begin{aligned}\mathrm{H}(f_k) &= 2\int_0^k c_k \log c_k d\gamma + 2\int_k^{k+\frac{1}{k}} c_k l_k(x)\log(c_k l_k(x))d\gamma \\ &\quad + 2\int_{k+\frac{1}{k}}^{\infty}(c_k\alpha g_b)\log(c_k\alpha g_b)d\gamma \\ &= (c_k\log c_k)(2\Phi(k)-1) + 2\int_k^{k+\frac{1}{k}} c_k l_k(x)\log(c_k l_k(x))d\gamma \\ &\quad + 2c_k\alpha\log(c_k\alpha)\Phi(b-k-\frac{1}{k}) + 2c_k\alpha b\gamma(b-k-\frac{1}{k}) + c_k\alpha b^2\Phi(b-k-\frac{1}{k}).\end{aligned} \tag{4.3.1}$$

Thus the deficit of the LSI is

$$\begin{aligned}\delta(f_k) = c_k\int_k^{k+\frac{1}{k}}\frac{|l_k'(x)|^2}{l_k(x)}d\gamma - 2\int_k^{k+\frac{1}{k}} c_k l_k(x)\log(c_k l_k(x))d\gamma - (c_k\log c_k)(2\Phi(k)-1) \\ - 2c_k\alpha\log(c_k\alpha)\Phi(b-k-\frac{1}{k}) - 2c_k\alpha b\gamma(b-k-\frac{1}{k}).\end{aligned}$$

Note that $c_k \to 1$ and $\alpha \to 0$, as $k \to \infty$. Since the limits of the map $t \mapsto t\log t$ at $t = 0$ and $t = 1$ is 0, we have

$$\lim_{k\to\infty}\big((c_k\log c_k)(2\Phi(k)-1) + 2c_k\alpha\log(c_k\alpha)\Phi(b-k-\frac{1}{k})\big) = 0.$$

By the construction of α_k and b_k, we have

$$\lim_{k\to\infty}\alpha b\gamma(b-k-\frac{1}{k}) = \lim_{k\to\infty}\frac{1}{\sqrt{2\pi}}vb^{1-w}e^{-\frac{1}{2}b\sqrt{k}} = 0.$$

By the construction of l_k, we have

$$l_k(x) \geq \alpha g_b(k+\frac{1}{k}),$$

which yields

$$\begin{aligned}\left|\int_k^{k+\frac{1}{k}}\frac{|l_k'(x)|^2}{l_k(x)}d\gamma\right| &\leq \frac{4k^2}{\alpha g_b(k+\frac{1}{k})}\int_k^{k+\frac{1}{k}} d\gamma \\ &\leq \frac{4k\gamma(k)}{\alpha g_b(k+\frac{1}{k})} \\ &= \frac{4k}{\sqrt{2\pi}\alpha}e^{-\frac{1}{2}(k^2-b\sqrt{k})} = o(1).\end{aligned}$$

Choose $k_0 \in \mathbb{N}$ such that $\frac{1}{2} \leq c_k \leq \frac{3}{2}$ for all $k \geq k_0$. Since $l_k(x) \leq 1$ for all k, there exists a constant C such that $|c_k l_k(x) \log(c_k l_k(x))| \leq C$ for all $k \geq k_0$. So we have

$$\left| \int_k^{k+\frac{1}{k}} c_k l_k(x) \log(c_k l_k(x)) d\gamma \right| \leq \frac{C}{k} \gamma(k) = o(1),$$

for $k \geq k_0$. Therefore we conclude that $\delta(f_k) \to 0$ as $k \to \infty$ as desired. □

Proposition 4.3.4. *Let f_k and ν_k be defined as in Example 4.3.2. Then, $f_k \to 1$ in $L^1(d\gamma)$. As a consequence, $\nu_k \rightharpoonup \gamma$ weakly as $k \to \infty$.*

Proof. By (2.1.6), it suffices to show that $\|\sqrt{f_k} - 1\|_{L^2(d\gamma)} \to 0$. A direct computation yields

$$\begin{aligned}
\|\sqrt{f_k} - 1\|^2_{L^2(d\gamma)} &= 2\int_0^k |\sqrt{c_k} - 1|^2 d\gamma + 2\int_k^{k+\frac{1}{k}} |\sqrt{c_k l_k} - 1|^2 d\gamma \\
&\quad + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - 1|^2 d\gamma \\
&= |\sqrt{c_k} - 1|^2 (2\Phi(k) - 1) + 2\int_k^{k+\frac{1}{k}} |\sqrt{c_k l_k} - 1|^2 d\gamma \\
&\quad + c_k \alpha \Phi(b - k - \frac{1}{k}) - 2\sqrt{c_k \alpha} e^{-\frac{b^2}{8}} \Phi(\frac{b}{2} - k - \frac{1}{k}) + \Phi(-k - \frac{1}{k}) \\
&= o(1) + 2\int_k^{k+\frac{1}{k}} |\sqrt{c_k l_k} - 1|^2 d\gamma.
\end{aligned}$$

It follows from the assumption on $l_k(x)$ that $\left| \int_k^{k+\frac{1}{k}} |\sqrt{c_k l_k} - 1|^2 d\gamma \right| \leq \frac{2(c_k+1)}{k} \gamma(k) = o(1)$, which leads to $\|\sqrt{f_k} - 1\|_{L^2(d\gamma)} = o(1)$ as desired. □

4.4 Proofs of the main results

4.4.1 Lemma

Let $\{\nu_k\}$ be a sequence of probability measures in $\mathcal{P}_p(\mathbb{R}^n)$. The next lemma gives a sufficient condition for the sequence $\{\nu_k\}$ not converging to a measure μ in the W_p metric. In the proof of Theorem 4.1.1 and Theorem 4.1.2, we control the second moments to conclude that the W_p distance does not converge to γ.

Lemma 4.4.1. *Let $p \geq 1$ and $\mu, \mu_k \in \mathcal{P}_p(\mathbb{R}^n)$ for $k \geq 1$. If there exists a constant $C_1 > 0$ such that*

$$\liminf_{k\to\infty}(m_p(\mu_k) - m_p(\mu)) \geq C_1,$$

then $\liminf_{k\to\infty} W_p^p(\mu, \mu_k) \geq C_2$, for some $C_2 > 0$.

Proof. Let $t > 0$, then there exists a constant $C_t > 0$ such that

$$|x|^p - |y|^p \leq t|y|^p + C_t|x - y|^p$$

for any $x, y \in \mathbb{R}^n$. Let π_k be a probability measure on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ_k and μ. Taking the integral with respect to $d\pi_k$, we get

$$m_p(\mu_k) - m_p(\mu) \leq tm_p(\mu) + C_t \int_{\mathbb{R}^n\times\mathbb{R}^n} |x - y|^p d\pi_k(x, y).$$

We take the infimum over all such π_k to get

$$m_p(\mu_k) - m_p(\mu) \leq tm_p(\mu) + C_t W_p^p(\mu, \mu_k).$$

Let $t_1 \in (0, C_1)$ and choose $k_0 \in \mathbb{N}$ large enough that

$$\liminf_{j\to\infty}(m_p(\mu_j) - m_p(\mu)) - t_1 < m_p(\mu_k) - m_p(\mu)$$

for all $k \geq k_0$. Put $C_3 = \liminf_{j\to\infty}(m_p(\mu_j) - m_p(\mu)) - t_1$, then

$$C_3 \leq tm_p(\mu) + C_t W_p^p(\mu, \mu_k)$$

for all $k \geq k_0$. We finish the proof by choosing $t = \frac{C_3}{2m_p(\mu)+1} > 0$ and $C_2 = \frac{C_3}{2C_t}$. □

4.4.2 Proof of Theorem 4.1.1

Let $w = 2$ and $v \in (0, (M-1)/4)$ be such that $vb_k^{-2} < \frac{1}{2}$ for all k. Define f_k and ν_k as in Example 4.3.2 with $b_k = 2(k + \frac{1}{k}) + \sqrt{k}$ and $\alpha_k = vb_k^{-2}$. The second moment of ν_k is

$$\begin{aligned} m_2(\nu_k) &= 2c_k \int_0^k x^2 d\gamma + 2c_k \int_k^{k+\frac{1}{k}} x^2 l_k(x) d\gamma + 2c_k\alpha \int_{k+\frac{1}{k}}^{\infty} x^2 g_b(x) d\gamma \qquad (4.4.1) \\ &= c_k(2\Phi(k) - 1 - 2k\gamma(k)) + 2c_k \int_k^{k+\frac{1}{k}} x^2 l_k(x) d\gamma + 2c_k\alpha b^2 \Phi(b - k - \frac{1}{k}) \\ &\quad + 2c_k\alpha \left(\Phi(b - k - \frac{1}{k}) - (b - k - \frac{1}{k})\gamma(b - k - \frac{1}{k}) + 2b\gamma(b - k - \frac{1}{k}) \right). \end{aligned}$$

Note that $c_k(2\Phi(k) - 1 - 2k\gamma(k)) \to 1$ and

$$\lim_{k\to\infty} 2c_k\alpha\left(\Phi(b-k-\frac{1}{k}) - (b-k-\frac{1}{k})\gamma(b-k-\frac{1}{k}) + 2b\gamma(b-k-\frac{1}{k})\right) = 0.$$

Since $\left|\int_k^{k+\frac{1}{k}} x^2 l_k(x)d\gamma\right| \leq \frac{(k+\frac{1}{k})^2}{k}\gamma(k) = o(1)$, we obtain

$$\begin{aligned}\lim_{k\to\infty} m_2(\nu_k) &= 1 + \lim_{k\to\infty} 2c_k\alpha b^2\Phi(b-k-\frac{1}{k}) \\ &= 1 + 2v\lim_{k\to\infty} c_k\Phi(b-k-\frac{1}{k}) \\ &= 1 + 2v.\end{aligned}$$

Since $1+2v < M$, there exists $N \in \mathbb{N}$ such that $\{\nu_k\}_{k\geq N} \subset \mathcal{P}_2^M$. Since $\mu_2(\gamma) = 1$ and

$$\lim_{k\to\infty}(m_2(\nu_k) - \mu_2(\gamma)) = 2v > 0,$$

it follows from Lemma 4.4.1 that $\liminf_{k\to\infty} W_2(\mu,\mu_k) \geq C$, for some $C > 0$. By (4.3.1), we have

$$\begin{aligned}\mathrm{H}(f_k) &= o(1) + 2c_k v b^{-2} b\gamma(b-k-\frac{1}{k}) + c_k v b^{-2} b^2\Phi(b-k-\frac{1}{k}) \\ &= o(1) + v c_k\Phi(b-k-\frac{1}{k}),\end{aligned}$$

which implies that $\mathrm{H}(f_k) \to v > 0$. By (2.1.2), we conclude that $\|f_k - 1\|_{L^p(d\gamma)}$ does not converge to zero for $p > 1$. □

4.4.3 Proof of Theorem 4.1.2

Let ν_k and f_k be defined as in Example 4.3.2 with $\alpha = b^{-\frac{1}{2}}$ (i.e. $v = 1$ and $w = \frac{1}{2}$). Note that $m_2(\nu_k) < \infty$ for all k and $m_2(\nu_k) \to \infty$ as $k \to \infty$ by (4.4.1). By Lemma 4.4.1 it is enough to show that $m_1(\nu_k)$ does not converge to $m_1(\gamma)$. By the construction of ν_k, we have

$$\begin{aligned}m_1(\nu_k) &= \int_{\mathbb{R}} |x| d\nu_k \\ &= 2c_k\int_0^k |x|d\gamma + 2c_k\int_k^{k+\frac{1}{k}} |x| l_k(x)d\gamma + 2c_k\alpha\int_{k+\frac{1}{k}-b}^{\infty} |x+b|d\gamma.\end{aligned}$$

We observe that $2c_k \int_0^k |x| d\gamma = o(1) + m_1(\gamma)$,

$$\left| \int_k^{k+\frac{1}{k}} |x| l_k(x) d\gamma \right| \leq \left(k + \frac{1}{k}\right) \frac{\gamma(k)}{k} = o(1),$$

and

$$\begin{aligned} 2c_k\alpha \int_{k+\frac{1}{k}-b}^{\infty} |x+b| d\gamma &\geq 2c_k\alpha \int_{k+\frac{1}{k}-b}^{\infty} (b - |x|) d\gamma \\ &\geq 2c_k\alpha b \Phi(b - k - \frac{1}{k}) - 2c_k\alpha m_1(\gamma). \end{aligned}$$

Since we have $\alpha b = b^{\frac{1}{2}} \to \infty$, we conclude that $m_1(\nu_k) \to \infty$. By Lemma 4.4.1, the proof is complete. □

Remark 4.4.2. We summarize what we have seen in this section. Let ν_k and f_k be as in Example 4.3.2. Note that $\alpha_k = v b_k^{-w}$ and $b = 2(k + \frac{1}{k}) + \sqrt{k}$. According to the computations above, we have

$$\mathrm{H}(\nu_k) = o(1) + c_k v b^{2-w} \Phi(b - k - \frac{1}{k})$$

and

$$o(1) + 2^{1-p} c_k v b^{p-w} \leq m_p(\nu_k) - m_p(\gamma) \leq o(1) + 2^p c_k v b^{p-w}$$

for all $p \geq 1$. For any $v, w > 0$, we have $\nu_k \in \mathcal{P}_2(\mathbb{R})$, $\delta(\nu_k) \to 0$, $\|f_k - 1\|_{L^1(d\gamma)} \to 0$, and $\nu_k \rightharpoonup \gamma$. The followings describe the behaviors of the relative entropy and the second moment of ν_k in terms of w.

(i) If $w > 2$, then $\mathrm{H}(\nu_k) \to 0$ and $m_2(\nu_k) \to m_2(\gamma)$ so that no instability results can be obtained.

(ii) If $w = 2$, then $m_2(\nu_k)$ does not converge to $m_2(\gamma)$ which implies that $W_2(\nu_k, \gamma) \not\to 0$. In this case, $m_2(\nu_k)$ can be bounded by some constant so that $\nu_k \in \mathcal{P}_2^M$.

(iii) If $w < 2$, then $m_2(\nu_k)$ goes to ∞ so that ν_k does not belong to $\mathcal{P}_2^M$ for any $M > 0$. In this case, we have $m_p(\nu_k) - m_p(\gamma) \not\to 0$ for any $p \geq w$. So $W_p(\nu_k, \gamma) \not\to 0$.

(iv) The relative entropy $\mathrm{H}(\nu_k) \not\to 0$ if and only if $w \leq 2$.

5. INSTABILITY FOR BECKNER–HIRSCHMAN INEQUALITY

In the previous chapter, we constructed a sequence of probability measures such that the LSI deficit converges to 0 but the distances from the Gaussian measure does not. Note that we have seen in (2.2.5) that the deficit of the BHI is bounded above by that of the LSI. As an application of Example 4.3.2, we prove that there are no stability for the Beckner–Hirschman inequality (the BHI) in terms of the normalized L^p distances with some weighted measures and range of p. This chapter is based on my work [85].

5.1 Main results

For a nonnegative function h on $\mathbb{R}$ with $\|h\|_2 = 1$, the Beckner–Hirschman inequality states that

$$\delta_{\mathrm{BH}}(h) = S(|h|^2) + S(|\widehat{h}|^2) - (1 - \log 2) \geq 0$$

where $S(\rho)$ is the entropy of ρ defined as in (2.2.2), and $\delta_{\mathrm{BH}}(h)$ is the deficit of the BHI. It is also called the entropic uncertainty principle. We say that a function h is an optimizer for the BHI if $\delta_{\mathrm{BH}}(h) = 0$. Let $\mathfrak{G}$ be the set of all nonnegative, L^2-normalized optimizers for the BHI. Using the fact that the optimizers are Gaussian (see [91] and [40, p.207]), we get

$$\mathfrak{G} = \{G_{a,r}(x) = \left(\frac{2a}{\pi}\right)^{\frac{1}{4}} e^{-a(x-r)^2} : a > 0, r \in \mathbb{R}\}. \tag{5.1.1}$$

We denote by $G_a(x) = G_{a,0}(x)$ and $g(x) = G_\pi(x)$. For a measure μ on $\mathbb{R}$ and $p > 0$, we define

$$\mathrm{dist}_{L^p(d\mu)}(h, \mathfrak{G}) = \inf_{u \in \mathfrak{G}} \|h - u\|_{L^p(d\mu)} = \inf_{a>0, r\in\mathbb{R}} \|h - G_{a,r}\|_{L^p(d\mu)}.$$

The key element of the application is that the deficit of the LSI is bounded below by that of the BHI. To be specific, we have $\delta(f) \geq \delta_{\mathrm{BH}}(h)$ where

$$h(x) = (f(2\sqrt{\pi}x))^{\frac{1}{2}} g(x). \tag{5.1.2}$$

Let f_k be a sequence of functions constructed in Example 4.3.2 and h_k the transformation of f_k by (5.1.2), then we have $\delta_{\mathrm{BH}}(h_k) \to 0$; see Lemma 4.3.3. Note that h_k is indeed a Gaussian function with small Gaussian bumps in the tails. In the proofs of Theorem 4.1.1 and Theorem 4.1.2, we have seen that the growth of the second moments of the probability measures $\{f_k d\gamma\}$ can be controlled by the choice of parameters. This implies that the Gaussian bumps of h_k in the tails are not negligible with respect to measures with some polynomial weights. This observation leads us to adopt the polynomial measure $d\eta_\lambda = |x|^\lambda dx$.

Theorem 5.1.1. *Let $\lambda > 0$, $d\eta_\lambda = |x|^\lambda dx$, and $p \geq 2(\lambda+1)$, then there exists a sequence of nonnegative functions $\{h_k\}_{k\geq 1}$ in $L^p(d\eta_\lambda)$ such that $\|h_k\|_2 = 1$, $\delta_{\mathrm{BH}}(h_k) \to 0$, $\|h_k\|_{L^p(d\eta_\lambda)} \to \infty$, and*

$$\liminf_{k\to\infty} \frac{\mathrm{dist}_{L^p(d\eta_\lambda)}(h_k, \mathfrak{G})}{\|h_k\|_{L^p(d\eta_\lambda)}} \geq C(p,\lambda) > 0.$$

Inspired by the transformation (5.1.2), it is natural to consider a reference measure with a Gaussian weight $g(x)$. It turns out that for $dm_\theta = g^{-\theta}dx$ with specific ranges of p and θ, we obtain an instability result for the BHI with respect to $L^p(dm_\theta)$.

Theorem 5.1.2. *Let $p > \theta > 0$ and $dm_\theta = g^{-\theta}dx$. There exists a sequence of nonnegative functions $\{h_k\}_{k\geq 1}$ in $L^p(dm_\theta)$ such that $\|h_k\|_2 = 1$, $\delta_{\mathrm{BH}}(h_k) \to 0$, $\|h_k\|_{L^p(dm_\theta)} \to \infty$, and*

$$\liminf_{k\to\infty} \frac{\mathrm{dist}_{L^p(dm_\theta)}(h_k, \mathfrak{G})}{\|h_k\|_{L^p(dm_\theta)}} \geq C(p,\theta) > 0.$$

We emphasize that $d\eta_\lambda$ is a more suitable reference measure than dm_θ in a sense that $L^p(d\eta_\lambda)$ contains all optimizers $\mathfrak{G}$ whereas $L^p(dm_\theta)$ does not (see (5.3.2)). If we choose the Lebesgue measure as a reference measure (that is, $\theta = 0$ in Theorem 5.1.1

or $\lambda = 0$ in Theorem 5.1.2), then the sequence of functions h_k converges to g in L^p (see Remark 5.3.4). It remains open to show L^p-stability for the BHI with respect to the Lebesgue measure.

5.2 Relation to stability of the Hausdorff–Young inequality

We briefly review the work of Christ [45] and discuss how it is related to the Beckner–Hirschman inequality. This consideration gives a glimpse of what stability of the Beckner–Hirschman inequality would be and the connection to our instability results of the BHI.

Let $p \in [1,2]$, $q = p/(p-1)$, and $\mathbf{A}_p = p^{1/2p} q^{-1/2q}$. For a complex-valued function $h \in L^p(\mathbb{R}^n)$, the sharp Hausdorff–Young inequality by Babenko [4] and Beckner [20] states that $\|\widehat{h}\|_q \le \mathbf{A}_p^n \|h\|_p$. Then Lieb [91] showed that equality holds if and only if a function h is of the form $h(x) = ce^{-Q(x)+x\cdot v}$ where $v \in \mathbb{C}^n$, $c \in \mathbb{C}$, and Q is a positive definite real quadratic form. Let $\mathscr{G}$ be the set of all optimizers for the Hausdorff–Young inequality. Define $\mathscr{P}(\mathbb{R}^n)$ to be the set of all polynomials $P : \mathbb{R}^n \to \mathbb{C}$ of the form $P(x) = -x \cdot Ax + b \cdot x + c$ where $b \in \mathbb{C}^n$, $c \in \mathbb{C}$, and A is a symmetric, positive definite real matrix. Note that $\mathscr{G} \setminus \{0\} = \{e^P : P \in \mathscr{P}(\mathbb{R}^n)\}$. Let $u \in \mathscr{G} \setminus \{0\}$. The real tangent space to $\mathscr{G}$ at u is $T_u\mathscr{G} = \{Pu : P \in \mathscr{P}(\mathbb{R}^n)\}$, and the normal space to $\mathscr{G}$ at u is

$$N_u\mathscr{G} = \{h \in L^p : \operatorname{Re}\Big(\int_{\mathbb{R}^n} hP\overline{u}|u|^{p-2}dx\Big) = 0\}. \tag{5.2.1}$$

Define $\operatorname{dist}_p(h,\mathscr{G}) = \inf_{u\in\mathscr{G}} \|h-u\|_p$. There exists $\delta_0 > 0$ such that if a nonzero function h satisfies $\operatorname{dist}_p(h,\mathscr{G}) \le \delta_0\|h\|_p$, then h can be written as $h = h^\perp + \pi(h)$ where $\pi(h) \in \mathscr{G}$ and $h^\perp \in N_{\pi(h)}\mathscr{G}$. Since $\|h^\perp\|_p = \|h - \pi(h)\|_p$ and $\pi(h) \in \mathscr{G}$, we have $\|h^\perp\|_p \ge \operatorname{dist}_p(h,\mathscr{G})$. For a function h satisfying $\operatorname{dist}_p(h,\mathscr{G}) \le \delta_0\|h\|_p$, we define $\operatorname{dist}_p^*(h,\mathscr{G}) = \|h^\perp\|_p$.

Let $p \in [1,2]$ and $h \in L^p(\mathbb{R}^n)$. The deficit of the Hausdorff–Young inequality is given by

$$\delta_{\mathrm{HY}}(h;p) = \mathbf{A}_p^n - \frac{\|\widehat{h}\|_q}{\|h\|_p}.$$

Let $\mathbf{B}_{p,n} = \frac{1}{2}(p-1)(2-p)\mathbf{A}_p^n$. For $\eta > 0$, we define

$$h_\eta^\perp = \begin{cases} h^\perp, & |h^\perp| \le \eta |\pi(h)|, \\ 0 & |h^\perp| > \eta |\pi(h)|. \end{cases}$$

In [45], Christ proved the following quantitative Hausdorff–Young inequality. He firstly showed a compactness result using combinatoric arguments, and then computed the second variation to obtain remainder terms for the Hausdorff–Young inequality.

Theorem 5.2.1 ([45, Theorem 1.3]). *For each $n \ge 1$ and $p \in (1,2)$, there exist $\eta_0, \gamma > 0$ and $C, c > 0$ such that for all $\eta \in (0, \eta_0)$, if a nonzero function $h \in L^p(\mathbb{R}^n)$ satisfies $\mathrm{dist}_p(h, \mathscr{G}) \le \eta^\gamma \|h\|_p$, then $\delta_{\mathrm{HY}}(h;p) \ge R_1(h;p) + R_2(h;p)$ where*

$$R_1(h;p) = (\mathbf{B}_{p,n} - C\eta)\|h\|_p^{-p}\Big(\int_{\mathbb{R}^n} |h_\eta^\perp|^2 |\pi(h)|^{p-2} dx\Big), \tag{5.2.2}$$

$$R_2(h;p) = c\eta^{2-p}\left(\frac{\mathrm{dist}_p(h,\mathscr{G})}{\|h\|_p}\right)^{p-2}\left(\frac{\|h^\perp - h_\eta^\perp\|_p}{\|h\|_p}\right)^2.$$

By differentiating the sharp Hausdorff–Young inequality, one can derive the BHI. Indeed, let $h \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ with $\|h\|_2 = 1$. Since $\delta_{\mathrm{HY}}(h;p) \ge 0$ and $\delta_{\mathrm{HY}}(h;2) = 0$, the derivative of $\delta_{\mathrm{HY}}(h;p)$ with respect to p at $p = 2$ is less than or equal to 0, which yields

$$-\frac{d}{dp}\delta_{\mathrm{HY}}(h,p)|_{p=2} = \frac{1}{4}\Big(S(|h|^2) + S(|\widehat{h}|^2) - n(1-\log 2)\Big) \ge 0.$$

A natural question is whether stability of the Hausdorff–Young inequality also yields that of the BHI. In what follows, we fix a function $h \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ that satisfies $\mathrm{dist}_p(h, \mathscr{G}) \le \delta_0 \|h\|_p$ and $\|h\|_2 = 1$ for all $p \in [1,2]$. Note that $h^\perp$ and $\pi(h)$ depend on p. We assume the following:

(i) We can choose a constant δ_0 to be uniform in $p \in [1,2]$.

(ii) The constant η in (5.2.2) is independent of $p \in (1,2)$.

(iii) We choose the constant $C = C(p)$ in (5.2.2) such that C is differentiable on $(1,2]$ and $C(2) = 0$.

(iv) $R_1(h;p) \geq 0$ for all $p \in (1,2)$.

(v) $h^\perp$ and $\pi(h)$ are differentiable with respect to p.

Based on these assumptions, we have $\delta_{\mathrm{HY}}(h;p) \geq R_1(h;p) + R_2(h;p) \geq R_1(h;p) \geq 0$ and $\delta_{\mathrm{HY}}(h;2) = R_1(h;2) = 0$. Taking the derivative with respect to p, we obtain

$$S(|h|^2) + S(|\widehat{h}|^2) - n(1-\log 2) = -4\frac{d}{dp}\Big(\mathbf{A}_p^n - \frac{\|\widehat{h}\|_q}{\|h\|_p}\Big)|_{p=2} \geq -4\frac{d}{dp}R_1(h;p)|_{p=2}$$

and

$$\begin{aligned}\frac{d}{dp}R_1(h;p)|_{p=2} &= \frac{d}{dp}(\mathbf{B}_{p,n} - C\eta)|_{p=2}\Big(\lim_{p\uparrow 2}\int_{\mathbb{R}^n}|h_\eta^\perp|^2|\pi(h)|^{p-2}dx\Big)\\ &= -(\frac{1}{2} + C'(2)\eta)\Big(\lim_{p\uparrow 2}\int_{\mathbb{R}^n}|h_\eta^\perp|^2|\pi(h)|^{p-2}dx\Big).\end{aligned}$$

Let h be a nonnegative function and $L_\eta = \{x : |h^\perp(x)| \leq \eta|\pi(h)(x)|\}$, then $h_\eta^\perp = h^\perp \cdot \mathbb{1}_{L_\eta}$. By Fatou's lemma, we get

$$\begin{aligned}\lim_{p\uparrow 2}\int_{\mathbb{R}}|h_\eta^\perp|^2|\pi(h)|^{p-2}dx &\geq \int_{\mathbb{R}}|h_\eta^\perp|^2dx\\ &= \int_{L_\eta}|h-\pi(h)|^2dx.\end{aligned}$$

Since $h - \pi(h) \in N_{\pi(h)}\mathscr{G}$, it follows from (5.2.1) that $\pi(h)$ is nonnegative with $\|\pi(h)\|_2 \leq 1$. Let

$$\widetilde{\mathfrak{G}} = \{u \in \mathscr{G} : u \geq 0, \|u\|_2 \leq 1\}.$$

Note that the set of the optimizers for the BHI defined in (5.1.1), $\mathfrak{G}$, is contained in $\widetilde{\mathfrak{G}}$ and $\pi(h) \in \widetilde{\mathfrak{G}}$. For η small such that $\frac{1}{2} + C'(2)\eta > 0$, we get

$$\delta_{\mathrm{BH}}(f) \geq C_\eta \mathrm{dist}_2(\widetilde{h}, \widetilde{\mathfrak{G}})^2$$

where $\mathrm{dist}_2(\widetilde{h}, \widetilde{\mathfrak{G}}) = \inf_{u\in\widetilde{\mathfrak{G}}}\|\widetilde{h} - u\|_2$ and

$$\widetilde{h}(x) = \begin{cases} h(x), & x \in L_\eta,\\ \pi(h)(x), & x \notin L_\eta.\end{cases}$$

Although we make strong assumptions, our observation suggests that there might be a stability estimate for the BHI in terms of L^2 or weaker distance than L^2 with

respect to the Lebesgue measure. We remark that Theorem 5.1.1 and 5.1.2 do not contradict to the observation. In Theorem 5.1.2, we show that the BHI is not stable in terms of $\mathrm{dist}_{L^p(dm_\theta)}(\cdot, \mathfrak{G})$ with normalization for $p > \theta > 0$. In Remark 5.3.4, we explain that our example constructed in Theorem 5.1.2 does not give any instability results for the BHI when $\theta = 0$. Note that $\mathrm{dist}_2(\cdot,\cdot)$ is the boundary case when $\theta = 0$ and $p = 2$. Compared to Theorem 5.1.1, $\mathrm{dist}_2(\cdot,\cdot)$ can be seen as the case when $\lambda = 0$ (so that $p \geq 2(\lambda+1) = 2$). Furthermore, Theorem 5.1.1 implies that an L^2-stability estimate would be best possible in terms of the L^p distances if exists.

5.3 Proofs of Theorem 5.1.1 and Theorem 5.1.2

5.3.1 Technical lemmas

To complete the proof of Theorem 5.1.2, we want to show that if k is large enough then

$$\mathrm{dist}_{L^p(dm_\theta)}(h_k, \mathfrak{G}) \geq C\|h_k\|_{L^p(dm_\theta)}$$

for some $C > 0$. Lemma 5.3.1 and Lemma 5.3.2 reduce the left hand side to the infimum of L^p norms over a finite interval when $p > 2$, which makes it easy to estimate a lower bound of the distance. To control the right hand side, we obtain a two-sided estimate of $\|h_k\|_{L^p(dm_\theta)}$ in Lemma 5.3.3.

Lemma 5.3.1. *Let $p > \theta > 0$, $a \geq a_0 > \pi$, $0 < t < (a_0/\pi)^{\frac{1}{4}}$, and $G_a(x) = G_{a,0}(x) = (\frac{2a}{\pi})^{\frac{1}{4}} e^{-ax^2}$. Let $M(a,t) = \{x : G_a(x) \geq tG_\pi(x)\}$, then there exist constants $C(p, a_0, t), C(p, \theta) > 0$ such that*

$$C(p, a_0, t) a^{\frac{p-2}{4p}} \leq \|G_a \cdot \mathbb{1}_{M(a,t)}\|_{L^p(dm_\theta)} \leq C(p,\theta) a^{\frac{p-2}{4p}}$$

for all $a \geq a_0$. In particular, if $p > 2$ then $\lim_{a\to\infty} \|G_a \cdot \mathbb{1}_{M(a,t)}\|_{L^p(dm_\theta)} = \infty$.

Proof. Since G_a is symmetric and decreasing in $[0,\infty)$, the level set $M_{a,t} = [-x_0, x_0]$ where $x_0 > 0$ satisfies $G_a(x_0) = tG_\pi(x_0)$. Solving the equation for x_0, we obtain

$$x_0 = \frac{1}{2}\sqrt{\frac{\log a - \log \pi - 4\log t}{a - \pi}}.$$

Let $\beta = ap - \theta\pi > 0$, then

$$\begin{aligned}\|G_a \cdot \mathbb{1}_{M(a,t)}\|^p_{L^p(dm_\theta)} &= \int_{-x_0}^{x_0} |G_a(x)|^p dm_\theta \\ &= \left(\frac{2a}{\pi}\right)^{\frac{p}{4}} \int_{-x_0}^{x_0} e^{-\beta x^2} dx \\ &= 2^{\frac{p}{4}} \pi^{-\frac{p-2}{4}} a^{\frac{p-2}{4}} (p - \frac{\theta\pi}{a})^{-\frac{1}{2}} (2\Phi(\sqrt{2\beta}x_0) - 1).\end{aligned}$$

Since $\sqrt{2\beta}x_0 \to \infty$ as $a \to \infty$, there exists a constant $C(a_0, t) > 0$ such that $C(a_0, t) \leq 2\Phi(\sqrt{2\beta}x_0) - 1 \leq 1$. We have

$$2^{\frac{1}{4}} \pi^{-\frac{p-2}{4p}} p^{-\frac{1}{2p}} C(a_0,t)^{\frac{1}{p}} a^{\frac{p-2}{4p}} \leq \|G_a \cdot \mathbb{1}_{M(a,t)}\|_{L^p(dm_\theta)} \leq 2^{\frac{1}{4}} \pi^{-\frac{p-2}{4p}} (p-\theta)^{-\frac{1}{2p}} a^{\frac{p-2}{4p}},$$

which completes the proof. □

Let f_k be the sequence of functions defined in Example 4.3.2 with $b = b_k = 2(k + \frac{1}{k}) + \sqrt{k}$ and $\alpha = \alpha_k = b_k^{-\frac{3}{2}}$. Recall that $b_k = 2(k + \frac{1}{k}) + \sqrt{k}$, $f_k(x) = f_k(-x)$, and

$$f_k(x) = \begin{cases} c_k, & x \in [0, k], \\ c_k l_k(x), & x \in (k, k + \frac{1}{k}], \\ c_k \alpha g_b(x), & x \in (k + \frac{1}{k}, \infty). \end{cases}$$

Here c_k is a normalization constant so that $\int_{\mathbb{R}} f_k d\gamma = 1$. Note that $\alpha_k \to 0$, $b_k \to \infty$, and $c_k \to 1$ as $k \to \infty$. Define $h_k(x) = \sqrt{f_k(2\sqrt{\pi}x)} g(x)$. It follows from change of variables that $\|h_k\|_2 = \|f_k\|_{L^1(d\gamma)} = 1$.

Lemma 5.3.2. *Let $p > 2$, $p > \theta > 0$, and h_k be defined as above. There exist $k_0 \in \mathbb{N}$ and $a_0 > \pi$ such that*

$$\|h_k - G_a\|_{L^p(dm_\theta)} \geq \|h_k - G_\pi\|_{L^p(dm_\theta)}$$

for all $a \geq a_0$ and $k \geq k_0$.

Proof. Let $\widetilde{G}_a(x) = G_a(\frac{x}{2\sqrt{\pi}})/G_\pi(\frac{x}{2\sqrt{\pi}})$, then

$$\|h_k - G_a\|^p_{L^p(dm_\theta)} = (4\pi)^{\frac{\beta-1}{2}} \int |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^\beta(x) dx$$

where $\gamma(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{|x|^2}{2}}$ and $\beta = \frac{p-\theta}{2}$. We choose $k_0 \in \mathbb{N}$ such that $\frac{1}{2} \le c_k \le \frac{3}{2}$ for all $k \ge k_0$. Since $l_k(x) \le 1$, we have $|\sqrt{c_k l_k(x)} - 1| \le 1$. Let $k \ge k_0$, then we get

$$\int |\sqrt{f_k(x)} - 1|^p \gamma^\beta(x)dx = \int_{-k}^{k} |\sqrt{c_k} - 1|^p \gamma^\beta(x)dx + 2\int_{k}^{k+\frac{1}{k}} |\sqrt{c_k l_k} - 1|^p \gamma^\beta(x)dx \tag{5.3.1}$$

$$\begin{aligned} &+ 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b} - 1|^p \gamma^\beta(x)dx \\ &\le 2^{-p}(2\pi)^{-\frac{\beta-1}{2}} \beta^{-\frac{1}{2}} (2\Phi(\sqrt{\beta}k) - 1) + \frac{2\gamma^\beta(k)}{k} \\ &+ 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^\beta(x)dx \\ &\le C_1(p,\theta) + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^\beta(x)dx. \end{aligned}$$

Choose $a_1 > \pi$ so that $\widetilde{G}_a(1) \le \frac{1}{2} \le \sqrt{c_k}$ for all $a \ge a_1$. Setting $A = \{x : \widetilde{G}_a(x) \ge \frac{3}{2}\}$, we see that $A \subseteq [-k, k]$ and

$$\begin{aligned} \int_{-k}^{k} |\sqrt{c_k} - \widetilde{G}_a(x)|^p \gamma^\beta(x)dx &\ge \int_{\widetilde{G_a}(x) \ge \frac{3}{2}} \left|\widetilde{G}_a(x) - \frac{3}{2}\right|^p \gamma^\beta(x)dx \\ &\ge 2^{1-p} \int_A |\widetilde{G}_a(x)|^p \gamma^\beta(x)dx - \left(\frac{3}{2}\right)^p (2\pi)^{-\frac{\beta-1}{2}} \beta^{-\frac{1}{2}} \end{aligned}$$

for all $a \ge a_1$. Let $B = \{x : \sqrt{c_k \alpha g_b(x)} \ge 1\}$. Note that $b = 2(k + \frac{1}{k}) + \sqrt{k}$, $\alpha_k \le \frac{1}{2}$, and $c_k \le \frac{3}{2}$. If $\sqrt{c_k \alpha g_b(x)} \ge 1$, then $x \ge \frac{b}{2} - \frac{1}{b}\log(c_k \alpha)$ and $B \subset [k + \frac{1}{k}, \infty)$. If $x \ge k + \frac{1}{k}$, then $\widetilde{G}_a(x) \le 1$; we have

$$\begin{aligned} \int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - \widetilde{G}_a(x)|^p \gamma^\beta(x)dx &\ge \int_B |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^\beta(x)dx \\ &\ge \int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^\beta(x)dx - \frac{1}{2}(2\pi)^{-\frac{\beta-1}{2}} \beta^{-\frac{1}{2}}. \end{aligned}$$

Combining our observation, we get

$$\begin{aligned} &\int |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^\beta(x)dx \\ &\quad \ge \int_{-k}^{k} |\sqrt{c_k} - \widetilde{G}_a(x)|^p \gamma^\beta(x)dx + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - \widetilde{G}_a(x)|^p \gamma^\beta(x)dx \\ &\quad \ge 2^{1-p} \int_A |\widetilde{G}_a(x)|^p \gamma^\beta(x)dx + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^\beta(x)dx - C_2(p,\theta). \end{aligned}$$

By Lemma 5.3.1, one can choose $a_0 \geq a_1$ such that

$$\int_A |\widetilde{G}_a(x)|^p \gamma^\beta(x) dx \geq 2^{p-1}(C_1(p,\theta) + C_2(p,\theta))$$

for all $a \geq a_0$. By (5.3.1), we have

$$\begin{aligned}\int |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^\beta(x) dx &\geq 2^{1-p} \int_A |\widetilde{G}_a(x)|^p \gamma^\beta(x) dx + \int |\sqrt{f_k(x)} - 1|^p \gamma^\beta(x) dx \\ &\quad - C_1(p,\theta) - C_2(p,\theta) \\ &\geq \int |\sqrt{f_k(x)} - 1|^p \gamma^\beta(x) dx,\end{aligned}$$

which finishes the proof. □

Lemma 5.3.3. *Let $p > \theta > 0$ and h_k be defined as above. There exists $k_0 \in \mathbb{N}$ such that*

$$\|h_k\|_{L^p(dm_\theta)} \asymp_{p,\theta} b^{-\frac{3}{4}} e^{\frac{\theta b^2}{4(p-\theta)}}$$

for all $k \geq k_0$.

Proof. Let $\beta = \frac{p-\theta}{2}$. A direct computation yields that

$$\begin{aligned}\|h_k\|^p_{L^p(dm_\theta)} &= (4\pi)^{\frac{\beta-1}{2}} \int |f_k(x)|^{\frac{p}{2}} \gamma^\beta(x) dx \\ &= |c_k|^{\frac{p}{2}} 2^{\frac{\beta-1}{2}} \beta^{-\frac{1}{2}} (2\Phi(\sqrt{\beta}k) - 1) + 2|c_k|^{\frac{p}{2}} \int_k^{k+\frac{1}{k}} |l_k(x)|^{\frac{p}{2}} \gamma^\beta(x) dx \\ &\quad + 2^{\frac{\beta+1}{2}} |c_k \alpha|^{\frac{p}{2}} \beta^{-\frac{1}{2}} e^{\frac{p\theta b^2}{4(p-\theta)}} \Phi(\frac{pb}{2\sqrt{\beta}} - \sqrt{\beta}(k + \frac{1}{k})).\end{aligned}$$

Choose $k_1 \in \mathbb{N}$ such that $c_k \in [\frac{1}{2}, \frac{3}{2}]$ and $\Phi(\frac{pb}{2\sqrt{\beta}} - \sqrt{\beta}(k+\frac{1}{k})) \geq \frac{1}{2}$ for all $k \geq k_1$. Then we have

$$\|h_k\|_{L^p(dm_\theta)} \geq C(p,\theta) b^{-\frac{3}{4}} e^{\frac{\theta b^2}{4(p-\theta)}}.$$

Since we have

$$|c_k|^{\frac{p}{2}} 2^{\frac{\beta-1}{2}} \beta^{-\frac{1}{2}} (2\Phi(\sqrt{\beta}k) - 1) + 2|c_k|^{\frac{p}{2}} \int_k^{k+\frac{1}{k}} |l_k(x)|^{\frac{p}{2}} \gamma^\beta(x) dx \leq C(p,\theta),$$

we can choose $k_2 \in \mathbb{N}$ such that

$$\|h_k\|_{L^p(dm_\theta)} \leq C(p,\theta) b^{-\frac{3}{4}} e^{\frac{\theta b^2}{4(p-\theta)}}$$

for all $k \geq k_2$. □

5.3.2 Proof of Theorem 5.1.2

Let f_k be the sequence of functions defined in Example 4.3.2 with $b = b_k = 2(k+\frac{1}{k}) + \sqrt{k}$ and $\alpha = \alpha_k = b_k^{-\frac{3}{2}}$. Define $h_k(x) = \sqrt{f_k(2\sqrt{\pi}x)}g(x)$. Note that $\|h_k\|_{L^2(dm_\theta)} = \|f_k\|_{L^1(d\gamma)} = 1$. By (2.2.5) and Lemma 4.3.3, we have $\delta_{\mathrm{BH}}(h_k) \to 0$ as $k \to \infty$. Since the function h_k and $g^{-\theta}$ are symmetric and the symmetric decreasing rearrangement of $G_{a,r}$ is G_a, it follows from the rearrangement inequality (see [92, Theorem 3.5]) that

$$\mathrm{dist}_{L^p(dm_\theta)}(h_k, \mathfrak{G}) = \inf_{a\in(\frac{\theta\pi}{p},\infty)} \|h_k - G_a\|_{L^p(dm_\theta)}$$

for all $k \geq 1$. Here we used the fact that

$$G_{a,r} \in L^p(dm_\theta) \text{ if and only if } a > \theta\pi/p. \tag{5.3.2}$$

Our goal is to show that there exists a constant $C = C(p,\theta) > 0$ such that

$$\|h_k - G_a\|_{L^p(dm_\theta)} \geq C\|h_k\|_{L^p(dm_\theta)}$$

for all $a \in (\frac{\theta\pi}{p}, \infty)$ and for large k.

Case 1: $a \geq \pi$

Suppose $p > 2$. By Lemma 5.3.2, there exists $a_0 > \pi$ such that

$$\mathrm{dist}_{L^p(dm_\theta)}(h_k, \mathfrak{G}) = \inf_{a\in(\frac{\theta\pi}{p},a_0]} \|h_k - G_a\|_{L^p(dm_\theta)}$$

for all large k. So it suffices to show that if k is large enough, then $\|h_k - G_a\|_{L^p(dm_\theta)} \geq C\|h_k\|_{L^p(dm_\theta)}$ for all $a \in (\pi, a_0]$. First we consider the case when $\pi \leq a \leq a_0$. Since $p > 2$,

$$\begin{aligned} \|G_a\|^p_{L^p(dm_\theta)} &= 2^{\frac{p-\theta}{4}}(\frac{a}{\pi})^{\frac{p-2}{4}}(p - \frac{\theta\pi}{a})^{-\frac{1}{2}} \\ &= C(p,\theta)a^{\frac{p-2}{4}}(p - \frac{\theta\pi}{a})^{-\frac{1}{2}} \end{aligned} \tag{5.3.3}$$

is uniformly bounded in $a \in [\pi, a_0]$. By Lemma 5.3.3, we can choose $k_1 \in \mathbb{N}$ so that for all $k \geq k_1$, $\|h_k\|_{L^p(dm_\theta)} \geq 2 \sup_{a\in[\pi,a_0]} \|G_a\|_{L^p(dm_\theta)}$. We obtain

$$\begin{aligned}\|h_k - G_a\|_{L^p(dm_\theta)} &\geq \|h_k\|_{L^p(dm_\theta)} - \sup_{a\in[\pi,a_0]} \|G_a\|_{L^p(dm_\theta)} \\ &\geq \frac{1}{2}\|h_k\|_{L^p(dm_\theta)}\end{aligned}$$

for all $a \in [\pi, a_0]$ and $k \geq k_1$.

If $p \leq 2$, then it follows from (5.3.3) that $\|G_a\|^p_{L^p(dm_\theta)} \leq C(p,\theta)\pi^{\frac{p-2}{4}}(p-\theta)^{-\frac{1}{2}}$ for all $a \geq \pi$. By Lemma 5.3.3, we choose $k_2 \in \mathbb{N}$ such that $\|h_k - G_a\|_{L^p(dm_\theta)} \geq \frac{1}{2}\|h_k\|_{L^p(dm_\theta)}$ for all $k \geq k_2$.

Case 2: $\frac{\theta\pi}{p} < a < \pi$

By Lemma 5.3.3, it suffices to show that there exists a constant $c > 0$ such that

$$\|h_k - G_a\|_{L^p(dm_\theta)} \geq c b^{-\frac{3}{4}} e^{\frac{\theta b^2}{4(p-\theta)}}$$

for all $a \in (\frac{\theta\pi}{p}, \pi)$ and large k. Let $\beta = \frac{p-\theta}{2}$ and $s = 1 - \frac{a}{\pi}$, then $0 < s < 1 - \frac{\theta}{p}$. We define $R_{s,k}(x) = \widetilde{G}_a(x)/\sqrt{f_k(x)}$, then

$$\begin{aligned}\|h_k - G_a\|_{L^p(dm_\theta)} &= (4\pi)^{\frac{\beta-1}{2}} \int |\sqrt{f_k} - \widetilde{G}_a|^p \gamma^\beta dx \\ &= (4\pi)^{\frac{\beta-1}{2}} \int |1 - R_{s,k}|^p |f_k|^{\frac{p}{2}} \gamma^\beta dx \\ &\geq (4\pi)^{\frac{\beta-1}{2}} |c_k\alpha|^{\frac{p}{2}} \int_{k+\frac{1}{k}}^{\infty} |1 - R_{s,k}|^p e^{\frac{pb}{2}x - \frac{pb^2}{4}} \gamma^\beta dx.\end{aligned}$$

Let $Q_{s,k}(x) = \frac{s}{4}(x - \frac{b}{s})^2 - (\frac{1-s}{4s})b^2$, then

$$R_{s,k}(x) = \frac{(1-s)^{\frac{1}{4}}}{(c_k\alpha)^{\frac{1}{2}}} e^{Q_{s,k}(x)}.$$

Choose $t \in (1, \frac{p}{p-\theta})$, then

$$Q_{s,k}(tb) = \frac{s}{4}(tb - \frac{b}{s})^2 - (\frac{1-s}{4s})b^2 = \frac{t^2b^2}{4}\Big(s - (\frac{2t-1}{t^2})\Big).$$

Since the map $t \mapsto \frac{2t-1}{t^2}$ is decreasing on $(1, \frac{p}{p-\theta})$, we know

$$\frac{2t-1}{t^2} \geq \frac{2(\frac{p}{p-\theta})-1}{(\frac{p}{p-\theta})^2} = \frac{p^2-\theta^2}{p^2} > \frac{p-\theta}{p}.$$

Since $s \in (0, \frac{p-\theta}{p})$, we have $Q_{s,k}(tb) < 0$. The function $Q_{s,k}(x)$ is symmetric about $x = \frac{b}{s}$ and $\frac{b}{s} > tb$. This yields that $Q_{s,k}(x) \leq Q_{s,k}(tb)$ for all $x \in [tb, \frac{2b}{s} - tb]$. Thus we can choose $k_3 \in \mathbb{N}$ so that $R_{s,k}(x) \leq \frac{1}{2}$ for all $k \geq k_3$ and $s \in (0, \frac{p-\theta}{p})$. Since $(t - \frac{p}{p-\theta}) < 0$ and $(\frac{2}{s} - t - \frac{p}{p-\theta}) \geq c > 0$ uniformly in s, we can choose $k_4 \in \mathbb{N}$ so that

$$\Phi((\frac{2}{s} - t - \frac{p}{p-\theta})b\sqrt{\beta}) - \Phi((t - \frac{p}{p-\theta})b\sqrt{\beta}) \geq \frac{1}{2}$$

for all $k \geq k_4$ and $s \in (0, \frac{p-\theta}{p})$. If k is large enough, then we obtain

$$\begin{aligned}
\|h_k - G_a\|^p_{L^p(dm_\theta)} &\geq (4\pi)^{\frac{\beta-1}{2}} 2^{-p} |c_k \alpha|^{\frac{p}{2}} \int_{k+\frac{1}{k}}^{\infty} e^{\frac{pb}{2}x - \frac{pb^2}{4}} \gamma^\beta dx \\
&\geq 2^{\frac{\beta-1}{2}-p} |c_k\alpha|^{\frac{p}{2}} e^{\frac{p\theta b^2}{4(p-\theta)}} \beta^{-\frac{1}{2}} (\Phi((\frac{2}{s} - t - \frac{p}{2\beta})b\sqrt{\beta}) - \Phi((t - \frac{p}{2\beta})b\sqrt{\beta})) \\
&\geq C(p,\theta) b^{-\frac{3p}{4}} e^{\frac{p\theta b^2}{4(p-\theta)}}.
\end{aligned}$$

By Lemma 5.3.3, we have

$$\|h_k - G_a\|_{L^p(dm_\theta)} \geq C \|h_k\|_{L^p(dm_\theta)}$$

for all $a \in (\frac{\theta\pi}{p}, \pi)$, which completes the proof. □

5.3.3 Proof of Theorem 5.1.1

We note that $G_{a,r} \in L^p(d\eta_\lambda)$ for all $a > 0$ and $r \in \mathbb{R}$. Indeed we have

$$\begin{aligned}
\|G_{a,r}\|^p_{L^p(d\eta_\lambda)} &= \int |G_{a,r}(x)|^p d\eta_\lambda \\
&\leq \int |G_a(x)|^p d\eta_\lambda \\
&= \Big(\frac{2a}{\pi}\Big)^{\frac{p}{4}} \int |x|^\lambda e^{-apx^2} dx \\
&= \Big(\frac{2a}{\pi}\Big)^{\frac{p}{4}} (2ap)^{-\frac{\lambda+1}{2}} \int |x|^\lambda e^{-\frac{x^2}{2}} dx \\
&= C(p,\lambda) a^{\frac{p-2\lambda-2}{4}} m_\lambda(\gamma)
\end{aligned} \tag{5.3.4}$$

where $m_\lambda(\gamma)$ is the λ-th moment of the standard Gaussian measure. Let $h_k(x) = \sqrt{f_k(2\sqrt{\pi}x)}g(x)$ with $\alpha_k = b_k^{-w}$ and $0 < w < \frac{2\lambda}{p}$, then

$$\begin{aligned}
\|h_k\|_{L^p(d\eta_\lambda)}^p &= C(p,\lambda)\int |f_k(x)|^{\frac{p}{2}}\gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\
&\geq C(p,\lambda)|c_k\alpha_k|^{\frac{p}{2}}\int_{k+\frac{1}{k}}^{\infty} |x|^\lambda e^{-\frac{p}{4}(x-b_k)^2}dx \\
&= C(p,\lambda)|c_k\alpha_k|^{\frac{p}{2}}\int_{k+\frac{1}{k}-b_k}^{\infty} |x+b_k|^\lambda e^{-\frac{p}{4}x^2}dx \\
&\geq C(p,\lambda)|c_k\alpha_k|^{\frac{p}{2}}(|b_k|^\lambda - m_\lambda(\gamma))
\end{aligned}$$

so that $\|h_k\|_{L^p(d\eta_\lambda)} \to \infty$ as $k \to \infty$. By the rearrangement inequality,

$$\operatorname{dist}_{L^p(d\eta_\lambda)}(h_k, \mathfrak{G}) = \inf_{a>0} \|h_k - G_a\|_{L^p(d\eta_\lambda)}.$$

Assume $p = 2\lambda + 2$, then $\|G_a\|_{L^p(d\eta_\lambda)} = C(p,\lambda)m_\lambda(\gamma)$ is independent of a. We pick $k_1 \in \mathbb{N}$ such that $\|h_k\|_{L^p(d\eta_\lambda)} \geq 2\|G_a\|_{L^p(d\eta_\lambda)}$ for all $k \geq k_1$, then

$$\|h_k - G_a\|_{L^p(d\eta_\lambda)} \geq \|h_k\|_{L^p(d\eta_\lambda)} - \|G_a\|_{L^p(d\eta_\lambda)} \geq \frac{1}{2}\|h_k\|_{L^p(d\eta_\lambda)}$$

for all $k \geq k_1$, as desired.

Suppose $p - 2\lambda - 2 > 0$. By (5.3.4), we have $\|G_a\|_{L^p(d\eta_\lambda)} \to \infty$ as $a \to \infty$. Since $\|h_k\|_{L^p(d\eta_\lambda)} \to \infty$ and $\|G_a\|_{L^p(d\eta_\lambda)}$ is bounded in $a \in (0, a_0]$ for a fixed a_0 by (5.3.4), it suffices to show that there exist k_0 and a_0 such that

$$\|h_k - G_a\|_{L^p(d\eta_\lambda)} \geq \|h_k - G_\pi\|_{L^p(d\eta_\lambda)}$$

for all $k \geq k_0$ and $a \geq a_0$. Let $\widetilde{G}_a(x) = G_a(\frac{x}{2\sqrt{\pi}})/G_\pi(\frac{x}{2\sqrt{\pi}})$, then

$$\|h_k - G_a\|_{L^p(d\eta_\lambda)}^p = C(p,\lambda)\int |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p\gamma^{\frac{p}{2}}(x)|x|^\lambda dx.$$

We choose $k_1 \in \mathbb{N}$ such that $\frac{1}{2} \leq c_k \leq \frac{3}{2}$ for all $k \geq k_1$. Let $I = [-x_0, x_0]$ with

$$x_0 = \frac{1}{2}\sqrt{\frac{\log a - \log\pi - 4\log(3/2)}{a-\pi}},$$

then $\widetilde{G}_a(x) \geq 3/2$ for all $x \in I$. Choose $a_1 > \pi$ so that $\widetilde{G}_a(1) \leq \frac{1}{2} \leq \sqrt{c_k}$ for all $a \geq a_1$, then $I \subset [-k, k]$. We get

$$\begin{aligned}\int_{-k}^{k} |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx &\geq \int_I |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &\geq C(p,\lambda) a^{\frac{p-2(\lambda+1)}{4}} \int_{-\sqrt{\frac{ap}{2\pi}}x_0}^{\sqrt{\frac{ap}{2\pi}}x_0} |x|^\lambda d\gamma - C_1(p,\lambda).\end{aligned}$$

Since $\sqrt{a}x_0 \to \infty$ as $a \to \infty$, there exist a_2 and $C > 0$ such that $\int_I |x|^\lambda d\gamma \geq C$ for all $a \geq a_2$. Let $B = \{x : \sqrt{c_k \alpha_k g_{b_k}(x)} \geq 1\}$. Note that $b_k = 2(k + \frac{1}{k}) + \sqrt{k}$, $\alpha_k \leq \frac{1}{2}$, and $c_k \leq \frac{3}{2}$. If $\sqrt{c_k \alpha g_b(x)} \geq 1$, then $x \geq \frac{b}{2} - \frac{1}{b}\log(c_k\alpha)$ and $B \subset [k + \frac{1}{k}, \infty)$. If $x \geq k + \frac{1}{k}$, then $\widetilde{G}_a(x) \leq 1$; thus we have

$$\begin{aligned}\int_{k+\frac{1}{k}}^{\infty} |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx &\geq \int_B |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &\geq \int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx - C(p,\lambda).\end{aligned}$$

Combining our observation, we get

$$\begin{aligned}&\int |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &\quad \geq \int_{-k}^{k} |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &\quad \geq C_1(p,\lambda) a^{\frac{p-2(\lambda+1)}{4}} + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b(x)} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx - C_2(p,\lambda).\end{aligned}$$

We choose k_2 large enough so that for all $k \geq k_2$, we have

$$\begin{aligned}\int |\sqrt{f_k(x)} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx &= 2\int_0^k |\sqrt{c_k} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &\quad + 2\int_k^{k+\frac{1}{k}} |\sqrt{c_k l_k} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &\quad + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &= C_3(p,\lambda) + 2\int_{k+\frac{1}{k}}^{\infty} |\sqrt{c_k \alpha g_b} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx.\end{aligned}$$

It then follows that

$$\begin{aligned}&\int |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \\ &\quad \geq C_1(p,\lambda) a^{\frac{p-2(\lambda+1)}{4}} + \int |\sqrt{f_k(x)} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx - C_2(p,\lambda) - C_3(p,\lambda).\end{aligned}$$

Letting a large enough, we obtain

$$\int |\sqrt{f_k(x)} - \widetilde{G}_a(x)|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx \geq \int |\sqrt{f_k(x)} - 1|^p \gamma^{\frac{p}{2}}(x)|x|^\lambda dx.$$

Therefore, we have $\|h_k - G_a\|_{L^p(d\eta_\lambda)} \geq \|h_k - G_\pi\|_{L^p(d\eta_\lambda)}$ as desired. □

Remark 5.3.4. For the Lebesgue measure and $p \geq 0$, we have

$$\begin{aligned}\|h_k - G_\pi\|_p^p &= (4\pi)^{\frac{p-2}{4}} \int |\sqrt{f_k} - 1|^p \gamma^{\frac{p}{2}}(x)dx \\ &= o(1) + 2(c_k\alpha)^{\frac{p}{2}} \int_{k+\frac{1}{k}}^{\infty} |\sqrt{g_b(x)} - 1|^p \gamma^{\frac{p}{2}}(x)dx\end{aligned}$$

and

$$\int_{k+\frac{1}{k}}^{\infty} |\sqrt{g_b(x)} - 1|^p \gamma^{\frac{p}{2}}(x)dx \leq 2^p \int_{k+\frac{1}{k}}^{\infty} g_b(x)^{\frac{p}{2}} \gamma^{\frac{p}{2}}(x)dx + 2^p \int_{k+\frac{1}{k}}^{\infty} \gamma^{\frac{p}{2}}(x)dx \leq C(p).$$

So we get

$$\lim_{k\to\infty} \mathrm{dist}_{L^p(dx)}(h_k, \mathfrak{G}) \leq \lim_{k\to\infty} \|h_k - G_\pi\|_p = 0,$$

which implies that our method does not give an instability result for the BHI when $\theta = 0$ in Theorem 5.1.2 and $\lambda = 0$ in Theorem 5.1.1.

6. STABILITY OF THE EXPECTED LIFETIME INEQUALITY

The isoperimetric inequalities for the expected lifetime of Brownian motion state that the L^p-norms of the expected lifetime in a bounded domain for $1 \leq p \leq \infty$ are maximized when the region is a ball with the same volume. In this chapter, we prove quantitative improvements of the inequalities. We also discuss related open problems that arise from these improvements. This chapter is based on my work [86].

6.1 Introduction

Let $\alpha \in (0, 2]$ and D a bounded domain in $\mathbb{R}^n$. Let X_t^α be the rotationally symmetric α–stable process with generator $-(-\Delta)^{\alpha/2}$. The first exit time of X_t^α from D is given by

$$\tau_D^\alpha = \inf\{t > 0 : X_t^\alpha \notin D\}.$$

The expected lifetime of X_t^α is denoted by $u_D^\alpha(x) = \mathbb{E}^x[\tau_D^\alpha]$ where $\mathbb{E}^x$ is the expectation associated with X_t^α starting at $x \in \mathbb{R}^n$. Note that $u_D^\alpha(x)$ is a solution to the equation

$$\begin{cases} (-\Delta)^{\frac{\alpha}{2}} u(x) = 1, & x \in D, \\ u(x) = 0, & x \notin D. \end{cases}$$

If B is a ball of radius R and centered at the origin, then $u_B^\alpha(x)$ is explicitly given by

$$u_B^\alpha(x) = C_{n,\alpha}(R^2 - |x|^2)^{\frac{\alpha}{2}}.$$

For $\alpha = 2$, X_t^α is Brownian motion with generator Δ. In this case, we drop the superscript α.

Bañuelos and Méndez-Hernández [16] showed that several isoperimetric type inequalities for Brownian motion continue to hold for a wide class of Lévy processes

using the symmetrization of Lévy processes and the multiple integral rearrangement inequalities of Brascamp–Lieb–Luttinger [31]. A particular case of this is that for all $t \geq 0$ and $x \in \mathbb{R}^n$,

$$\mathbb{P}^0(\tau_B^\alpha > t) \geq \mathbb{P}^x(\tau_D^\alpha > t), \tag{6.1.1}$$

which yields in turn that

$$u_B^\alpha(0) \geq u_D^\alpha(x), \tag{6.1.2}$$

where B is a ball centered at 0 with $|B| = |D|$. In fact, (6.1.1) gives

$$\mathbb{E}^0(\tau_B^\alpha)^p \geq \mathbb{E}^x(\tau_D^\alpha)^p \tag{6.1.3}$$

for all $p > 0$.

Talenti [110] proved that the L^p norm of a solution of a second-order elliptic equation is maximized when the elliptic operator and the domain are symmetrically rearranged. In particular, the result yields that for $p > 0$, $\alpha = 2$, and a bounded domain D,

$$\|u_B\|_p \geq \|u_D\|_p \tag{6.1.4}$$

where B is a ball with $|B| = |D|$.

Given the above isoperimetric type inequalities for the first exit time of the stable processes and their connection to the classical torsion function, it is interesting to find quantitative versions of these inequalities: for example, quantitative versions of (6.1.1) and (6.1.2), and their implications to quantitative versions of the torsional rigidity inequality, not only for the stable processes but even for the more general Lévy processes studied in [16]. The goal of this chapter is to study quantitative versions of the expected lifetime inequalities (6.1.2) for $\alpha = 2$ and (6.1.4) for $p \geq 1$.

6.2 Main results

We define the deficit of (6.1.2) by

$$\delta(x, D) = 1 - \frac{u_D(x)}{u_B(0)} \geq 0 \tag{6.2.1}$$

where B is a ball centered at 0 with $|B| = |D|$. We provide a lower bound of the deficit $\delta(x, D)$ in terms of the deviations of x and D from the optimizers. Note that equality holds in (6.2.1) if D is a ball and $u_D(x) = \max_{y\in D} u_D(y)$. The deviation of x is represented by the level set $|\{y \in D : u_D(y) > u_D(x)\}|$, and the deviation of D by the Fraenkel asymmetry, which is defined by

$$A(D) = \inf\left\{\frac{|D\triangle B|}{|D|} : B \text{ is a ball with } |B| = |D|\right\}. \tag{6.2.2}$$

Theorem 6.2.1. *Let $D \subseteq \mathbb{R}^n$ be a bounded domain with $A(D) > 0$. Let $D_t = \{y \in D : u_D(y) > t\}$, $\mu(t) = |D_t|$, and*

$$t_* = t_*(D) = \sup\{t > 0 : \mu(t) > |D|(1 - \frac{1}{4}A(D))\}. \tag{6.2.3}$$

Then we have

$$\delta(x, D) \geq |D|^{-\frac{2}{n}}\Big(\mu(u_D(x))^{\frac{2}{n}} + \mathsf{C}_n(u_D(x) \wedge t_*)A(D)^2\Big), \tag{6.2.4}$$

where $\mathsf{C}_n = \beta_n\omega_n^{\frac{1}{n}}$, β_n is a dimensional constant in (6.3.3), and ω_n is the volume of a unit ball in $\mathbb{R}^n$.

The proof is based on the proof of (6.1.2) for $\alpha = 2$ in [6, 110], and the sharp quantitative isoperimetric inequality [64]. In order to estimate the asymmetry of the level sets, we use the idea of Hansen and Nadirashvili [76] as in the proof of the boosted Pólya–Szegö inequality [33, Lemma 2.9].

Remark 6.2.2. We note that (6.2.4) with the first remainder term follows from the pointwise estimate $u_B(x) \geq (u_D)^*(x)$ of [109]. For simplicity, we assume that $|D| = 1$. For each $x \in D$, we define $r(x) \geq 0$ by $\mu(u_D(x)) = |B_{r(x)}|$ where $B_{r(x)}$ is a ball of radius $r(x)$. For a nonnegative measurable function f on D, the symmetric decreasing rearrangement $f^*(x) = f^*(|x|)$ satisfies $f^*(r(x)) \geq f(x)$ for each $x \in D$. Since u_B is rotationally symmetric, we use the notation $u_B(x) = u_B(|x|)$. Using $u_B(x) \geq (u_D)^*(x)$, one has

$$u_D(x) \leq (u_D)^*(r(x)) \leq u_B(r(x)) = u_B(0)(1 - (\omega_n^{\frac{1}{n}}r(x))^2) = u_B(0)(1 - \mu(u_D(x))^{\frac{2}{n}}).$$

Notice that (6.2.4) can be written as $u_B(r(x)) - u_D(x) \geq \mathsf{C}_n(u_D(x) \wedge t_*)A(D)^2$.

Remark 6.2.3. Note that if $A(D) > 0$, then $t_* > 0$. Suppose $\delta(x, D) = 0$. If $A(D) > 0$, then (6.2.4) implies $\mu(u_D(x)) = 0$ and $u_D(x) = 0$. This contradicts to the assumption $|D| > 0$ and thus D is a ball with $|B| = |D|$. As a consequence, one sees that equality holds (6.2.1) *only if* D is a ball and $u_D(x) = \max_{y\in D} u_D(y)$.

Remark 6.2.4. One can extend the result to an uniformly elliptic operator as in [110]. Let $\mathcal{L} = \partial_i(a_{ij}(x)\partial_j)$ where $a_{ij}(x)$ is a bounded measurable function with

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \sum_{i=1}^{n} \xi_i^2 \tag{6.2.5}$$

for each $x \in \mathbb{R}^n$ and $\xi = (\xi_1, \cdots, \xi_n) \in \mathbb{R}^n$. Consider a weak solution $u_D^{\mathcal{L}}$ of

$$\begin{cases} -\mathcal{L}u(x) = 1, & x \in D, \\ u(x) = 0, & x \in \partial D. \end{cases}$$

Following the proof of Theorem 6.2.1 and modifying (6.3.6) with inequality, which follows from the elliptic condition (6.2.5), one obtains

$$1 - \frac{u_D^{\mathcal{L}}(x)}{u_B(0)} \geq |D|^{-\frac{2}{n}} \Big(\mu(u_D^{\mathcal{L}}(x))^{\frac{2}{n}} + C_n(u_D^{\mathcal{L}}(x) \wedge t_*)A(D)^2 \Big).$$

The second result is a quantitative inequality for the L^p norm of the expected lifetime, $1 \leq p \leq \infty$. We define the L^p deficit of the expected lifetime inequality (6.1.4) for $1 \leq p \leq \infty$ by

$$\delta_p(D) = 1 - \left(\frac{\|u_D\|_p}{\|u_B\|_p} \right)^{\kappa(p)}$$

where $\kappa(p) = p$ for $1 \leq p < \infty$, $\kappa(\infty) = 1$, and B is a ball centered at 0 with $|B| = |D|$.

Theorem 6.2.5. *Let $n \geq 2$ and D be a bounded domain in $\mathbb{R}^n$. For $1 \leq p \leq \infty$, we have*

$$\delta_p(D) \geq \mathsf{C}_{n,p} A(D)^{2+\kappa(p)} \tag{6.2.6}$$

where $\mathsf{C}_{n,p}$ is explicitly given in (6.3.12) and (6.3.13). In particular, if $p = 1$, we have

$$T(B) - T(D) \geq \mathsf{C}_{n,1} T(B) A(D)^3. \tag{6.2.7}$$

Remark 6.2.6. Let $n = 2$ and $\varepsilon > 0$. Consider an ellipse $D = \{(x, y) \in \mathbb{R}^2 : x = \cos t, y = (1+\varepsilon)\sin t, t \in \mathbb{R}\}$. The asymmetry of D is $A(D) = \frac{1}{\pi}\varepsilon + O(\varepsilon^2)$ (see [75, pp. 88–89]). Note that the torsion function of D is

$$u_D(x) = \frac{(1+\varepsilon)^2}{2(1+(1+\varepsilon)^2)}\Big(1 - x^2 - \frac{y^2}{(1+\varepsilon)^2}\Big).$$

Let B be a ball with $|B| = |D| = (1+\varepsilon)\pi$. Let $p \in [1, \infty)$. Direct computations yield

$$\begin{aligned}
\|u_B\|_p^p - \|u_D\|_p^p &= \frac{\pi}{2^{2p}(p+1)}(1+\varepsilon)^{p+1} - \frac{\pi}{2^p(p+1)(1+(1+\varepsilon)^2)^p}(1+\varepsilon)^{2p+1} \\
&= \frac{\pi}{2^{2p}(p+1)}(1+\varepsilon)^{p+1}\Big(1 - \Big(1 - \frac{\varepsilon^2}{1+(1+\varepsilon)^2}\Big)^p\Big) \\
&= C_p\varepsilon^2 + o(\varepsilon^2)
\end{aligned}$$

for some $C_p > 0$, and

$$\delta_\infty(D) = 1 - \frac{\|u_D\|_\infty}{\|u_B\|_\infty} = 1 - \frac{2(1+\varepsilon)}{1+(1+\varepsilon)^2} = \frac{\varepsilon^2}{1+(1+\varepsilon)^2}$$

for $p = \infty$. This implies that the exponent of $A(D)$ in (6.2.6) cannot be replaced by smaller number than 2. It is open to show the inequality (6.2.6) with power 2.

Brasco, De Philippis, and Velichkov [34] showed that the sharp exponent of (6.2.7) is 2 in the sense that the power cannot be replaced by any smaller number. Their method, however, does not give an explicit dimensional constant because the proof relies on the selection principle of Cicalese and Leonardi [49].

The key step in the proof of Theorem 6.2.5 is the removal of t_* defined in (6.2.3). In [33], the authors proved the non-sharp quantitative Saint-Venant inequality (6.2.7) using transfer of asymmetry (Lemma 6.3.1) and the boosted Pólya–Szegö inequality. Thus t_* also appears in their proof. To replace t_* by $A(D)$ (up to a dimensional constant), they made use of the variational representation for $T(D)$ (2.5.2). In our case, however, the L^p norm of the expected lifetime does not have an appropriate variational formula for $1 < p \le \infty$. Instead, we estimate the distribution function of u_D when t_* is sufficiently small, and apply the layer cake representation and the strong Markov property. It turns out that this enables us to replace t_* by $A(D)$.

The fractional analogue of (6.2.7) is proven in [32]. Brasco, Cinti, and Vita showed that if $n \geq 2$, $\alpha \in (0,2)$, and D is an open set with $|D| = 1$, then

$$T_\alpha(B) - T_\alpha(D) \geq C_{n,\alpha} A(D)^{\frac{6}{\alpha}}$$

where $C_{n,\alpha}$ is explicit and B is a ball with $|B| = 1$. Furthermore, if D has Lipschitz boundary and satisfies the exterior ball condition, then the exponent can be lowered to $2 + \frac{2}{\alpha}$. It turns out that our method for removing t_* yields the same exponent without any additional geometric assumptions on D.

Theorem 6.2.7. *If $n \geq 2$, $\alpha \in (0,2)$, and D is an open set with $|D| = 1$, then*

$$T_\alpha(B) - T_\alpha(D) \geq C_{n,\alpha} A(D)^{2+\frac{2}{\alpha}}$$

where B is a ball with $|B| = |D|$.

6.3 Proofs of the main results

6.3.1 Transfer of asymmetry

The following lemma is essentially from [76, Lemma 5.1], which provides an estimate of asymmetries of two sets when these sets are close in L^1 sense. We refer the reader to [32, Lemma 4.1] for its generalization.

Lemma 6.3.1 ([33, Lemma 2.8])**.** *Let $D \subseteq \mathbb{R}^n$ be an open set with finite measure, $U \subseteq D$, $|U| > 0$, and*

$$\frac{|D \setminus U|}{|D|} \leq k A(D)$$

for $k \in (0, \frac{1}{2})$. Then, $A(U) \geq (1 - 2k) A(D)$.

Proof. Let B_1 be a ball centered at 0 with $|B_1| = |U|$ satisfying

$$A(U) = \frac{|U \triangle (x + B_1)|}{|U|}$$

for some $x \in \mathbb{R}^n$ and B_2 a ball centered at 0 with $|B_2| = |D|$. Note that $|U \triangle D| = |D \setminus U| = |B_1 \triangle B_2|$. Using the triangular inequality for the symmetric difference, one can see that

$$\begin{aligned} A(U) &= \frac{|U \triangle (x + B_1)|}{|U|} \\ &\geq \frac{|D \triangle (x + B_2)| - |U \triangle D| - |B_1 \triangle B_2|}{|D|} \\ &\geq A(D) - 2\frac{|D \setminus U|}{|D|} \\ &\geq (1 - 2k)A(D). \end{aligned}$$

□

Remark 6.3.2. Let D be a bounded domain in $\mathbb{R}^n$, u a nonnegative function defined in D, and $D_t = \{x : u(x) > t\}$ for $t > 0$. Assume $A(D) > 0$ and

$$t_* = \sup\{t > 0 : \mu(t) > |D|(1 - \frac{1}{4}A(D))\} > 0.$$

If $t < t_*$, then we have

$$\frac{|D \setminus D_t|}{|D|} = 1 - \frac{\mu(t)}{|D|} \leq 1 - (1 - \frac{1}{4}A(D)) = \frac{1}{4}A(D), \tag{6.3.1}$$

which yields $A(D_t) \geq \frac{1}{2}A(D)$ by Lemma 6.3.1.

6.3.2 Proof of Theorem 6.2.1

We assume that $|D| = 1$. Let $D_t = \{x \in D : u(x) > t\}$, $\mu(t) = |D_t|$, and $u(x) = u_D(x)$. By the coarea formula, we have

$$\Big(-\frac{d}{dt} \int_{D_t} |\nabla u| \, dx \Big)^2 \geq P(D_t)^2 \tag{6.3.2}$$

for almost every $t > 0$. Note that the sharp quantitative isoperimetric inequality [64] states

$$P(D) \geq P(B) + \beta_n A(D)^2 \tag{6.3.3}$$

where B is a ball with $|B| = |D| = 1$ and β_n is a dimensional constant. A simple manipulation gives

$$\begin{aligned} P(D_t)^2 &\geq P(D_t^*)^2 + 2P(D_t^*)(P(D_t) - P(D_t^*)) \\ &\geq P(D_t^*)^2 + (2n\omega_n^{\frac{1}{n}}\beta_n)\mu(t)^{2-\frac{2}{n}}A(D_t)^2 \\ &\geq n^2\omega_n^{\frac{2}{n}}\mu(t)^{2-\frac{2}{n}}\Big(1 + \frac{2}{n}\beta_n\omega_n^{-\frac{1}{n}}A(D_t)^2\Big) \end{aligned} \tag{6.3.4}$$

where ω_n is the volume of the unit ball in $\mathbb{R}^n$ and D_t^* is a ball with $|D_t| = |D_t^*|$. It follows from Cauchy–Schwarz inequality that

$$(-\mu'(t))^{\frac{1}{2}}\Big(-\frac{d}{dt}\int_{D_t}|\nabla u|^2\,dx\Big)^{\frac{1}{2}} \geq -\frac{d}{dt}\int_{D_t}|\nabla u|\,dx. \tag{6.3.5}$$

By (6.3.2), (6.3.4), and (6.3.5), we get

$$-\mu'(t)\Big(-\frac{d}{dt}\int_{D_t}|\nabla u|^2\,dx\Big) \geq n^2\omega_n^{\frac{2}{n}}\mu(t)^{2-\frac{2}{n}}\Big(1 + \frac{2}{n}\beta_n\omega_n^{-\frac{1}{n}}A(D_t)^2\Big)$$

for almost every $t > 0$. Since u is a weak solution of $-\Delta u = 1$ in D,

$$\int_D \varphi\,dx = \int_D \nabla u\cdot\nabla\varphi\,dx$$

for all $\varphi \in W_0^{1,2}(D)$. Let $\varphi(x) = (u(x) - t)_+$, then it belongs to $\varphi \in W_0^{1,2}(D)$ and

$$\int_{D_t}(u - t)\,dx = \int_{D_t}|\nabla u|^2\,dx.$$

Let $h \in \mathbb{R}$ be small enough, then

$$\frac{1}{h}\Big(\int_{D_t}|\nabla u|^2\,dx - \int_{D_{t+h}}|\nabla u|^2\,dx\Big) = \mu(t+h) + \int_{D_t\triangle D_{t+h}}\Big|\frac{u-t}{h}\Big|\,dx.$$

Since $0 \leq |u - t| \leq |h|$ in $D_t\triangle D_{t+h}$ and $|D_t\triangle D_{t+h}| \to 0$ as $h \to 0$, we obtain

$$\mu(t) = -\frac{d}{dt}\int_{D_t}|\nabla u|^2\,dx. \tag{6.3.6}$$

Therefore, we have

$$-\mu(t)^{\frac{2}{n}-1}\mu'(t) \geq n^2\omega_n^{\frac{2}{n}}(1 + \frac{2}{n}\beta_n\omega_n^{-\frac{1}{n}}A(D_t)^2) \tag{6.3.7}$$

for almost every $t > 0$.

For each $t > 0$, choose $R(t) > 0$ such that $\mu(t) = |B_{R(t)}(0)|$, where $B_{R(t)}(0)$ is the ball of radius $R(t)$, centered at 0. Let $\tau_{R(t)}$ be the first exit time from the ball $B_{R(t)}(0)$. Since $\mathbb{E}^x[\tau_{R(t)}] = \frac{1}{2n}(R(t)^2 - |x|^2)$, we have

$$\mathbb{E}^0[\tau_{R(t)}] = \frac{1}{2n}\omega_n^{-\frac{2}{n}}\mu(t)^{\frac{2}{n}}. \tag{6.3.8}$$

Differentiating of the both sides in t and applying (6.3.7), we have

$$-\frac{d}{dt}\mathbb{E}^0[\tau_{R(t)}] = -\frac{1}{n^2}\omega_n^{-\frac{2}{n}}\mu(t)^{\frac{2}{n}-1}\mu'(t) \geq 1 + \frac{2}{n}\beta_n\omega_n^{-\frac{1}{n}}A(D_t)^2.$$

Taking the integral over $[0, u_D(x)]$ and applying (6.3.8), we have

$$\begin{aligned} u_B(0) - \frac{1}{2n\omega_n^{2/n}}\mu(u_D(x))^{2/n} &= \mathbb{E}^0[\tau_{R(0)}] - \mathbb{E}^0[\tau_{R(u_D(x))}] \\ &\geq u_D(x) + \frac{2}{n}\beta_n\omega_n^{-\frac{1}{n}}\int_0^{u(x)} A(D_t)^2\,dt. \end{aligned}$$

By Lemma 6.3.1 and Remark 6.3.2, we have $A(D_t) \geq \frac{1}{2}A(D)$ for $t < t_*$ and

$$\int_0^{u(x)} A(D_t)^2\,dt \geq \int_0^{u(x)\wedge t_*} A(D_t)^2\,dt \geq \frac{1}{4}(u(x)\wedge t_*)A(D)^2.$$

Therefore, we obtain

$$\begin{aligned} u_B(0) - u_D(x) &\geq \frac{1}{2n\omega_n^{n/2}}\mu(u_D(x))^{\frac{2}{n}} + \frac{2}{n}\beta_n\omega_n^{-\frac{1}{n}}\int_0^{u(x)} A(D_t)^2\,dt \\ &\geq u_B(0)\Big(\mu(u_D(x))^{\frac{2}{n}} + \mathsf{C}_n(u(x)\wedge t_*)A(D)^2\Big) \end{aligned}$$

where $\mathsf{C}_n = \beta_n\omega_n^{\frac{1}{n}}$.

Suppose that $|D| = r^{-n}$ for some $r > 0$. By translation invariance, we assume $0 \in D$ without loss of generality. For $r > 0$, we denote by $rD = \{ry : y \in D\}$. Note that the Fraenkel asymmetry is scaling invariant, i.e. $A(D) = A(rD)$. By the

scaling property of X_t, we have $r^2 u_D(x) = u_{rD}(rx)$. This leads to the following scaling identities

$$\begin{aligned}
\delta(x, D) &= \delta(rx, rD), \\
\mu_D(t) &= |\{y : u_D(y) > t\}| = |\{y : u_{rD}(ry) > r^2 t\}| = r^{-n}\mu_{rD}(r^2 t), \\
t_*(D) &= \sup\{t > 0 : \mu_D(t) > |D|(1 - \frac{1}{4}A(D))\} \\
&= \sup\{t > 0 : \mu_{rD}(r^2 t) > |rD|(1 - \frac{1}{4}A(rD))\} \\
&= r^{-2} t_*(rD).
\end{aligned}$$

Since $|rD| = 1$, we have

$$\begin{aligned}
\delta(x, D) &= \delta(rx, rD) \\
&\geq \mu(u_{rD}(rx))^{\frac{2}{n}} + \mathsf{C}_n(u_{rD}(rx) \wedge t_*(rD))A(rD)^2 \\
&= r^2\Big(\mu(u_D(x))^{\frac{2}{n}} + \mathsf{C}_n(u(x) \wedge t_*)A(D)^2\Big) \\
&= |D|^{-\frac{2}{n}}\Big(\mu(u_D(x))^{\frac{2}{n}} + \mathsf{C}_n(u(x) \wedge t_*)A(D)^2\Big),
\end{aligned}$$

as desired. □

6.3.3 Proof of Theorem 6.2.5

If $A(D) = 0$, the results follow from (6.1.4). From now on, we assume $A(D) > 0$. By scaling invariance, we assume $|D| = 1$ without loss of generality. Let B be a ball centered at 0 with $|B| = 1$.

Consider $p \in [1, \infty)$. Let $D_t = \{x \in D : u_D(x) > t\}$ and $\mu(t) = |D_t|$. Note that Theorem 6.2.1 reads

$$\frac{1}{2n\omega_n^{2/n}}(1 - \mu(u_D(x))^{2/n}) - u_D(x) \geq \tilde{\mathsf{C}}_n(u_D(x) \wedge t_*)A(D)^2$$

where $\tilde{\mathsf{C}}_n = \frac{1}{2n\omega_n^{2/n}}\mathsf{C}_n$. By the coarea formula, we have

$$\begin{aligned}
&\frac{1}{(2n)^p\omega_n^{2p/n}}\int_D (1-\mu(u_D(x))^{2/n})^p\,dx \\
&\quad= \frac{1}{(2n)^p\omega_n^{2p/n}}\int_0^\infty \int_{\partial D_t} (1-\mu(u_D(x))^{2/n})^p |\nabla u_D|^{-1} d\mathcal{H}^{n-1}(x)\,dt \\
&\quad= -\frac{1}{(2n)^p\omega_n^{2p/n}}\int_0^\infty (1-\mu(t)^{2/n})^p \mu'(t)\,dt \\
&\quad= \frac{1}{2^{p+1}n^{p-1}\omega_n^{2p/n}} B(p,(n-2)/2) \\
&\quad= \|u_B\|_p^p
\end{aligned}$$

where $B(a,b)$ is the Beta function. Using $a^p - b^p \geq pb^{p-1}(a-b)$ for $a \geq b$, we get

$$\begin{aligned}
\|u_B\|_p^p - \|u_D\|_p^p &\geq \tilde{\mathsf{C}}_n A(D)^2 \int_D p u_D(x)^{p-1}(u_D(x)\wedge t_*)\,dx \\
&\geq \tilde{\mathsf{C}}_n A(D)^2 \int_0^{t_*} p t^{p-1}\mu(t)\,dt \\
&\geq \frac{1}{2}\tilde{\mathsf{C}}_n A(D)^2\,(t_*)^p.
\end{aligned} \tag{6.3.9}$$

In the last inequality, we used the fact that $\mu(t) > |D|(1-\frac{1}{4}A(D)) \geq \frac{1}{2}$ for $0 < t < t_*$.

Let $\mu_0(t) = |\{x \in B : u_B(x) > t\}|$. Since $u_B(x) = \frac{1}{2n}(r_n^2 - |x|^2)$ with $r_n = \omega_n^{-\frac{1}{n}}$, we have

$$\mu_0(t) = \left(1 - 2n\omega_n^{\frac{2}{n}} t\right)^{\frac{n}{2}}. \tag{6.3.10}$$

Choose $t_0 > 0$ so that $\mu_0(2t_0) = 1 - \frac{1}{8}A(D)$. By (6.3.10) and the inequality $1-(1-x)^a \geq ax$ for $0 \leq x, a \leq 1$, we have

$$t_0 = \frac{1}{4n\omega_n^{\frac{2}{n}}}\left(1-(1-\frac{1}{8}A(D))^{\frac{2}{n}}\right) \geq \frac{1}{16n^2\omega_n^{\frac{2}{n}}}A(D). \tag{6.3.11}$$

Suppose $t_* < t_0$, then $\mu(t) \le 1 - \frac{1}{4}A(D)$ for all $t \ge t_0$ by definition. Since $\mu_0(t) \ge 1 - \frac{1}{8}A(D)$ for $t \le 2t_0$, we get $\mu_0(t) - \mu(t) \ge \frac{1}{8}A(D)$ for $t \in [t_0, 2t_0]$. By the layer cake representation and (6.3.11), we have

$$\begin{aligned}
\|u_B\|_p^p - \|u_D\|_p^p &= \int_0^\infty pt^{p-1}(\mu_0(t) - \mu(t))\,dt \\
&\ge \int_{t_0}^{2t_0} pt^{p-1}(\mu_0(t) - \mu(t))\,dt \\
&\ge \frac{p}{8}(t_0)^p A(D) \\
&\ge \frac{p}{2^{4p+3}n^{2p}\omega_n^{\frac{2p}{n}}} A(D)^{1+p} \\
&\ge \frac{p}{2^{4(p+1)}n^{2p}\omega_n^{\frac{2p}{n}}} A(D)^{2+p}.
\end{aligned}$$

If $t_* \ge t_0$, then it follows from (6.3.9) and (6.3.11) that

$$\|u_B\|_p^p - \|u_D\|_p^p \ge \frac{\tilde{\mathsf{C}}_n}{2^{4p+1}n^{2p}\omega_n^{\frac{2p}{n}}} A(D)^{2+p}.$$

For $1 \le p < \infty$, we finish the proof of (6.2.6) by letting

$$\begin{aligned}
\mathsf{C}_{n,p} &= \frac{1}{2^{4(p+1)}n^{2p}\omega_n^{\frac{2p}{n}}\|u_B\|_p^p} \min\{p, 8\tilde{\mathsf{C}}_n\} \\
&= \frac{1}{2^{3(p+1)}n^{p+1}B(p,(n-2)/2)} \min\left\{p, \frac{4\beta_n}{n\omega_n^{\frac{1}{n}}}\right\}
\end{aligned} \tag{6.3.12}$$

where β_n is the constant in (6.3.3).

Consider the case $p = \infty$. By translation invariance, we assume that $0 \in D$ and $u_D(0) = \max_{y\in D} u_D(y)$ without loss of generality. Putting $x = 0$ in (6.2.4), we get

$$\delta_\infty(D) \ge \mathsf{C}_n t_* A(D)^2.$$

Let $\mu_0(t) = |\{x \in B : u_B(x) > t\}|$ and choose $t_0 > 0$ so that $\mu_0(2t_0) = 1 - \frac{1}{8}A(D)$ as above. If $t_* \ge t_0$, then it follows from (6.3.11) that

$$\delta_\infty(D) \ge \frac{\mathsf{C}_n}{16n^2\omega_n^{\frac{2}{n}}} A(D)^3.$$

Let $t_* < t_0$. Let $\varepsilon > 0$ be small enough that $t_1 := t_* + \varepsilon < t_0$ and $D_1 = \{x \in D : u_D(x) > t_1\}$, then D_1 is open. Let $\tilde{B}$ be a ball centered at 0 with $|\tilde{B}| = |D_1|$ and $\tilde{t}$

be such that $\mu_0(\tilde{t}) = \mu(t_1)$. Since $1 - \frac{1}{4}A(D) > \mu(\tilde{t})$, we have $\tilde{t} > 2t_0$. Recall that the strong Markov property of X_t yields for any $x \in U \subset D$ that

$$\mathbb{E}^x[\tau_D] = \mathbb{E}^x[\tau_U] + \mathbb{E}^x[\mathbb{E}^{X_{\tau_U}}[\tau_D]].$$

Since the paths of X_t are continuous a.s., we have $X_{\tau_{D_1}} \in \partial D_1$ a.s. Since D_1 is open, $\partial D_1 \subset \mathbb{R}^n \setminus D_1$ and $u_D(y) \leq t_1$ for $y \in \partial D_1$. Then we obtain

$$\mathbb{E}^0[\tau_D] = \mathbb{E}^0[\tau_{D_1}] + \mathbb{E}^0[\mathbb{E}^{X_{\tau_{D_1}}}[\tau_D]] \leq \mathbb{E}^0[\tau_{D_1}] + t_1.$$

On the other hand, it follows from a direct computation that $\mathbb{E}^0[\tau_B] = \mathbb{E}^0[\tau_{\tilde{B}}] + \tilde{t}$. Since $\mathbb{E}^0[\tau_{\tilde{B}}] \geq \mathbb{E}^0[\tau_{D_1}]$ by (6.1.2), we get

$$\begin{aligned}
\|u_B\|_\infty - \|u_D\|_\infty &= u_B(0) - u_D(0) \\
&\geq (\mathbb{E}^0[\tau_{\tilde{B}}] - \mathbb{E}^0[\tau_{D_1}]) + t_0 \\
&\geq \frac{1}{16n^2\omega_n^{\frac{2}{n}}} A(D) \\
&\geq \frac{\|u_B\|_\infty}{32n} A(D)^3.
\end{aligned}$$

We complete the proof by letting

$$\mathsf{C}_{n,\infty} = \min\left\{ \frac{\beta_n}{16n^2\omega_n^{\frac{1}{n}}}, \frac{1}{32n} \right\}. \tag{6.3.13}$$

□

6.3.4 Proof of Theorem 6.2.7

Since $A(D) < 2$, it suffices to consider the case $\frac{1}{2}T_\alpha(B) \leq T_\alpha(D)$. Let u_D^α be the expected lifetime of the α-stable process in D, $\mu(t) = |\{y \in D : u_D^\alpha(y) > t\}|$, and $t_* = \sup\{t > 0 : \mu(t) > |D|(1 - \frac{1}{9}A(D))\}$. By the proof of [32, Theorem 1.3], one has

$$T_\alpha(B) - T_\alpha(D) \geq C_{n,\alpha} T_\alpha(B)^2 (t_*)^{\frac{4}{\alpha}} A(D)^{\frac{2}{\alpha}}. \tag{6.3.14}$$

Let $\mu_0(t) = |\{y \in D : u_B^\alpha(y) > t\}|$. Since $u_B^\alpha(x) = C_{n,\alpha}(r^2 - |x|^2)^{\frac{\alpha}{2}}$ and $r = \omega_n^{-\frac{1}{n}}$, we have

$$\mu_0(t) = (1 - C_{n,\alpha} t^{\frac{2}{\alpha}})^{\frac{n}{2}}.$$

Choose $t_0 > 0$ such that $\mu_0(2t_0) = 1 - \frac{1}{18}A(D)$, then

$$t_0 = C_{n,\alpha}(1 - (1 - \frac{1}{18}A(D))^{\frac{2}{n}})^{\frac{\alpha}{2}} \geq C_{n,\alpha}A(D)^{\frac{\alpha}{2}}. \tag{6.3.15}$$

If $t_* < t_0$, then $\mu(t) \leq 1 - \frac{1}{9}A(D)$ for all $t \geq t_0$ by definition. Since $\mu_0(t) \geq 1 - \frac{1}{18}A(D)$ for $t \leq 2t_0$, we get $\mu_0(t) - \mu(t) \geq \frac{1}{18}A(D)$ for $t \in [t_0, 2t_0]$. By the layer cake representation and (6.3.11), we have

$$\begin{aligned} T_\alpha(B) - T_\alpha(D) &= \int_0^\infty (\mu_0(t) - \mu(t))\, dt \\ &\geq \frac{1}{18}t_0 A(D) \\ &\geq C_{n,\alpha}A(D)^{1+\frac{\alpha}{2}}. \end{aligned}$$

If $t_* \geq t_0$, then by (6.3.14) and (6.3.15) we have

$$T_\alpha(B) - T_\alpha(D) \geq C_{n,\alpha}T_\alpha(B)^2 A(D)^{2+\frac{2}{\alpha}},$$

which completes the proof. □

6.4 Related open problems

6.4.1 Brownian motion

It is open to find quantitative improvement of (6.1.1) and (6.1.3) even for Brownian motion. In particular, it is unclear what is the right statement for stability of (6.1.1). Having a small deficit of (6.1.1) at some t is not enough to obtain the proximity of the region to a ball, which implies that the deficit should be defined in a strong sense.

As discussed in Remark 6.2.6, it is expected that the sharp exponent of (6.2.6) is 2 for $1 < p \leq \infty$. For $p = 1$, the sharp result was derived in [34]. It is, however, not obvious how to apply the method of [34] to the case $1 < p \leq \infty$ because the proof strongly replies on the variational formula (2.5.2), whereas the L^p norm of the expected lifetime does not have such formula.

In Theorem 6.2.1, our quantitative result of (6.1.2) for $\alpha = 2$ depends on t_*. It is unclear whether this dependence is necessary. Removing t_* in (6.2.4) is an interesting open problem.

It was shown in [34] that the sharp exponent of $A(D)$ in (6.2.7) is 2. Since the proof is based on the selection principle of [49], the constant is not explicit. The best-known exponent with an explicit constant is 3. It is still open to prove a sharp quantitative result of (6.2.7) with an explicit dimensional constant.

6.4.2 Symmetric stable processes

As mentioned above, it is an open problem to extend Theorems 6.2.1 and Theorem 6.2.5 to the case $0 < \alpha < 2$. At this moment, a fractional analogue of the inequality (6.1.4) for $0 < \alpha < 2$ and $1 < p \leq \infty$ is not known. Our approach of Theorem 6.2.1 may not work for this case since it is not obvious how to apply the coarea formula in the fractional setting. A standard way of avoiding this difficulty is to consider the extension of Caffarelli–Silvestre [39]. Fusco, Millot, and Morini [65] considered the rearrangement inequality for the extension to show the quantitative isoperimetric inequality for the fractional perimeter. Recently, Brasco, Cinti, and Vita [32] proved stability of the fractional Faber–Krahn inequality using a similar argument. As a corollary, they also showed stability of the fractional Saint-Venant inequality.

We introduce some notations. The fractional Laplacian of order $\frac{\alpha}{2}$ is given by

$$(-\Delta)^{\frac{\alpha}{2}} f(x) = \mathsf{A}_{n,\alpha} \int_{\mathbb{R}^n} \frac{f(x) - f(y)}{|x-y|^{n+\alpha}}\, dy \tag{6.4.1}$$

where

$$\mathsf{A}_{n,\alpha} = \frac{2^\alpha \Gamma\big(\frac{n+\alpha}{2}\big)}{\pi^{\frac{n}{2}} |\Gamma\big(-\frac{\alpha}{2}\big)|}. \tag{6.4.2}$$

The space $\widetilde{W}_0^{\alpha,p}(D)$ is the closure of $C_0^\infty(D)$ with respect to the norm $u \mapsto [u]_{\alpha,p} + \|u\|_{L^p(D)}$ where

$$[u]_{\alpha,p} = \Big(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^p}{|x-y|^{n+\alpha p/2}}\, dxdy \Big)^{\frac{1}{p}}.$$

The fractional torsional rigidity of order α is defined by $\|u_D^\alpha\|_1$. We have the following variational representations

$$T_\alpha(D) = \max_{u \in \widetilde{W}_0^{\alpha,2}(D)\setminus\{0\}} \Big(2\|u\|_{L^1(D)} - \frac{\mathsf{A}_{n,\alpha}}{2} [u]_{\alpha,2}^2 \Big) = \max_{u \in \widetilde{W}_0^{\alpha,2}(D)\setminus\{0\}} \frac{2}{\mathsf{A}_{n,\alpha}} [u]_{\alpha,2}^{-2} \|u\|_{L^1(D)}^2$$

where $\mathsf{A}_{n,\alpha}$ is given by (6.4.2). In particular, since $u_D^\alpha \in \widetilde{W}_0^{\alpha,2}(D)$ we have

$$T_\alpha(D) = \|u_D^\alpha\|_{L^1(D)} = \frac{\mathsf{A}_{n,\alpha}}{2}[u_D^\alpha]_{\alpha,2}^2 = \frac{\mathsf{A}_{n,\alpha}}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{|u_D^\alpha(x)-u_D^\alpha(y)|^2}{|x-y|^{n+\alpha}}\,dxdy. \tag{6.4.3}$$

Consider a solution of the equation

$$\begin{cases} \operatorname{div}(z^{1-\alpha}\nabla U) = 0, & (x,z)\in\mathbb{R}^{n+1}_+, \\ U(x,0) = u_D^\alpha(x), & x\in\mathbb{R}^n. \end{cases}$$

Then we have

$$[u_D^\alpha]_{\alpha,2}^2 = \gamma_{n,\alpha}\iint_{\mathbb{R}^{n+1}_+} z^{1-\alpha}|\nabla U|^2\,dxdz$$

for some constant $\gamma_{n,\alpha}$. Let $U^*(x,z) = (U(\cdot,z))^*(x)$ be the symmetric decreasing rearrangement of U with respect to x, then it was shown in [65, Lemma 2.6] that

$$\iint_{\mathbb{R}^{n+1}_+} z^{1-\alpha}|\nabla_x U|^2\,dxdz \geq \iint_{\mathbb{R}^{n+1}_+} z^{1-\alpha}|\nabla_x U^*|^2\,dxdz \tag{6.4.4}$$

and

$$\iint_{\mathbb{R}^{n+1}_+} z^{1-\alpha}|\partial_z U|^2\,dxdz \geq \iint_{\mathbb{R}^{n+1}_+} z^{1-\alpha}|\partial_z U^*|^2\,dxdz.$$

In [32], the authors improved (6.4.4) quantitatively as in the local case, which leads to a quantitative fractional Saint-Venant inequality.

To generalize Theorems 6.2.1 and Theorem 6.2.5 to the α-stable processes, one might need to apply this extension and symmetrization argument at the level of the function U, not the seminorm $[u_D^\alpha]_{\alpha,2}$. Then it is required to show that a quantitative improvement can be transferred as z tends to 0. For $\alpha = 1$, this approached was also used in [12–14] to study spectral gap estimates and properties of nodal domains. Because of its connection to the Cauchy process and the Steklov problem, this special case may be more tractable with such an approach.

6.4.3 A fractional Pólya–Szegö inequality

We discuss stability of fractional Pólya–Szegö inequalities. The fractional α–perimeter of D is defined by

$$P_\alpha(D) = \int_D \int_{\mathbb{R}^n \setminus D} \frac{1}{|x-y|^{n+\alpha/2}} dxdy = \frac{1}{2}[\mathbb{1}_D]_{\alpha,1}.$$

Note that $P_\alpha(D) \geq C_{n,\alpha}|D|^{\frac{2n-\alpha}{2n}}$ by the fractional Sobolev embedding. The quantitative isoperimetric inequality for fractional perimeter [65] states that for $n \geq 1$ and $\alpha \in (0,2)$, there exists a constant $\mathsf{B}_{n,\alpha}$ such that for all Borel set $D \subset \mathbb{R}^n$ with $0 < |D| < \infty$,

$$P_\alpha(D) \geq P_\alpha(D^*)(1 + \mathsf{B}_{n,\alpha} A(D)^{\frac{2}{\alpha}}). \tag{6.4.5}$$

By layer cake representation, we obtain a fractional version of the coarea formula [35, Lemma 4.7]. Indeed, if $u \in L^1(\mathbb{R}^n)$ is a nonnegative function vanishing at ∞, then

$$[u]_{\alpha,1} = 2\int_0^\infty P_\alpha(\{x : u(x) > t\})dt. \tag{6.4.6}$$

We have a fractional version of the Pólya–Szegö inequality with a remainder term.

Proposition 6.4.1. *Let $\alpha \in (0,2)$ and D be a bounded domain in $\mathbb{R}^n$ with $A(D) > 0$. If $u \in \widetilde{W}_0^{\alpha,1}(D)$, then there exists $t_* > 0$ such that*

$$[u]_{\alpha,1} \geq [u^*]_{\alpha,1} + C_{n,\alpha} A(D)^{\frac{2}{\alpha}} \max\{t_* |D|^{\frac{2n-\alpha}{2n}}, \|u \wedge t_*\|_{\frac{2n-\alpha}{2n}}\}.$$

Proof. Let $D_t = \{x : u(x) > t\}$ and $\mu(t) = |D_t|$. Using the coarea formula (6.4.6) and the quantitative isoperimetric inequality for fractional perimeter (6.4.5), we have

$$\begin{aligned}
[u]_{\alpha,1} &= 2\int_0^\infty P_\alpha(D_t)dt \\
&\geq 2\int_0^\infty P_\alpha(D_t^*)dt + 2\mathsf{B}_{n,\alpha}\int_0^\infty P_\alpha(D_t^*)A(D_t)^{\frac{2}{\alpha}}dt \\
&\geq [u^*]_{\alpha,1} + C_{n,\alpha}\int_0^\infty \mu(t)^{\frac{2n-\alpha}{2n}} A(D_t)^{\frac{2}{\alpha}} dt
\end{aligned}$$

for some constant $C_{n,\alpha}$. Let $t_* = \sup\{t > 0 : \mu(t) \geq |D|(1 - \frac{1}{4}A(D))\}$. By Lemma 6.3.1 and (6.3.1), we have $A(D_t) \geq \frac{1}{2}A(D)$ for $t < t_*$ and

$$[u]_{\alpha,1} \geq [u^*]_{\alpha,1} + C_{n,\alpha} t_* |D|^{\frac{2n-\alpha}{2n}} A(D)^{\frac{2}{\alpha}}.$$

Using the inequality

$$\Big(\int_0^\infty f(x)dx\Big)^r \geq \int_0^\infty r f(x)^r x^{r-1}dx$$

for $r \geq 1$ and a nonnegative, non-increasing function f on $(0,\infty)$ (see [93, p.49]), we get

$$\int_0^{t_*} \mu(t)^{\frac{1}{r}}dt \geq \Big(\int_0^{t_*} rt^{r-1}\mu(t)dt\Big)^{\frac{1}{r}} = \|u \wedge t_*\|_r$$

where $r = \frac{2n}{2n-\alpha} > 1$, which implies

$$[u]_{\alpha,1} \geq [u^*]_{\alpha,1} + C_{n,\alpha}\|u \wedge t_*\|_{\frac{2n}{2n-\alpha}} A(D)^{\frac{2}{\alpha}}.$$

□

A natural question is a quantitative improvement of the inequality $[u]_{\alpha,2} \geq [u^*]_{\alpha,2}$ in terms of $A(D)$. This open problem is interesting because it yields a quantitative Saint-Venant inequality. Suppose that we have $[u]_{\alpha,2} \geq [u^*]_{\alpha,2} + \Phi(t_*, A(D))$ for some function Φ. By (6.4.3) and the rearrangement inequality [62], we get

$$\begin{aligned} T_\alpha(D) &\leq \frac{2}{\mathsf{A}_{n,\alpha}} \frac{\|u^*\|_1^2}{[u^*]_{\alpha,2}^2 + \Phi(t_*, A(D))} \\ &\leq T_\alpha(B)\Big(1 + \frac{\Phi(t_*, A(D))}{[u^*]_{\alpha,2}^2}\Big)^{-1} \end{aligned}$$

where $u = u_D^\alpha$ is the α–torsion function and B is a ball with $|D| = |B|$. Using the fact that $[u^*]_{\alpha,2}^2 \leq [u]_{\alpha,2}^2$, we get

$$T_\alpha(B) - T_\alpha(D) \geq \Phi(t_*, A(D)).$$

Under mild assumption on Φ, t_* can be removed as in Theorem 6.2.5.

7. THE HARDY–LITTLEWOOD–SOBOLEV INEQUALITY

7.1 Introduction

The classical Hardy–Littlewood–Sobolev inequality [77,78,105] (the HLS inequality) states that if $0 < \alpha < d$ and $1 = \frac{1}{p} + \frac{1}{r} - \frac{\alpha}{d}$, then there exists a constant $C_{\alpha,p,d}$ such that

$$\left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{f(x)h(y)}{|x-y|^{d-\alpha}} \, dxdy \right| \leq C_{\alpha,p,d} \|f\|_p \|h\|_r \tag{7.1.1}$$

for $f \in L^p(\mathbb{R}^d)$ and $h \in L^r(\mathbb{R}^d)$. Lieb [90] showed the existence of maximizing functions in (7.1.1) based on the rearrangement inequalities and a compactness argument. Also, he explicitly computed the maximizing functions f and h and so the sharp constant $C_{\alpha,p,d}$, for the spacial cases $p = r$, $p = 2$, and $r = 2$. After this, there has been many effort to find a different proof of the sharp result: competing symmetry [43], inversion positivity [60], fast diffusion flows [42]. Frank and Lieb [61] introduced a rearrangement-free proof of the sharp HLS inequality, which leads to an analogue of the sharp inequality on the Heisenberg group. For the recent progress on the extension of the sharp HLS inequality, we refer to [46,47,54,95,96].

In this chapter, we give a probabilistic representation for fractional integrals for symmetric Markov semigroups and derive an analogue of the Hardy–Littlewood–Sobolev inequality using the background radiation process, which was exploited in [71–73], together with time reversal, to represent the Riesz transforms via harmonic extensions. To prove the HLS inequality, we introduce a fractional analogue of the Littlewood–Paley function for symmetric Markov semigroups and prove Littlewood–Paley type inequalities. This chapter is based on my work [84].

Our representation is a variation of the one used in [3] based on the space-time Brownian motion often used for the second order Riesz transforms. In [3], Applebaum

and Bañuelos give a probabilistic proof of the HLS inequality on $\mathbb{R}^d$ using their representation and the martingale inequalities of Doob and Burkholder–Davis–Gundy. Unlike the space-time Brownian motion representation which requires the gradient of the harmonic extension in the space variable (or a carré du champ), our representation only requires the time derivative which is well-defined for symmetric Markov semigroups.

The probabilistic representation of the fractional integrals can be thought of as a martingale transform where the predictable sequence is not bounded. Martingale transform techniques have been used quite effectively in the study of singular integral operators, particularly in obtaining optimal, or near optimal, inequalities. For some of this extensive literature on this subject, we refer the reader to [5,7,15,17,66,89,97] and references therein. Given the powerful martingale and Bellman function methods pioneered by Burkholder in [37] to obtain sharp inequalities for martingale transforms and their many subsequent uses in various problems in analysis and probability (see for example Osękowski [98]), it is natural to ask if those techniques can be extended to martingale transforms with unbounded multipliers and provide a different proof of the sharp HLS inequalities which could be extended to other settings. Unfortunately, as of now, we have not been able to obtain sharp results with the Bellman function methods. This remains an interesting challenging problem.

7.2 Main results

Let $\mathcal{S}$ be a locally compact space with a countable base equipped with a positive Radon measure dx on $\mathcal{S}$ and $\{T_t\}_{t\geq 0}$ a strongly continuous symmetric Markov semigroup. We assume that the semigroup is Feller and has the *Varopoulos* dimension d that we will define below. The fractional integral of order α $(0 < \alpha < d)$ associated to $\{T_t\}_{t\geq 0}$ is defined by

$$\mathcal{I}_\alpha(f)(x) = \frac{1}{\Gamma(\frac{\alpha}{2})} \int_0^\infty t^{\frac{\alpha}{2}-1} T_t f(x) dt. \tag{7.2.1}$$

Note that if $\{T_t\}_{t\geq 0}$ is the standard heat semigroup on $\mathbb{R}^d$ then (7.2.1) reads

$$\mathcal{I}_\alpha(f)(x) = \frac{\Gamma(\frac{d-\alpha}{2})}{2^\alpha \pi^{d/2}\Gamma(\frac{\alpha}{2})} \int_{\mathbb{R}^d} \frac{f(y)}{|x-y|^{d-\alpha}} dy,$$

which is called the Riesz potential associated with the symmetric Markov semigroup $\{T_t\}$. If $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$, $\frac{1}{q} + \frac{1}{q'} = 1$, and $0 < \alpha < d$, then the HLS inequality for $\mathcal{I}_\alpha$ states that

$$|\langle \mathcal{I}_\alpha(f), h\rangle| \leq C_{\alpha,p,d}\|f\|_p\|h\|_{q'} \tag{7.2.2}$$

for $f \in L^p$ and $h \in L^{q'}$.

Suppose that $(X_t)_{t\geq 0}$ is a stochastic process associated to $\{T_t\}_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ is the standard 1-dimensional Brownian motion independent of $(X_t)_{t\geq 0}$. Let $Z_t = (X_t, Y_t)$. Since $\{T_t\}_{t\geq 0}$ is Feller, $(X_t)_{t\geq 0}$ is right continuous with left limits and has the strong Markov property. Fix $s > 0$ and assume that the initial distribution of $(Z_t)_{t\geq 0}$ is given by $dx \otimes \delta_s$. We denote by $\mathbf{E}^s$ the expectation of $(Z_t)_{t\geq 0}$. Let τ be the hitting time of Y_t at 0 and $\{P_y\}_{y\geq 0}$ the Poisson semigroup associated with $\{T_t\}_{t\geq 0}$ (see (7.3.1)). Let $u_f(x,y) = P_y f(x)$ be the harmonic extension of f defined on $\mathcal{S} \times [0,\infty)$. We set

$$\mathcal{T}^s_\alpha(f)(x) = \mathbf{E}^s[\int_0^\tau Y_t^\alpha \frac{\partial u_f}{\partial y}(Z_t) dY_t | X_\tau = x]. \tag{7.2.3}$$

The main result of this chapter is to show that $\mathcal{T}^s_\alpha$ gives a probabilistic representation of the fractional integral, and that it satisfies the analogue of the HLS inequality (7.2.2).

Theorem 7.2.1. *Let $s > 0$ and $f, h \in C_0(\mathcal{S})$. If $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$, $1 < p < q < \infty$, $0 < \alpha < d$, and q' is the conjugate exponent of q, then we have*

$$|\langle \mathcal{T}^s_\alpha f, h\rangle| = \left|\mathbf{E}^s\left[\int_0^\tau Y_t^\alpha \frac{\partial u_f}{\partial y}(Z_t)\frac{\partial u_h}{\partial y}(Z_t)\, dt\right]\right| \leq C_{\alpha,p,d}\|f\|_p\|h\|_{q'} \tag{7.2.4}$$

where $C_{\alpha,p,d}$ depends only on α, p and d. As a consequence, we have

$$\lim_{s\to\infty} \mathcal{T}^s_\alpha(f) = \frac{\Gamma(\alpha+2)}{2^{\alpha+2}}\mathcal{I}_\alpha(f)$$

in the distributional sense.

The proof of Theorem 7.2.1 relies on an auxiliary function which satisfies an HLS-type inequality. To be specific, we define the fractional Littlewood–Paley function $\mathcal{G}_\alpha$ by

$$\mathcal{G}_\alpha(f)(x) = \left(\int_0^\infty y^{2\alpha+1} \left| \frac{\partial u_f}{\partial y}(x,y) \right|^2 dy \right)^{1/2}. \tag{7.2.5}$$

The next theorem says that the fractional Littlewood–Paley function satisfies an HLS-type inequality, which leads to the HLS inequality for $\mathcal{T}_\alpha^s$.

Theorem 7.2.2. *Let $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d} > 0$, $1 < p < q < \infty$, and $0 < \alpha < d$. If $f \in L^p(\mathcal{S})$, then the fractional Littlewood–Paley function $\mathcal{G}_\alpha(f)$ defined in (7.2.5) satisfies*

$$\|\mathcal{G}_\alpha(f)\|_q \le C_{\alpha,p,d} \|f\|_p.$$

7.3 Preliminaries

7.3.1 Notations

The space of all continuous functions on $\mathcal{S}$ vanishing at ∞ is denoted by $C_0(\mathcal{S})$. We also use $C_c(\mathcal{S})$ to denote the space of all compactly supported continuous functions. The lower case letter c, c_1, $c_2, \cdots$ denote generic constants which may change from line to line. We use the notation $C_{p,q,r}$ to specify that the constant depends on p, q and r. We denote the inner product by $\langle f, g \rangle = \int_{\mathcal{S}} f(x)g(x)dx$ for notational convenience. The domain of an operator A is denoted by $\mathrm{Dom}(A)$.

7.3.2 General semigroup theory

We recall some facts about semigroups that we will call upon later. Particularly, we review the definition of a strongly continuous symmetric Markov semigroup and the construction of the Poisson semigroup used in the probabilistic representation of the fractional integral (7.2.3).

We say that a semigroup $\{T_t\}_{t\ge 0}$ on $\mathcal{S}$ is a *symmetric Markov semigroup* if it has the following properties:

(S1) If $f \geq 0$, then $T_t f \geq 0$.

(S2) $T_t 1 = 1$ for all $t \geq 0$.

(S3) (Symmetry) If $f, g \in L^2(\mathcal{S})$, then $\langle T_t f, g\rangle = \langle f, T_t g\rangle$ for all $t \geq 0$.

(S4) (L^p-contraction) If $1 \leq p \leq \infty$ and $f \in L^p(\mathcal{S})$, then $\|T_t f\|_p \leq \|f\|_p$ for all $t \geq 0$.

In what follows, we assume that there exists a symmetric Markov semigroup $\{T_t\}_{t\geq 0}$ on $\mathcal{S}$. We also assume that the semigroup is *strongly continuous* on $L^2(\mathcal{S})$ and a *Feller semigroup*:

(S5) (Strong continuity) If $f \in L^2(\mathcal{S})$, then $\lim_{t\to 0} \|T_t f - f\|_2 = 0$.

(S6) (Feller) If $f \in C_0(\mathcal{S})$, then $T_t f \in C_0(\mathcal{S})$ for all $t \geq 0$ and $\lim_{t\to 0} \|T_t f - f\|_\infty = 0$.

We assume that $\{T_t\}_{t\geq 0}$ has the *Varopoulos dimension* d ($d > 2$) in the sense of [112]:

(S7) (Varopoulos dimension) If $1 \leq p < \infty$ and $f \in L^p(\mathcal{S})$, there exists $C > 0$ such that

$$\|T_t f\|_\infty \leq C t^{-\frac{d}{2p}} \|f\|_p \tag{7.3.1}$$

for all $t > 0$.

For instance, the heat semigroup $e^{-t\Delta}$ on $\mathbb{R}^d$ ($d \geq 3$) has the Varopoulos dimension d.

Given a symmetric Markov semigroup $\{T_t\}_{t\geq 0}$, the Poisson semigroup associated to $\{T_t\}_{t\geq 0}$ is defined in the following ways. The first way is to use the spectral decomposition on $L^2(\mathcal{S})$. For $f \in L^2(\mathcal{S})$, $\{T_t\}_{t\geq 0}$ can be written as

$$T_t f(x) = \int_0^\infty e^{-\lambda t} dE_\lambda f(x)$$

where $\{E_\lambda : \lambda \geq 0\}$ is the spectral resolution associated to the infinitesimal generator of $\{T_t\}_{t\geq 0}$. The corresponding Poisson semigroup on $L^2(\mathcal{S})$ is defined by

$$P_t f(x) = \int_0^\infty e^{-\sqrt{\lambda} t} dE_\lambda f(x). \tag{7.3.2}$$

Another way of defining the Poisson semigroup is to subordinate $\{T_t\}_{t\geq 0}$ in the sense of Bochner [28]. For $1 \leq p \leq \infty$ and $f \in L^p(\mathcal{S})$, the Poisson semigroup is defined by

$$P_t f(x) = \int_0^\infty T_s f(x) \mu_t(ds) \tag{7.3.3}$$

where $\mu_t(ds) = \frac{t}{2\sqrt{\pi}} e^{-t^2/4s} s^{-3/2} ds$. For $p = 2$, it follows from a direct calculation that (7.3.3) is equivalent to (7.3.2). We notice that this construction is a special case of the subordination. Generally speaking, one obtains a new semigroup by subordinating with a convolution measure on $[0, \infty)$, which is a Lévy process on $[0, \infty)$ from the probabilistic point of view. In (7.3.3), we adopted the convolution measure $\mu_t(ds)$ called the $\frac{1}{2}$-stable subordinator. The harmonic extension of f is defined by $u_f(x, y) = P_y f(x)$.

Lemma 7.3.1. *Let $\{T_t\}_{t\geq 0}$ be a strongly continuous symmetric Markov semigroup and $\{P_y\}$ the Poisson semigroup defined by (7.3.2). Then $\{P_y\}$ is also a strongly continuous symmetric Markov semigroup. In addition, if $\{T_t\}_{t\geq 0}$ has the Varopoulos dimension d, then there exists $C > 0$ such that*

$$\|P_y f\|_\infty = \|u_f(\cdot, y)\|_\infty \leq \frac{C}{y^{d/p}} \|f\|_p \tag{7.3.4}$$

for all $f \in L^p$, $1 \leq p < \infty$, and $y > 0$. (That is, $\{P_y\}$ has the Varopoulos dimension $2d$)

Proof. The assumptions (S1), (S2), and (S3) follow from the definition (7.3.3). By Jensen's inequality, we see

$$\begin{aligned} \|P_y f\|_p^p &= \int_{\mathcal{S}} |P_y f(x)|^p dx \\ &\leq \int_{\mathcal{S}} \int_0^\infty |T_s f(x)|^p \mu_y(ds) dx \\ &= \int_0^\infty \|T_s f\|_p^p \mu_y(ds) \leq \|f\|_p^p. \end{aligned}$$

Similarly, one can show that P_y is strongly continuous on L^2. Since $\{T_t\}_{t\geq 0}$ has the dimension d, we have

$$\begin{aligned}|P_y f(x)| &= \left|\int_0^\infty T_s f(x)\mu_y(ds)\right| \\ &\leq \int_0^\infty |T_s f(x)|\mu_y(ds) \\ &\leq C\|f\|_p \int_0^\infty s^{-\frac{d}{2p}}\mu_y(ds) \\ &\leq Cy^{-\frac{d}{p}}\|f\|_p,\end{aligned}$$

which yields (7.3.4) as desired. □

Note that for each $x \in \mathcal{S}$ and $f \in L^p$ $(1 < p < \infty)$, $u_f(x,\cdot)$ is real-analytic [107, p.67, p.72]. Next lemma is concerned with a derivative estimate for the harmonic extension u_f.

Lemma 7.3.2. *Let f be a bounded measurable function on $\mathcal{S}$, then there exists $c_1 > 0$ such that*

$$\left|y\frac{\partial u_f}{\partial y}(x,y)\right| \leq c_1 u_{|f|}(x, \frac{y}{\sqrt{2}}).$$

Proof. Let $\mu_y(ds) = \frac{1}{2\sqrt{\pi}}\eta_y(s)ds$, then we have

$$y\frac{\partial \eta_y(s)}{\partial y} = (1 - \frac{y^2}{2s})ye^{-y^2/4s}s^{-3/2}.$$

Since there exists a constant c_1 such that $|1-\frac{y^2}{2s}| \leq c_1 e^{y^2/8s}$ for every $y > 0$ and $s > 0$, we have

$$\left|y\frac{\partial \eta_y}{\partial y}(s)\right| \leq c_1 ye^{-y^2/8s}s^{-3/2} = c_1\eta_{\frac{y}{\sqrt{2}}}(s)$$

for every $y > 0$ and $s > 0$. We finish the proof by interchanging the differentiation and the integral. □

Let A_T and A_P be the infinitesimal generators of $\{T_t\}_{t\geq 0}$ and $\{P_t\}_{t\geq 0}$ respectively, then we have $A_P = -(-A_T)^{\frac{1}{2}}$. Let

$$\begin{aligned}R_0 &= \{f \in \mathrm{Dom}(A_T) : A_T(f) \in \mathrm{Dom}(A_T)\}, \\ R_n &= \bigcap_{k=1}^{n} \mathrm{Dom}(A_P^k)\end{aligned} \tag{7.3.5}$$

for $n \geq 1$. If $1 \leq k \leq n$ and $f \in R_n$, then the $\frac{\partial^k}{\partial y^k} u_f \in R_{n-k}$. Since $\{T_t\}_{t\geq 0}$ and $\{P_t\}_{t\geq 0}$ are Feller, R_n is contained in $C_0(\mathcal{S})$ for every $n \geq 0$, which implies that R_n is dense in L^p for $p \geq 1$ and $n \geq 0$. Thus it suffices to consider $C_0(\mathcal{S})$ in what follows. We refer the reader to [111, p.29] and [114, Chap. IV §10, §11] for further discussion.

We recall the maximal ergodic theorem, which plays an important role in the proof of Theorem 7.2.2. Stein [107] gives two different proofs. One is to use the Hopf–Dunford–Schwartz ergodic theorem with an interpolation argument. The other way is to rely on the martingale inequalities via the result of Rota [101]. For the completeness, we provide a continuous martingale version of the second proof, which is a special case of [104, Theorem 3.1].

Proposition 7.3.3 (Maximal ergodic theorem)**.** *If $1 < p \leq \infty$ and $f \in L^p(\mathcal{S})$, then*

$$\left\| \sup_{y>0} |u_f(\cdot, y)| \right\|_p \leq C(p) \|f\|_p,$$

where $C(p) = \frac{p}{p-1}$ for $1 < p < \infty$ and $C(\infty) = 1$.

Proof. We prove the result for a general symmetric Markov semigroup $\{Q_t\}_{t\geq 0}$. Let $(X_t)_{t\geq 0}$ be the stochastic process corresponding to $\{Q_t\}_{t\geq 0}$, that is, $Q_t f(x) = \mathbb{E}^x[f(X_t)]$ for $f \in L^p$. We assume $1 < p < \infty$ since the case $p = \infty$ is trivial. Let $T > 0$ be fixed and $\{\mathcal{F}_t : t \geq 0\}$ the natural filtration of X_t. By the Markov property, we have

$$\begin{aligned} Q_{2(T-t)} f(X_T) &= Q_{T-t}(Q_{T-t} f)(X_T) \\ &= \mathbb{E}^{X_T}[Q_{T-t} f(X_{T-t})] \\ &= \mathbb{E}^x[Q_{T-t} f(X_{2T-t}) | \mathcal{F}_T]. \end{aligned}$$

Since

$$\sup_{0\leq t\leq T} |Q_{2(T-t)} f(X_T)|^p \leq \mathbb{E}^x[\sup_{0\leq t\leq T} |Q_{T-t} f(X_{2T-t})|^p | \mathcal{F}_T],$$

we have

$$\begin{aligned} \int_{\mathcal{S}} \mathbb{E}^x[\sup_{0\leq t\leq T} |Q_{2(T-t)} f(X_T)|^p]\, dx &\leq \int_{\mathcal{S}} \mathbb{E}^x[\sup_{0\leq t\leq T} |Q_{T-t} f(X_{2T-t})|^p]\, dx \\ &= \int_{\mathcal{S}} \mathbb{E}^x[\sup_{0\leq t\leq T} |Q_{T-t} f(X_t)|^p]\, dx. \end{aligned} \tag{7.3.6}$$

We have used the reversibility of X_t in the equality. Note that $Q_{T-t}f(X_t^x)$ is a martingale because $Q_{T-t}f(X_t) = \mathbb{E}^x[f(X_T)|\mathcal{F}_t]$. Then Doob's maximal inequality yields

$$\mathbb{E}^x[\sup_{0\le t\le T}|Q_{T-t}f(X_T)|^p] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}^x[|f(X_T)|^p]. \tag{7.3.7}$$

Since Q_t is self-adjoint and $Q_t 1 = 1$, we have

$$\int_{\mathcal{S}} \mathbb{E}^x[g(X_T)]dx = \int_{\mathcal{S}} Q_T g(x)dx = \int_{\mathcal{S}} g(x)dx$$

for any bounded measurable function g. Applying this to (7.3.6) and (7.3.7), we get

$$\begin{aligned}\|\sup_{0\le t\le T}|Q_{2(T-t)}f(x)|\|_p &\le \left(\int_{\mathcal{S}} \mathbb{E}^x[\sup_{0\le t\le T}|Q_{T-t}f(X_t)|^p]dx\right)^{\frac{1}{p}} \\ &\le \frac{p}{p-1}\left(\int_{\mathcal{S}} \mathbb{E}^x[|f(X_T)|^p]dx\right)^{\frac{1}{p}} \\ &= \frac{p}{p-1}\|f\|_p.\end{aligned}$$

We complete the proof by letting $T \to \infty$. □

For a function $f \in L^p(\mathcal{S})$ and $k \ge 1$, the Littlewood–Paley function of order k is defined by

$$g_k(f)(x) = \Big(\int_0^\infty y^{2k-1}\Big|\frac{\partial^k u_f}{\partial y^k}(x,y)\Big|^2 dt\Big)^{\frac{1}{2}}.$$

Proposition 7.3.4. *Let $1 < p < \infty$ and $k \ge 1$. If $f \in L^p(\mathcal{S})$, then $g_k(f) \in L^p(\mathcal{S})$ and satisfies*

$$\|g_k(f)\|_p \le C_{p,k}\|f\|_p$$

for some constant $C_{p,k}$ depending only on p and k.

We refer the reader to [106, p.111, p.120] for the proof. In what follows, we only use the Littlewood–Paley function of order 1.

7.3.3 Stochastic analysis

Let $\{T_t\}_{t\ge 0}$ be a strongly continuous symmetric Markov semigroup of the Varopoulos dimension d and $\{H_t\}_{t\ge 0}$ the heat semigroup on $\mathbb{R}$ defined by

$$H_t f(x) = \frac{1}{(2\pi t)^{1/2}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2t}} f(y)dy.$$

Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be the stochastic processes on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ associated with $\{T_t\}_{t\geq 0}$ and $\{H_t\}_{t\geq 0}$ respectively. We assume that $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are independent each other and their paths are right-continuous with left limits a.s. Let $Z_t = (X_t, Y_t) \in \mathcal{S} \times \mathbb{R}$. For example, if T_t is the standard heat semigroup on $\mathbb{R}^n$, then Z_t is Brownian motion on $\mathbb{R}^{n+1}$. Let $\tau = \inf\{t \geq 0 : Y_t = 0\}$ be the hitting time of Y_t at 0. From now on, we consider the killed process $(Z_{t\wedge\tau})_{t\geq 0}$.

For fixed $s > 0$, we assume that the initial distribution of $(Z_t)_{t\geq 0}$ is given by $dx \otimes \delta_s$ where δ_s is the Dirac delta measure at fixed $s > 0$. In other words, $(Z_t)_{t\geq 0}$ starts at $(x_0, s) \in \mathcal{S} \times \mathbb{R}$ where x_0 is randomly chosen with respect to the measure dx. The probability and expectation of Z_t with the initial distribution are denoted by $\mathbf{E}^s$ and $\mathbf{P}^s$ respectively. Explicitly, we have

$$\mathbf{E}^s = \int_{\mathcal{S}} \mathbb{E}^{(x,s)} dx, \quad \mathbf{P}^s = \int_{\mathcal{S}} \mathbb{P}^{(x,s)} dx.$$

Note that even though $\mathbf{P}^s$ may not be a probability measure, all the results from probability theory connected with this context remain valid as explained in [111].

Let $h \in L^1(\mathcal{S})$ and P_y be the Poisson semigroup associated with T_t. Since P_y is invariant and symmetric, we have

$$\mathbf{E}^s[h(X_\tau)] = \int_{\mathcal{S}} \mathbb{E}^{(x,s)} h(X_\tau) dx = \int_{\mathcal{S}} P_s h(x) dx = \int_{\mathcal{S}} h(x) dx.$$

We recall the Green function formula for Z_t.

Lemma 7.3.5 ([111, Proposition 3.1]). *For a Borel measurable function f on $\mathcal{S} \times \mathbb{R}$, we have*

$$\mathbf{E}^s[\int_0^\tau f(Z_t) dt] = 2 \int_0^\infty \int_{\mathcal{S}} (y \wedge s) f(x, y)\, dx dy. \tag{7.3.8}$$

Definition 7.3.6. We say a stochastic process $(A_t)_{t\geq 0}$ on $(\Omega, \mathbf{P}^s, \mathscr{F}_t)$ is in $L^2(\Omega, \mathbf{P}^s)$ if the map $A : \Omega \times [0, \infty) \to \mathbb{R}$ is jointly measurable, $A_t \in \mathscr{F}_t$ for every $t \in [0, \infty)$, and

$$\mathbf{E}^s[\int_0^\infty |A_t|^2\, dt] < \infty. \tag{7.3.9}$$

Let $(A_t)_{t\geq 0} \in L^2(\Omega, \mathbf{P}^s)$. If the "probability" $\mathbf{P}^s$ is finite, then we define

$$I(A)_t := \int_0^t A_s \, dY_s$$

as a L^2-limit of martingale transforms using Itô's isometry. If $\mathbf{P}^s$ is infinite, we decompose the Radon measure dx into a countable family of finite measures dx_n and define the stochastic integral for each finite measure dx_n as above. Then we define $I(A)_t$ by the sum of the stochastic integrals. The sum is well-defined by (7.3.9). We refer the reader to [111, pp.37-38].

We recall *the projection lemma*, which is an analogue of Itô's formula for the $(d+1)$-dimensional Brownian motion. We omit the proof and refer to [111, pp.50-59]. Let V be the set of stochastic processes in $L^2(\Omega, \mathbf{P}^s)$ of the form $(I(A)_t)_{t\geq 0}$. Note that V is a closed subspace. Let Φ_V be the orthogonal projection from $L^2(\Omega, \mathbf{P}^s)$ onto V.

Proposition 7.3.7. *Let R_n be defined as in (7.3.5). If $f \in R_5$, then*

$$\Phi_V(u_f(Z_{t\wedge\tau}) - u_f(Z_0)) = \int_0^{t\wedge\tau} \frac{\partial u_f}{\partial y}(Z_s) dY_s$$

for all $t > 0$.

7.4 Proofs of the main results

7.4.1 Proof of Theorem 7.2.2

For $\delta > 0$, we divide $\mathcal{G}_\alpha(f)^2$ into two parts

$$\begin{aligned}\mathcal{G}_\alpha(f)(x)^2 &= \int_0^\infty y^{2\alpha+1} \left|\frac{\partial u_f}{\partial y}(x,y)\right|^2 dy \\ &= \int_0^\delta y^{2\alpha+1} \left|\frac{\partial u_f}{\partial y}(x,y)\right|^2 dy + \int_\delta^\infty y^{2\alpha+1} \left|\frac{\partial u_f}{\partial y}(x,y)\right|^2 dy.\end{aligned}$$

Applying Lemma 7.3.2 to the first integral, we obtain

$$\begin{aligned}\int_0^\delta y^{2\alpha+1} \left|\frac{\partial u_f}{\partial y}(x,y)\right|^2 dy &\leq c_1 \int_0^\delta y^{2\alpha-1} \left|u_{|f|}(x, \frac{1}{\sqrt{2}}y)\right|^2 dy \\ &\leq C_\alpha \sup_{y>0} \left|u_{|f|}(x,y)\right|^2 \delta^{2\alpha}.\end{aligned}$$

For the second integral, we apply Lemma 7.3.1 and Lemma 7.3.2 to see

$$\int_\delta^\infty y^{2\alpha+1}\left|\frac{\partial u_f}{\partial y}(x,y)\right|^2 dy \le c_1\int_\delta^\infty y^{2\alpha-1}|u_{|f|}(x,\frac{1}{\sqrt{2}}y)|^2 dy \le C_\alpha\|f\|_p^2\delta^{2(\alpha-\frac{d}{p})},$$

which yields

$$\mathcal{G}_\alpha(f)(x) \le C_{\alpha,p,d}(\sup_{y>0}|u_{|f|}(x,y)|\delta^\alpha + \|f\|_p\delta^{\alpha-\frac{d}{p}})$$

for some constant $C_{\alpha,p,d}$. Optimizing the RHS in δ yields

$$\mathcal{G}_\alpha(f)(x) \le C_{\alpha,p,d}(\sup_{y>0}|u_{|f|}(x,y)|)^{1-\frac{\alpha p}{d}}\|f\|_p^{\frac{\alpha p}{d}}.$$

Proposition 7.3.3 yields

$$\|(\sup_{y>0}|u_{|f|}(x,y)|)^{1-\frac{\alpha p}{d}}\|_q = \|\sup_{y>0}|u_{|f|}(x,y)|\|_p^{\frac{p}{q}} \le C_p\|f\|_p^{\frac{p}{q}}$$

because $1-\frac{\alpha p}{d}=\frac{p}{q}$. Therefore, we obtain

$$\begin{aligned}\|\mathcal{G}_\alpha(f)\|_q &\le C_{\alpha,p,d}\|(\sup_{y>0}|u_{|f|}(x,y)|)^{p/q}\|_q\|f\|_p^{1-p/q}\\ &= C_{\alpha,p,d}\|(\sup_{y>0}|u_{|f|}(x,y)|)\|_p^{p/q}\|f\|_p^{1-p/q}\\ &\le C_{\alpha,p,d}\|f\|_p,\end{aligned}$$

which finishes the proof. □

7.4.2 Proof of Theorem 7.2.1

We claim that

$$\mathbf{E}^s[\int_0^\tau Y_t^\alpha\left|\frac{\partial u_f}{\partial y}(Z_t)\right|\left|\frac{\partial u_h}{\partial y}(Z_t)\right|dt] \le C_{\alpha,p,d}\|f\|_p\|h\|_{q'}. \tag{7.4.1}$$

Applying the Green function formula (7.3.8), we see

$$\begin{aligned}\mathbf{E}^s[\int_0^\tau Y_t^\alpha\left|\frac{\partial u_f}{\partial y}(Z_t)\right|\left|\frac{\partial u_h}{\partial y}(Z_t)\right|dt] &= 2\int_{\mathcal{S}}\int_0^\infty (y\wedge s)y^\alpha\left|\frac{\partial u_f}{\partial y}\right|\left|\frac{\partial u_h}{\partial y}\right|dydx\\ &\le \int_{\mathcal{S}}\int_0^\infty y^{\alpha+\frac{1}{2}}\left|\frac{\partial u_f}{\partial y}\right|y^{\frac{1}{2}}\left|\frac{\partial u_h}{\partial y}\right|dydx\\ &\le \int_{\mathcal{S}}\mathcal{G}_\alpha(f)g_1(h)dx\\ &\le \|\mathcal{G}_\alpha(f)\|_q\|g_1(h)\|_{q'}.\end{aligned}$$

The claim follows from Proposition 7.3.4 and Theorem 7.2.2.

For $N > 0$, we define

$$\mathcal{T}_\alpha^{s,N}(f)(x) = \mathbf{E}^s[\int_0^\tau (Y_t^\alpha \wedge N)\frac{\partial u_f}{\partial y}(Z_t)dY_t | X_\tau = x].$$

By Lemma 7.3.5, we have

$$\begin{aligned}
\langle \mathcal{T}_\alpha^{s,N}(f), h\rangle &= \mathbf{E}^s[\mathcal{T}_\alpha^{s,N}(f)(X_\tau)h(X_\tau)] \\
&= \mathbf{E}^s[\mathbf{E}^s[\int_0^\tau (Y_t^\alpha \wedge N)\frac{\partial u_f}{\partial y}(Z_t)dY_t | X_\tau]h(X_\tau)] \\
&= \mathbf{E}^s[\mathbf{E}^s[h(X_\tau)\int_0^\tau (Y_t^\alpha \wedge N)\frac{\partial u_f}{\partial y}(Z_t)dY_t | X_\tau]] \\
&= \mathbf{E}^s[h(X_\tau)\int_0^\tau (Y_t^\alpha \wedge N)\frac{\partial u_f}{\partial y}(Z_t)dY_t].
\end{aligned}$$

Note that

$$I_t := \int_0^{t\wedge\tau} (Y_t^\alpha \wedge N)\frac{\partial u_f}{\partial y}(Z_t)dY_t \in L^2(\Omega, \mathbf{P}^s).$$

Indeed, it follows from the Green formula (7.3.8) that

$$\begin{aligned}
\mathbf{E}^s[\int_0^{t\wedge\tau} (Y_t^{2\alpha} \wedge N^2)\left|\frac{\partial u_f}{\partial y}(Z_t)\right|^2 dt] &\le 2N^2 \int_0^\infty \int_{\mathcal{S}} y \left|\frac{\partial u_f}{\partial y}(x,y)\right|^2 dxdy \\
&= 2N^2\|g_1(f)\|_2^2 \\
&\le cN^2\|f\|_2^2 < \infty.
\end{aligned}$$

Furthermore, $I_t \in V$, where V is the closed subspace of $L^2(\Omega, \mathbf{P}^s)$ of stochastic integrals with respect to $(Y_t)_{t\ge 0}$. Thus, Proposition 7.3.7 yields that

$$\begin{aligned}
\langle \mathcal{T}_\alpha^{s,N}(f), g\rangle &= \mathbf{E}^s[\left(\int_0^\tau \frac{\partial u_h}{\partial y}(Z_t)dY_t\right)\left(\int_0^\tau (Y_t^\alpha \wedge N)\frac{\partial u_f}{\partial y}(Z_t)dY_t\right)] \\
&= \mathbf{E}^s[\int_0^\tau (Y_t^\alpha \wedge N)\frac{\partial u_f}{\partial y}(Z_t)\frac{\partial u_h}{\partial y}(Z_t)dt].
\end{aligned}$$

By (7.4.1), the dominated convergence theorem, and letting $N \to \infty$, we obtain

$$\langle \mathcal{T}_\alpha^{s}(f), h\rangle = \mathbf{E}^s[\int_0^\tau Y_t^\alpha \frac{\partial u_f}{\partial y}(Z_t)\frac{\partial u_h}{\partial y}(Z_t)dt].$$

Finally, we show that $\mathcal{T}_\alpha^s f$ converges to $c_\alpha \mathcal{I}_\alpha(f)$ as s tends to ∞ in the distributional sense. By (7.2.4) and the Green function formula (7.3.8), we see

$$\langle \mathcal{T}_\alpha^s f, h\rangle = 2\int_0^\infty \int_{\mathcal{S}} (y \wedge s)y^\alpha \frac{\partial u_f}{\partial y}(x,y)\frac{\partial u_h}{\partial y}(x,y)dxdy.$$

Thus it suffices to show

$$\langle \mathcal{I}_\alpha f, h\rangle = C_\alpha \int_0^\infty \int_{\mathcal{S}} y^{\alpha+1} \frac{\partial u_f}{\partial y}(x,y)\frac{\partial u_h}{\partial y}(x,y)dxdy.$$

Since f and g are in L^2, it follows from (7.3.2) that

$$\begin{aligned}\int_{\mathcal{S}} \frac{\partial u_f}{\partial y}(x,y)\frac{\partial u_h}{\partial y}(x,y)dx &= \left\langle \frac{\partial u_f}{\partial y}(\cdot,y), \frac{\partial u_h}{\partial y}(\cdot,y)\right\rangle \\ &= \left\langle \int_0^\infty \lambda^{1/2} e^{-\lambda^{1/2}y} dE_\lambda f, \int_0^\infty \lambda^{1/2} e^{-\lambda^{1/2}y} dE_\lambda h\right\rangle \\ &= \int_0^\infty \lambda e^{-2\lambda^{1/2}y} d\langle E_\lambda f, E_\lambda h\rangle.\end{aligned}$$

By Fubini's theorem, we get

$$\begin{aligned}\int_0^\infty \int_{\mathcal{S}} y^{\alpha+1} \frac{\partial u_f}{\partial y}(x,y)\frac{\partial u_h}{\partial y}(x,y)dxdy &= \int_0^\infty y^{\alpha+1} \left\langle \frac{\partial u_f}{\partial y}(\cdot,y), \frac{\partial u_h}{\partial y}(\cdot,y)\right\rangle dy \\ &= \int_0^\infty y^{\alpha+1} \left(\int_0^\infty \lambda e^{-2\lambda^{1/2}y} d\langle E_\lambda f, E_\lambda h\rangle \right) dy \\ &= \int_0^\infty \lambda \left(\int_0^\infty y^{\alpha+1} e^{-2\lambda^{1/2}y} dy \right) d\langle E_\lambda f, E_\lambda h\rangle \\ &= \frac{\Gamma(\alpha+2)}{2^{\alpha+2}} \int_0^\infty \lambda^{-\alpha/2} d\langle E_\lambda f, E_\lambda h\rangle \\ &= C_\alpha \langle I_\alpha f, h\rangle,\end{aligned}$$

which completes the proof. □

8. HARDY–STEIN IDENTITY FOR NON-SYMMETRIC LÉVY PROCESSES AND FOURIER MULTIPLIERS

8.1 Introduction

Littlewood–Paley square (quadratic) functions have been of interest for many years with many applications in harmonic analysis and probability. On the analysis side, these include the classical square functions obtained from the Poisson semigroup as in [106] and more general heat semigroups as in [107]. On the probability side, these correspond to the celebrated Burkholder–Gundy inequalities which are of fundamental importance in modern stochastic analysis.

In [10], the authors extend some of the classical Littlewood–Paley L^p inequalities for $1 < p < \infty$ to symmetric pure jump Lévy processes and apply them to prove L^p bounds for a certain class of Fourier multipliers that arise from transformations of symmetric Lévy processes. The key to the proof in [10] is a Hardy–Stein identity, which is proved from properties of the semigroup. In the classical case of the Laplacian, such Hardy–Stein identity follows from, essentially, Green's theorem and the chain rule as in Lemmas 1 and 2 in [106, pp.86-87]. In the case of Brownian motion, a probabilistic Burkholder–Gundy type version of this Hardy–Stein identity can be proved (see [5], [100, p.152]) as a simple application of Itô's formula.

The goal of this chapter is to extend the results of [10] to non-symmetric pure jump Lévy processes. The first result is a Hardy–Stein identity for non-symmetric Lévy measure (Theorem 8.3.1). The proof is based on the Itô's formula for jump processes (Theorem 8.2.1). It turns out that this method gives a Hardy–Stein type identity for uniformly integrable martingales ((Theorem 8.3.5). Furthermore, the proof contains additional information, further illuminating the origins of the function $F(a, b; p)$ (see (8.3.1)) used in [10].

In the second part, we introduce a certain class of the Fourier multipliers for non-symmetric pure jump Lévy measures and prove the L^p boundedness of the Fourier multipliers (Theorem 8.4.1). It is important to emphasize that although the Hardy–Stein identity holds for non-symmetric Lévy measures, the full comparability of the L^p-norms between the function itself and its Littlewood–Paley square function proved in [10] requires symmetry and hence the main application given there to the boundedness of the Fourier multipliers requires it too. Thus we use a symmetrization of the Littlewood–Paley function (see (8.4.2) and (8.4.3)) to obtain the L^p inequalities for the Littlewood–Paley functions (Lemma 8.4.2), which leads to the L^p boundedness of the Fourier multipliers.

This chapter is based on joint work with Rodrigo Bañuelos [11].

8.2 Preliminaries

8.2.1 Notations

The indicator function of a set A is denoted by $\mathbb{1}_A$. For $a, b \in \mathbb{R}$, we denote by $a \wedge b = \min\{a, b\}$. The real part of a complex number ξ is denoted by $\mathrm{Re}(\xi) = x$ where $\xi = x + iy$. For a set $B \subseteq \mathbb{R}^d$, we define $-B = \{-x : x \in B\}$. An open ball in $\mathbb{R}^d$ of radius r, centered at $x_0 \in \mathbb{R}^d$ is denoted by $B_r(x_0)$. We denote by $B_r(0) = B_r$. For $f, g \in L^2(\mathbb{R}^d)$, we define the inner product of f and g in $L^2(\mathbb{R}^d)$ by $\langle f, g\rangle = \int_{\mathbb{R}^d} f(x)g(x)dx$. Let $\mathcal{S}(\mathbb{R}^d)$ be the Schwartz space on $\mathbb{R}^d$ and $f \in \mathcal{S}(\mathbb{R}^d)$. We define the Fourier transform and the inverse Fourier transform of f by

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-ix\cdot\xi}\,dx,$$
$$\mathcal{F}^{-1}(f)(x) = f^{\vee}(x) = (2\pi)^{-d}\int_{\mathbb{R}^d} f(\xi)e^{i\xi\cdot x}\,d\xi.$$

With our definition, Parseval's formula takes the form

$$\int_{\mathbb{R}^d} f(x)g(x)\,dx = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi, \tag{8.2.1}$$

for $f, g \in L^2(\mathbb{R}^d)$. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For a multi-index $\alpha \in \mathbb{N}_0^d$, we use the notations $|\alpha| = \alpha_1 + \cdots + \alpha_d$ and $\nabla^\alpha = \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}$. The space of continuous functions vanishing at infinity is denoted by $C_0(\mathbb{R}^d)$. For $k \in \mathbb{N}$, $C_0^k(\mathbb{R}^d)$ is the space of functions $f \in C^k(\mathbb{R}^d)$ such that $\nabla^\alpha f \in C_0(\mathbb{R}^d)$ for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$, and $C_0^\infty(\mathbb{R}^d)$ is the intersection of all $C_0^k(\mathbb{R}^d)$ over $k \in \mathbb{N}$.

8.2.2 Lévy processes

A d-dimensional stochastic process $(X_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *Lévy process* if

(i) for $0 \leq t_0 < t_1 < \cdots < t_n < \infty$, $\{X_{t_k} - X_{t_{k-1}}\}_{k\geq 1}$ are independent,

(ii) for $0 < s < t < \infty$ and a Borel set $A \subseteq \mathbb{R}^n$, $\mathbb{P}(X_t - X_s \in A) = \mathbb{P}(X_{t-s} \in A)$, and

(iii) for all $\delta > 0$ and $s \geq 0$,

$$\lim_{t\to s} \mathbb{P}(|X_t - X_s| > \delta) = 0. \tag{8.2.2}$$

The characteristic exponent $\psi(\xi)$ of a Lévy process $(X_t)_{t\geq 0}$ is defined by $\mathbb{E}[e^{i\xi\cdot X_t}] = e^{-t\psi(\xi)}$ for $\xi \in \mathbb{R}^d$. The Lévy–Khintchine theorem tells us that $(X_t)_{t\geq 0}$ is a Lévy process with characteristic exponent $\psi(\xi)$ if and only if there exists a triplet (b, A, ν) such that

$$\psi(\xi) = ib\cdot\xi + \frac{1}{2}\xi\cdot A\xi + \int_{\mathbb{R}^d} (1 - e^{i\xi\cdot y} + i\xi\cdot y \mathbb{1}_{B_1}(y))\, \nu(dy),$$

where $b \in \mathbb{R}^d$, A is a positive semi-definite $d \times d$ matrix, and ν is a σ-finite Borel measure on $\mathbb{R}^d \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}^d\setminus\{0\}} (1 \wedge |y|^2)\, \nu(dy) < \infty.$$

We call ν the *Lévy measure*. This gives a large class of stochastic processes that have been extensively studied. For instance, Brownian motion is the case where $b = 0$,

$\nu = 0$, and A is the identity matrix. We say that $(X_t)_{t\geq 0}$ is a *pure jump Lévy process* if $b = 0$ and $A = 0$, and *symmetric* if ν is symmetric. We refer the reader to [2] for further information on these processes. The jump of X_t at time s is denoted by $\Delta X_s = X_s - X_{s-}$. For $t \geq 0$ and a Borel subset $A \subseteq \mathbb{R}^n \setminus \{0\}$, we define the jump measure of $(X_t)_{t\geq 0}$ by

$$\begin{aligned} N(t, A) &= \text{the number of jumps during time } [0, t] \text{ of size in } A \\ &= \#\{s \in [0, t] : \Delta X_s \in A\}. \end{aligned}$$

Note that $N(t, A)$ is a Poisson random measure with intensity $dt \otimes d\nu$. By the Lévy–Itô decomposition theorem [2, Theorem 2.4.16], one can decompose X_t into

$$X_t = bt + G_t + \int_{|x|\geq 1} x\, N(t, dx) + \int_{|x|<1} x\, \widetilde{N}(t, dx),$$

where $b \in \mathbb{R}^d$, G_t is a Gaussian process, and $\widetilde{N}(t, A) = N(t, A) - t\nu(A)$. Following the standard terminology, we call $\widetilde{N}(t, A)$ the compensated jump measure. Let $P_t f(x) = \mathbb{E}^x[f(X_t)]$, then the semigroup P_t has the Feller property: for $f \in C_0(\mathbb{R}^d)$, $P_t f \in C_0(\mathbb{R}^d)$ and $\lim_{t\to 0} |P_t f(x) - f(x)| = 0$ uniformly in x (see [19, Theorem 3.1.9] and [22, p.19]). The infinitesimal generator $\mathcal{L}$ for the semigroup $(P_t)_{t\geq 0}$ is given by

$$\mathcal{L}f(x) = \lim_{t\downarrow 0} \frac{P_t f(x) - f(x)}{t}$$

whenever the limit exists. Here the limit is taken in the supremum norm. Let $D(\mathcal{L})$ be the domain of $\mathcal{L}$, then $C_0^2(\mathbb{R}^d) \subset D(\mathcal{L})$ and $\mathcal{L}$ can be explicitly written as

$$\begin{aligned} \mathcal{L}f(x) = b \cdot \nabla f(x) + \frac{1}{2} \sum_{i,j=1}^{d} A_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \\ + \int_{\mathbb{R}^d} (f(x + y) - f(x) - y \cdot \nabla f(x) \mathbb{1}_{B_1}(y))\, \nu(dy) \end{aligned} \quad (8.2.3)$$

for $f \in C_0^2(\mathbb{R}^d)$, where (b, A, ν) is the triplet of X_t (see [102, Theorem 31.5]).

8.2.3 Itô's formula

We recall Itô's formula for a general stochastic process Z_t from [81, Theorem 5.1, p. 66]. Let M_t be a continuous square integrable local martingale and A_t a

continuous adapted process of bounded variation with $A_0 = 0$. Let $(X_t)_{t\geq 0}$ be a Lévy process with its jump measure $N(t,\cdot)$. Let $G(t,x) = (G_1(t,x),\cdots,G_d(t,x))$ and $H(t,x) = (H_1(t,x),\cdots,H_d(t,x))$ be d-dimensional predictable processes such that $G_i(t,x)H_j(t,x) = 0$,

$$\int_0^t \int_{\mathbb{R}^d} |G_i(s,x)|\, N(ds,dx) < \infty \quad \text{a.s.}, \tag{8.2.4}$$

and

$$\mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |H_i(s\wedge\tau_n,x)|^2\, \nu(dx)ds\Big] < \infty, \tag{8.2.5}$$

for all $t > 0$ and $i,j = 1,2,\cdots,d$, where (τ_n) is a sequence of stopping times such that $\tau_n \to \infty$ as $n \to \infty$ almost surely. Let $(Z_t)_{t\geq 0}$ be the d-dimensional stochastic process defined by

$$Z_t = Z_0 + M_t + A_t + \int_0^t \int_{\mathbb{R}^d} G(s,x)\, N(ds,dx) + \int_0^t \int_{\mathbb{R}^d} H(s,x)\, \widetilde{N}(ds,dx). \tag{8.2.6}$$

Theorem 8.2.1. *Let $(Z_t)_{t\geq 0}$ be given by (8.2.6) and $\varphi \in C^2(\mathbb{R}^d)$. Assume that for all $1 \leq i,j \leq d$ and $T > 0$,*

$$\sup_{0\leq t\leq T} \sup_{x\in\mathbb{R}^d} |H(t,x)| < \infty \tag{8.2.7}$$

almost surely. Then we have

$$\begin{aligned}
\varphi(Z_t) - \varphi(Z_0) = &\int_0^t \nabla\varphi(Z_s)\cdot dM_s \\
&+ \int_0^t \nabla\varphi(Z_s)\cdot dA_s + \frac{1}{2}\int_0^t D^2\varphi(Z_s)\cdot d[M]_s \\
&+ \int_0^t \int_{\mathbb{R}^d} (\varphi(Z_{s-} + G(s,y)) - \varphi(Z_{s-}))\, N(ds,dy) \\
&+ \int_0^t \int_{\mathbb{R}^d} (\varphi(Z_{s-} + H(s,y)) - \varphi(Z_{s-}))\, \widetilde{N}(ds,dy) \\
&+ \int_0^t \int_{\mathbb{R}^d} (\varphi(Z_{s-} + H(s,y)) - \varphi(Z_{s-}) - H(s,y)\cdot\nabla\varphi(Z_{s-}))\, \nu(dy)ds
\end{aligned} \tag{8.2.8}$$

where $[M]_t$ is the quadratic variation of M_t.

8.2.4 Hartman–Wintner condition

In what follows, we assume that $(X_t)_{t\geq 0}$ is a pure jump Lévy process with càdlàc path and its Lévy measure ν satisfies the *Hartman–Wintner condition*

$$\lim_{|\xi|\to\infty} \frac{\operatorname{Re}(\psi(\xi))}{\log(1+|\xi|)} = \infty. \tag{HW}$$

In [87, Theorem 2.1], Knopova and Schilling proved that a Lévy process $(X_t)_{t\geq 0}$ satisfies (HW) if and only if for all $t > 0$, the transition density $p_t(x,y) = p_t(y-x)$ exists and $p_t, \nabla^\alpha p_t \in C_0^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, for all $\alpha \in \mathbb{N}_0^d$. By (HW), P_t is an L^p-contraction for $1 \leq p \leq \infty$ and $P_t f \in L^p(\mathbb{R}^d) \cap C_0^\infty(\mathbb{R}^d)$ for $f \in L^p(\mathbb{R}^d)$ and $1 \leq p < \infty$ (see [87, Theorem 2.1] and [10, p. 466]).

8.2.5 Fourier multipliers

Let $m : \mathbb{R}^n \to \mathbb{C}$ be a function in L^∞. For $1 \leq p \leq \infty$ and $f \in L^2 \cap L^p$, we define an operator T_m by $\widehat{T_m f}(\xi) = m(\xi)\widehat{f}(\xi)$. If $\|T_m f\|_p \lesssim \|f\|_p$ for all $f \in L^2 \cap L^p$, then T_m can be extended to all of L^p uniquely. We say T_m is an *L^p-Fourier multiplier operator with symbol m*. For many of the classical examples of L^p–Fourier multipliers, we refer the reader to [106].

8.3 The Hardy–Stein identity

The purpose of this section is to give a proof of the Hardy–Stein identity based on Itô's formula. For $a, b \in \mathbb{R}$, $\varepsilon > 0$, and $p \in (1,\infty)$, we define

$$F(a,b;p) = |b|^p - |a|^p - pa|a|^{p-2}(b-a) \tag{8.3.1}$$

and

$$F_\varepsilon(a,b;p) = (b^2+\varepsilon^2)^{\frac{p}{2}} - (a^2+\varepsilon^2)^{\frac{p}{2}} - pa(a^2+\varepsilon^2)^{\frac{p-2}{2}}(b-a). \tag{8.3.2}$$

We note that $F(a,b;p)$ and $F_\varepsilon(a,b;p)$ are the second-order Taylor remainders of the maps $x \mapsto |x|^p$ and $x \mapsto (x^2+\varepsilon^2)^{\frac{p}{2}}$ respectively. Since the maps are convex, it follows from Taylor's theorem that $F(a,b;p) \geq 0$ and $F_\varepsilon(a,b;p) \geq 0$ for any $a, b \in \mathbb{R}$.

Theorem 8.3.1 (The Hardy–Stein identity)**.** *Let $1 < p < \infty$ and $F(a, b; p)$ be defined as in (8.3.1). If $f \in L^p(\mathbb{R}^d)$, then we have*

$$\int_{\mathbb{R}^d} |f(x)|^p \, dx = \int_{\mathbb{R}^d} \int_0^\infty \int_{\mathbb{R}^d} F(P_t f(x), P_t f(x+y); p) \, \nu(dy) dt dx. \tag{8.3.3}$$

Again we note that our proof of this result does not require that ν is symmetric as is the case in [10]. Before we present the proof of Theorem 8.3.1, we give the following lemmas. The first lemma concerns basic properties of F and F_ε which allow us to use a limiting argument when we consider the case $1 < p < 2$. This lemma is proved in [30].

Lemma 8.3.2 ([30, Lemma 6, p.198])**.** *Let $p > 1$, $F(a, b; p) = |b|^p - |a|^p - pa|a|^{p-2}(b - a)$, and $K(a, b; p) = (b - a)^2(|a| \vee |b|)^{p-2}$. Then we have*

$$c_p K(a, b; p) \le F(a, b; p) \le C_p K(a, b; p),$$

for some positive constants c_p, C_p that depend only on p. If $1 < p < 2$, then we have

$$0 \le F_\varepsilon(a, b; p) \le \frac{1}{p-1} F(a, b; p)$$

for all $\varepsilon > 0$ and $a, b \in \mathbb{R}$.

Next lemma is an application of Itô's formula, which is presented in [10, (4.4)] and [3, p. 1118] for general Lévy processes without proof. For the completeness, we give a proof.

Lemma 8.3.3. *Let $T > 0$, $t \in [0, T)$, $1 \le p < \infty$, and $f \in L^p(\mathbb{R}^d)$. For $P_t f(x) = \mathbb{E}^x[f(X_t)]$ and $Y_t = P_{T-t} f(X_t)$, we have*

$$Y_t = Y_0 + \int_0^t \int_{\mathbb{R}^d} (P_{T-s} f(X_{s-} + y) - P_{T-s} f(X_{s-})) \, \widetilde{N}(ds, dy) \tag{8.3.4}$$

for $t \in [0, T)$, where $\widetilde{N}$ is the compensated jump measure of $(X_t)_{t \ge 0}$.

Proof. Fix $t > 0$ and choose $r \in (0, T - t)$. Let $\widetilde{T} = T - r$ and $g(x) = P_r f(x)$, then it follows from (HW) and [87, Theorem 2.1] that $g \in C_0^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$. Since (8.3.4) can be written as

$$P_{\widetilde{T}-t} g(X_t) = P_{\widetilde{T}} g(X_0) + \int_0^t \int_{\mathbb{R}^d} (P_{\widetilde{T}-s} g(X_{s-} + y) - P_{\widetilde{T}-s} g(X_{s-})) \, \widetilde{N}(ds, dy),$$

it suffices to prove (8.3.4) for $f \in C_0^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$.

We claim that if $h(x)$ is $C_0(\mathbb{R}^d)$ then the map $(s,x) \mapsto P_{T-s}h(x)$ is $C_0([0,t]\times\mathbb{R}^d)$. Let $(s_0,x_0) \in [0,t]\times\mathbb{R}^d$ and $\varepsilon > 0$. By the Feller property, $P_{T-s}h$ is continuous in s uniformly in x, and $P_{T-s}h(x)$ is continuous in x for each s. Thus there exists $\delta > 0$ such that for $(s,x) \in B_\delta((s_0,x_0))$, a $(d+1)$-dimensional ball of radius δ centered at (s_0,x_0),

$$\begin{aligned} &|P_{T-s}h(x) - P_{T-s_0}h(x_0)| \\ &\le |P_{T-s}h(x) - P_{T-s_0}h(x)| + |P_{T-s_0}h(x) - P_{T-s_0}h(x_0)| + |P_{T-s}h(x_0) - P_{T-s_0}h(x_0)| \\ &< \varepsilon, \end{aligned} \tag{8.3.5}$$

which proves the claim.

Let $\varphi(s,x) = P_{T-s}f(x)$. Since $f \in C_0(\mathbb{R}^d)$, we have $\varphi(s,x) \in C([0,t]\times\mathbb{R}^d)$. Let $i \in \{1,2,\cdots,d\}$. For $h > 0$ and $i = 1,\cdots,d$, we have

$$\frac{1}{h}(P_{T-s}f(x+he_i) - P_{T-s}f(x)) = \int_{\mathbb{R}^d} \Big(\frac{f(x+y+he_i) - f(x+y)}{h}\Big) p_{T-s}(y)\,dy.$$

Since $f \in C_0^\infty(\mathbb{R}^d)$, we have

$$\Big|\frac{f(x+y+he_i) - f(x+y)}{h}\Big| \le \Big\|\frac{\partial f}{\partial x_i}\Big\|_\infty < \infty.$$

By the dominated convergence theorem and the claim, we conclude that $\frac{\partial\varphi}{\partial x_i}(s,x) = P_{T-s}(\partial_i f)(x) \in C([0,t]\times\mathbb{R}^d)$. Since X_t is a pure jump Lévy process, $\mathcal{L}f$ can be written as

$$\mathcal{L}f(x) = \int_{\mathbb{R}^d} (f(x+y) - f(x) - y\cdot\nabla f(x)\mathbb{1}_{B_1}(y))\,\nu(dy)$$

by (8.2.3). By Taylor's theorem, we have

$$|f(x+y) - f(x) - y\cdot\nabla f(x)\mathbb{1}_{B_1}(y)| \le 2\|f\|_\infty \mathbb{1}_{\mathbb{R}^d\setminus B_1}(y) + \frac{1}{2}|y|^2 \sum_{i,j=1}^d \Big\|\frac{\partial^2 f}{\partial x_i \partial x_j}\Big\|_\infty \mathbb{1}_{B_1}(y).$$

Since the RHS is integrable with respect to the Lévy measure ν, $\mathcal{L}f$ is $C_0(\mathbb{R}^d)$ by the dominated convergence theorem. It then follows from

$$\frac{\partial\varphi}{\partial s}(s,x) = \frac{\partial}{\partial s}P_{T-s}f(x) = -\mathcal{L}P_{T-s}f(x) = -P_{T-s}\mathcal{L}f(x), \tag{8.3.6}$$

that $\frac{\partial\varphi}{\partial s} \in C([0,t]\times\mathbb{R}^d)$. Therefore we have $\varphi \in C^1([0,t]\times\mathbb{R}^d)$.

Note that X_t can be written as

$$X_t = \int_{|x|\geq 1} x\, N(t,dx) + \int_{|x|<1} x\, \widetilde{N}(t,dx)$$

by the Lévy–Itô decomposition. Since X_t has no continuous martingale part, we can apply Theorem 8.2.1 for $\varphi \in C^1([0,t]\times\mathbb{R}^d)$ and the process $Z_t = (t, X_t)$. Note that Z_t is a $(d+1)$-dimensional stochastic process of the form (8.2.6) and satisfies the assumptions of Theorem 8.2.1. Thus we have

$$\begin{aligned}
\varphi(t,X_t) - \varphi(0,X_0) &= \int_0^t \frac{\partial\varphi}{\partial s}(s,X_{s-})\,ds \\
&\quad + \int_0^t \int_{|y|\geq 1} (\varphi(s,X_{s-}+y) - \varphi(s,X_{s-}))\, N(ds,dy) \\
&\quad + \int_0^t \int_{|y|<1} (\varphi(s,X_{s-}+y) - \varphi(s,X_{s-}))\, \widetilde{N}(ds,dy) \\
&\quad + \int_0^t \int_{|y|<1} (\varphi(s,X_{s-}+y) - \varphi(s,X_{s-}) - y\cdot\nabla_x\varphi(s,X_{s-}))\, \nu(dy)ds \\
&= \int_0^t \frac{\partial\varphi}{\partial s}(s,X_{s-})\,ds + \int_0^t \int_{\mathbb{R}^d} (\varphi(s,X_{s-}+y) - \varphi(s,X_{s-}))\, \widetilde{N}(ds,dy) \\
&\quad + \int_0^t \mathcal{L}\varphi(s,X_{s-})\,ds.
\end{aligned}$$

The result follows from (8.3.6). □

Although not explicitly written, the next lemma follows from [87]. Since its proof is quite simple, we present it here for the completeness.

Lemma 8.3.4. *The semigroup P_t defined by $P_tf(x) = \mathbb{E}^x[f(X_t)]$ is ultracontractive on L^p, $1 \leq p < \infty$. That is, for every $t > 0$, there exists a constant $C_t > 0$ such that for all $f \in L^p(\mathbb{R}^d)$,*

$$\|P_tf\|_\infty \leq C_t^{\frac{1}{p}} \|f\|_p. \tag{8.3.7}$$

Furthermore, C_t converges to zero as t tends to ∞.

Proof. Fix $t > 0$. Note that $e^{-t\psi(\xi)} = \mathbb{E}^0[e^{i\xi\cdot X_t}] = (2\pi)^d\mathcal{F}^{-1}(p_t(\cdot))(\xi)$. Since p_t is in $L^1(\mathbb{R}^d)$, one sees that $e^{-t\psi(\xi)}$ belongs to $L^\infty(\mathbb{R}^d)$. We claim that $e^{-t\psi(\xi)}$ is in $L^1(\mathbb{R}^d)$.

To see this, it suffices to show that $e^{-t\,\mathrm{Re}\,\psi(\xi)} \in L^1(\mathbb{R}^d)$. Let $h : \mathbb{R}^d \to \mathbb{R}$ be a function satisfying $\mathrm{Re}\,\psi(\xi) = \log(1+|\xi|)h(\xi)$. Since we have $h(\xi) \to \infty$ as $|\xi| \to \infty$ by the Hartman–Wintner condition (HW), there exists $R > 0$ such that $th(\xi) > d+1$ holds whenever $|\xi| \geq R$. Let B_R be an open ball centered at 0 and radius R. Denote its Lebesgue measure by $|B_R|$. Using the definition of h, one sees that

$$\int_{\mathbb{R}^d \setminus B_R} e^{-t\,\mathrm{Re}\,\psi(\xi)}\,d\xi = \int_{|\xi|\geq R} \frac{1}{(1+|\xi|)^{th(\xi)}}\,d\xi \leq \int_{|\xi|\geq R} \frac{1}{(1+|\xi|)^{d+1}}\,d\xi.$$

Since we have

$$e^{-t\,\mathrm{Re}\,\psi(\xi)} = |e^{-t\psi(\xi)}| = \Big|\int_{\mathbb{R}^d} e^{i\xi\cdot x} p_t(x)\,dx\Big| \leq 1, \tag{8.3.8}$$

we obtain

$$\int_{\mathbb{R}^d} e^{-t\,\mathrm{Re}\,\psi(\xi)}\,d\xi \leq \int_{|\xi|\geq R} \frac{1}{(1+|\xi|)^{d+1}}\,d\xi + |B_R| < \infty.$$

So we have $e^{-t\psi(\xi)} \in L^1(\mathbb{R}^d)$ as desired. By the Fourier inversion formula, we have

$$p_t(x) = \frac{1}{(2\pi)^d}\mathcal{F}(e^{-t\psi(\xi)}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-t\psi(\xi)} e^{-ix\cdot\xi}\,d\xi$$

and $p_t \in L^\infty(\mathbb{R}^d)$. Define

$$C_t = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-t\,\mathrm{Re}\,\psi(\xi)}\,d\xi, \tag{8.3.9}$$

then it is obvious to see that C_t is finite and $|p_t(x)| \leq C_t$ for all $x \in \mathbb{R}^d$. Using Jensen's inequality, we obtain that

$$|P_t f(x)| = \Big|\int_{\mathbb{R}^d} f(y)p_t(x,y)\,dy\Big| \leq \Big|\int_{\mathbb{R}^d} |f(y)|^p p_t(x,y)\,dy\Big|^{\frac{1}{p}} \leq C_t^{\frac{1}{p}}\|f\|_p,$$

for any $x \in \mathbb{R}^d$, which yields (8.3.7).

We now prove the second assertion that $C_t \to 0$ as $t \to \infty$. First, we note that $\mathrm{Re}\,\psi(\xi)$ is nonnegative by (8.3.8) and in fact the Lebesgue measure of the set $\{\xi : \mathrm{Re}\,\psi(\xi) = 0\}$ is zero (see [9, §3]). Thus $e^{-t\,\mathrm{Re}\,\psi(\xi)}$ tends to 0, a.e., as $t \to \infty$. Since $e^{-t\,\mathrm{Re}\,\psi(\xi)}$ is integrable for all $t \geq 1$ and bounded by $e^{-\mathrm{Re}\,\psi(\xi)}$, it follows from the dominated convergence theorem that

$$\lim_{t\to\infty} C_t = \lim_{t\to\infty} \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-t\,\mathrm{Re}\,\psi(\xi)}d\xi = 0.$$

□

We are ready to prove the Hardy–Stein identity.

Proof of Theorem 8.3.1. Let $p \geq 2$. Fix $T > 0$ and let $0 < T_0 < T$. Consider $\varphi(x) = |x|^p$, $Y_t = P_{T-t}f(X_t)$, and $H(t,x) = P_{T-t}f(X_{t-} + x) - P_{T-t}f(X_{t-})$ for $0 \leq t \leq T_0$ and $x \in \mathbb{R}^d$. It follows from Lemma 8.3.3 that

$$Y_t = Y_0 + \int_0^t \int_{\mathbb{R}^d} H(s,y)\, \widetilde{N}(ds,dy)$$

for $0 \leq t \leq T_0$. By Lemma 8.3.4,

$$\mathbb{E}|Y_t|^2 = \mathbb{E}|P_{T-t}f(X_t)|^2 \leq C_{T-T_0}^{\frac{2}{p}} \|f\|_p^2 < \infty$$

for $0 \leq t \leq T_0$ and $H(t,x)$ satisfies (8.2.5) and (8.2.7). Applying Itô's formula to $\varphi(Y_t)$, we obtain

$$\begin{aligned}\varphi(Y_t) - \varphi(Y_0) &= \int_0^t \int_{\mathbb{R}^d} (\varphi(Y_{s-} + H(s,y)) - \varphi(Y_{s-}))\, \widetilde{N}(ds,dy) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} (\varphi(Y_{s-} + H(s,y)) - \varphi(Y_{s-}) - H(s,y) \cdot \nabla\varphi(Y_{s-}))\, \nu(dy)ds \quad (8.3.10)\end{aligned}$$

for all $0 \leq t \leq T_0$. Note that $Y_{s-} + H(s,y) = P_{T-s}f(X_{s-} + y)$, $Y_{s-} = P_{T-s}f(X_{s-})$, and

$$\mathbb{E}|Y_t - Y_0|^2 = \mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |P_{T-s}f(X_{s-} + y) - P_{T-s}f(X_{s-})|^2\, \nu(dy)ds\Big] < \infty \quad (8.3.11)$$

for all $0 \leq t \leq T_0$. By Lemma 8.3.4, we have

$$\begin{aligned}|\varphi(Y_{s-} + H(s,y)) - \varphi(Y_{s-})| &= ||P_{T-s}f(X_{s-} + y)|^p - |P_{T-s}f(X_{s-})|^p| \\ &\leq p|P_{T-s}f(X_{s-} + y)|^{p-1}|P_{T-s}f(X_{s-} + y) - P_{T-s}f(X_{s-})| \\ &\leq pC_{T-T_0}^{(p-1)/p}\|f\|_p^{p-1}|P_{T-s}f(X_{s-} + y) - P_{T-s}f(X_{s-})|.\end{aligned}$$

Here we used the fact that the constant C_t in (8.3.9) is decreasing in t. By (8.3.11), we see that

$$\begin{aligned}&\mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |\varphi(Y_{s-} + H(s,y)) - \varphi(Y_{s-})|^2\, \nu(dy)ds\Big] \\ &\leq p^2 C_{T-T_0}^{2(p-1)}\|f\|_p^{2(p-1)}\mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |P_{T-s}f(X_{s-} + y) - P_{T-s}f(X_{s-})|^2\, \nu(dy)ds\Big] < \infty\end{aligned}$$

for $t \in [0, T_0]$, which implies that $\int_0^t \int_{\mathbb{R}^d} (\varphi(Y_{s-} + H(s,y)) - \varphi(Y_{s-}))\, \widetilde{N}(ds, dy)$ is a martingale for $t \in [0, T_0]$. Note that

$$\begin{aligned}
&\varphi(Y_{s-} + H(s,y)) - \varphi(Y_{s-}) - H(s,y) \cdot \nabla\varphi(Y_{s-}) \\
&\quad = |P_{T-s}f(X_{s-} + y)|^p - |P_{T-s}f(X_{s-})|^p \\
&\qquad - pP_{T-s}f(X_{s-})|P_{T-s}f(X_{s-})|^{p-2}(P_{T-s}f(X_{s-} + y) - P_{T-s}f(X_{s-})) \\
&\quad = F(P_{T-s}f(X_{s-} + y), P_{T-s}f(X_{s-}); p).
\end{aligned}$$

Putting $t = T_0$ and taking the expectation of both sides in (8.3.10), we have

$$\mathbb{E}^x|Y_{T_0}|^p - \mathbb{E}^x|Y_0|^p = \mathbb{E}^x\Big[\int_0^{T_0} \int_{\mathbb{R}^d} F(P_{T-s}f(X_{s-} + y), P_{T-s}f(X_{s-}); p)\, \nu(dy)ds\Big]. \tag{8.3.12}$$

Integrating both sides in (8.3.12), we see

$$\begin{aligned}
&\|P_{T-T_0}f\|_p^p - \|P_Tf\|_p^p \\
&\quad = \int_{\mathbb{R}^d} \mathbb{E}^x\Big[\int_0^{T_0} \int_{\mathbb{R}^d} F(P_{T-s}f(X_{s-} + y), P_{T-s}f(X_{s-}); p)\, \nu(dy)ds\Big]\, dx \\
&\quad = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_0^{T_0} \int_{\mathbb{R}^d} F(P_{T-s}f(z + y), P_{T-s}f(z); p) p_s(x,z)\, \nu(dy)dsdzdx \\
&\quad = \int_{\mathbb{R}^d} \int_0^{T_0} \int_{\mathbb{R}^d} F(P_{T-s}f(z + y), P_{T-s}f(z); p)\, \nu(dy)dsdz \\
&\quad = \int_{\mathbb{R}^d} \int_{T-T_0}^{T} \int_{\mathbb{R}^d} F(P_sf(z + y), P_sf(z); p)\, \nu(dy)dsdz.
\end{aligned}$$

First, we let $T \to T_0$. Since $F(a, b; p)$ is nonnegative, we have

$$\begin{aligned}
&\lim_{T \to T_0} \int_{\mathbb{R}^d} \int_{T-T_0}^{T} \int_{\mathbb{R}^d} F(P_sf(z + y), P_sf(z); p)\, \nu(dy)dsdz \\
&\qquad\qquad = \int_{\mathbb{R}^d} \int_0^{T} \int_{\mathbb{R}^d} F(P_sf(z + y), P_sf(z); p)\, \nu(dy)dsdz.
\end{aligned}$$

We claim that $\|P_{T-T_0}f\|_p \to \|f\|_p$ as $T \to T_0$. It suffices to show that $\|P_tf - f\|_p \to 0$ as $t \to 0$. Let $\varepsilon > 0$. Using the continuity of the translation operator on $L^p(\mathbb{R}^d)$, we choose $\delta > 0$ small enough such that $\|T_yf - f\|_p^p < \varepsilon$ where $T_yf(x) = f(x + y)$. By (8.2.2), there exists $t_0 > 0$ such that for all $t \in [0, t_0]$

$$\mathbb{P}^0(|X_t| > \delta) = \int_{|y| > \delta} p_t(y)\, dy < \varepsilon.$$

For $0 \le t \le t_0$, we get

$$\begin{aligned}\|P_t f - f\|_p^p &\le \iint |f(x+y) - f(x)|^p p_t(y)\, dydx \\ &\le 2^{p-1}\|f\|_p^p \int_{|y|>\delta} p_t(y)\, dy + \int_{|y|\le\delta} \|T_y f - f\|_p^p p_t(y)\, dy \\ &\le (2^{p-1}\|f\|_p^p + 1)\varepsilon,\end{aligned}$$

which proves the claim and yields

$$\|f\|_p^p - \|P_T f\|_p^p = \int_{\mathbb{R}^d}\int_0^T \int_{\mathbb{R}^d} F(P_s f(z+y), P_s f(z); p)\, \nu(dy) ds dz.$$

Let $f^*(x) = \sup_t |P_t f(x)|$, then it follows from Proposition 7.3.3 that $\|f^*\|_p \le \frac{p}{p-1}\|f\|_p$. Since $|P_T f(x)| \le |f^*(x)|$ and $P_T f(x) \to 0$, as $T \to \infty$ for each $x \in \mathbb{R}^d$ by Lemma 8.3.4, the dominated convergence theorem yields $\|P_T f\|_p \to 0$ as $T \to \infty$. Since $F(a, b; p)$ is nonnegative, we have

$$\|f\|_p^p = \int_{\mathbb{R}^d}\int_0^\infty \int_{\mathbb{R}^d} F(P_s f(z+y), P_s f(z); p)\, \nu(dy) ds dz \tag{8.3.13}$$

as desired.

Let $1 < p < 2$ and $\varepsilon > 0$. Following the same argument as in the case $p > 2$ with the function $\varphi(x) = (|x|^2 + \varepsilon^2)^{\frac{p}{2}}$, we arrive at

$$\begin{aligned}\int_{\mathbb{R}^d} \left((|f(x)|^2 + \varepsilon^2)^{\frac{p}{2}} - (|P_T f(x)|^2 + \varepsilon^2)^{\frac{p}{2}}\right) dx \\ = \int_{\mathbb{R}^d}\int_0^T \int_{\mathbb{R}^d} F_\varepsilon(P_s f(z+y), P_s f(z); p)\, \nu(dy) ds dz,\end{aligned}$$

where F_ε is the function defined by (8.3.2). Since the function $x \mapsto x^{\frac{p}{2}}$ is $\frac{p}{2}$-Hölder continuous on $[0, \infty)$ for $1 < p < 2$, we have $(|f(x)|^2 + \varepsilon^2)^{\frac{p}{2}} - \varepsilon^p \le C_p |f(x)|^p$ and $(|P_T f(x)|^2 + \varepsilon^2)^{\frac{p}{2}} - \varepsilon^p \le C_p |P_T f(x)|^p$. Thus the left hand side converges to $\|f\|_p^p - \|P_T f\|_p^p$ as $\varepsilon \to 0$ by the dominated convergence theorem. On the other hand, $0 \le F_\varepsilon(a, b; p) \to F(a, b; p)$, as $\varepsilon \to 0$, and $0 \le F_\varepsilon(a, b; p) \le \frac{1}{p-1} F(a, b; p)$, by Lemma 8.3.2. Since the integral

$$I(\varepsilon, T) = \int_{\mathbb{R}^d}\int_0^T \int_{\mathbb{R}^d} F_\varepsilon(P_s f(z+y), P_s f(z); p)\, \nu(dy) ds dz$$

is bounded for each $\varepsilon > 0$, Fatou's lemma and the dominated convergence theorem give (see [30, p.199]) that

$$\lim_{\varepsilon \to 0} I(\varepsilon, T) = \int_{\mathbb{R}^d} \int_0^T \int_{\mathbb{R}^d} F(P_s f(z+y), P_s f(z); p)\, \nu(dy) ds dz.$$

We finish the proof by letting $T \to \infty$. □

Following the same argument, we obtain a more general result for martingales of which Theorem 8.3.1 is a special case.

Theorem 8.3.5 (A Hardy–Stein identity for martingales)**.** *Let $1 < p < \infty$ and $H(t,x)$ be a d-dimensional predictable process satisfying (8.2.5) and (8.2.7). Assume that a martingale M_t defined by*

$$M_t = M_0 + \int_0^t \int_{\mathbb{R}^d} H(s,y)\, \widetilde{N}(ds, dy)$$

is uniformly integrable in $L^2 \cap L^p$, that is,

$$\sup_{t>0} \mathbb{E}|M_t|^{\max\{2,p\}} < \infty.$$

Then we have

$$\mathbb{E}|M_\infty|^p - \mathbb{E}|M_0|^p = \int_0^\infty \int_{\mathbb{R}^d} \mathbb{E}[F(M_{s-}, M_{s-} + H(s,y); p)]\, \nu(dy) ds. \tag{8.3.14}$$

Proof. Let $p \geq 2$. Let $T > 0$ and $\varphi(x) = |x|^p$. By Itô's formula, we have

$$\begin{aligned}\varphi(M_t) - \varphi(M_0) &= \int_0^t \int_{\mathbb{R}^d} (\varphi(M_{s-} + H(s,y)) - \varphi(M_{s-}))\, \widetilde{N}(ds, dy) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} (\varphi(M_{s-} + H(s,y)) - \varphi(M_{s-}) - H(s,y) \cdot \nabla\varphi(M_{s-}))\, \nu(dy) ds\end{aligned}$$

for $0 \leq t \leq T$. Since

$$\begin{aligned}|\varphi(M_{s-} + H(s,y)) - \varphi(M_{s-})|^2 &\leq p^2 2^{2p-3} (|M_{s-}|^{2p-2} + |H(s,y)|^{2p-2}) |H(s,y)|^2 \\ &\leq C(p,T) |H(s,y)|^2\end{aligned}$$

for $0 \le s \le T$, we get

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |\varphi(M_{s-} + H(s,y)) - \varphi(M_{s-})|^2 \, \nu(dy)ds\Big] \\
&\quad \le C(p,T)\mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |H(s,y)|^2 \, \nu(dy)ds\Big] \\
&\quad = C(p,T)\mathbb{E}|M_t - M_0|^2 \\
&\quad < \infty,
\end{aligned}$$

which implies that $\int_0^t \int_{\mathbb{R}^d} (\varphi(M_{s-} + H(s,y)) - \varphi(M_{s-})) \, \widetilde{N}(ds,dy)$ is a martingale and its expectation is zero. Since $F(a,b;p) \ge 0$ and ν is σ-finite, it follows from Fubini–Tonelli theorem that

$$\begin{aligned}
\mathbb{E}|M_t|^p - \mathbb{E}|M_0|^p &= \int_0^t \int_{\mathbb{R}^d} \mathbb{E}[\varphi(M_{s-} + H(s,y)) - \varphi(M_{s-}) - H(s,y) \cdot \nabla\varphi(M_{s-})] \, \nu(dy)ds \\
&= \int_0^t \int_{\mathbb{R}^d} \mathbb{E}[F(M_{s-}, M_{s-} + H(s,y); p)] \, \nu(dy)ds.
\end{aligned}$$

Letting $t \to \infty$, we get the result.

Suppose $1 < p < 2$. Let $\varepsilon > 0$, $T > 0$, and $\varphi_\varepsilon(t) = (t^2 + \varepsilon^2)^{\frac{p}{2}}$. By Itô's formula, we have

$$\mathbb{E}[\varphi_\varepsilon(M_t)] - \mathbb{E}[\varphi_\varepsilon(M_0)] = \int_0^t \int_{\mathbb{R}^d} \mathbb{E}[F_\varepsilon(M_{s-}, M_{s-} + H(s,y); p)] \, \nu(dy)ds$$

for $0 \le t \le T$. Since $\varphi_\varepsilon(t) \le C_p|t|^p + \varepsilon^p$ and $\sup_{t \ge 0} \mathbb{E}|M_t|^p < \infty$, we have

$$\lim_{\varepsilon \to 0}(\mathbb{E}[\varphi_\varepsilon(M_t)] - \mathbb{E}[\varphi_\varepsilon(M_0)]) = \mathbb{E}|M_t|^p - \mathbb{E}|M_0|^p.$$

Let $I(\varepsilon,t) = \int_0^t \int_{\mathbb{R}^d} \mathbb{E}[F_\varepsilon(M_{s-}, M_{s-} + H(s,y); p)] \, \nu(dy)ds$, then $\liminf_{\varepsilon \to 0} I(\varepsilon,t) < \infty$ by Lemma 8.3.2. Using Fatou's lemma, we have $I(0,t) < \infty$. Thus the result follows from $\varphi_\varepsilon(t) \le C_p|t|^p + \varepsilon^p$ and the dominated convergence theorem. □

8.4 Fourier multipliers and square functions

The main application of the results in [10] was to show the L^p boundedness of the Fourier multipliers introduced in [8], $1 < p < \infty$, without appealing to martingale

transforms. Of course, a disadvantage of such a proof is that we do not obtain the sharp bounds given in [8, 9], which follow from Burkholder's sharp inequalities. In addition, the Littlewood–Paley inequalities proved in [10] only apply to symmetric pure jump Lévy processes and therefore the Fourier multiplier proof given there also has this restriction. In this section, we prove, via a symmetrization of the Littlewood–Paley inequalities, the general result for Fourier multipliers.

We recall that $(X_t)_{t\geq 0}$ is a pure jump Lévy process with càdlàc path and ν is its Lévy measure that satisfies the *Hartman–Wintner condition*

$$\lim_{|\xi|\to\infty} \frac{\operatorname{Re}(\psi(\xi))}{\log(1+|\xi|)} = \infty. \tag{HW}$$

Let P_t be a semigroup defined by $P_t f(x) = \mathbb{E}^x[f(X_t)]$. Let $\phi : (0,\infty)\times\mathbb{R}^d \to \mathbb{R}$ be a bounded measurable function and $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let $m : \mathbb{R}^d \to \mathbb{C}$ be a measurable function. The Fourier multiplier operator with symbol m is denoted by T_m. Note that T_m is determined by $\mathcal{F}(T_m f)(\xi) = m(\xi)\widehat{f}(\xi)$. For $f, g \in L^2(\mathbb{R}^d)$, we denote by $\langle f, g\rangle = \int_{\mathbb{R}^d} f g\, dx$. By Parseval's formula (8.2.1), we have

$$\begin{aligned}
\langle T_m f, g\rangle &= \int_{\mathbb{R}^d} T_m f(x) g(x)\, dx \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathcal{F}(T_m f)(\xi)\overline{\mathcal{F}(g)(\xi)}\, d\xi \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} m(\xi)\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\, d\xi.
\end{aligned}$$

We are ready to state our result on Fourier multipliers.

Theorem 8.4.1. *Let $\phi : (0,\infty)\times\mathbb{R}^d \to \mathbb{R}$ be a bounded measurable function, $p \in (1,\infty)$, and q the conjugate exponent of p. Then for $f \in L^2(\mathbb{R}^d)\cap L^p(\mathbb{R}^d)$ and $g \in L^2(\mathbb{R}^d)\cap L^q(\mathbb{R}^d)$,*

$$\Lambda_\phi(f,g) = \int_{\mathbb{R}^d}\int_0^\infty \int_{\mathbb{R}^d} (P_t f(x+y) - P_t f(x))(P_t g(x+y) - P_t g(x))\phi(t,y)\,\nu(dy)dtdx \tag{8.4.1}$$

is well-defined. Furthermore, there is a unique bounded linear operator S_ϕ on $L^p(\mathbb{R}^d)$ such that $\Lambda_\phi(f,g) = \langle S_\phi(f), g\rangle$ and $S_\phi = T_{m_\phi}$ with symbol m_ϕ given by

$$m_\phi(\xi) = \int_0^\infty \int_{\mathbb{R}^d} |e^{i\xi\cdot y} - 1|^2 e^{-2t\operatorname{Re}(\psi(\xi))}\phi(t,y)\,\nu(dy)dt.$$

When ν is symmetric, this result was proved in [10] as an application of the boundedness on L^p of the Littlewood–Paley square functions which itself was the main application of the Hardy–Stein inequality, completely bypassing the martingale transform arguments used earlier. The question left open in [10] was whether Littlewood–Paley arguments can be used to prove the result for general ν. We answer this in the affirmative.

Let us introduce the dual process and the symmetrization of the Lévy process $(X_t)_{t\geq 0}$ with the Lévy measure ν. Let $(\widehat{X}_t)_{t\geq 0}$ be a càdlàg stochastic process having the same finite dimensional distribution as $(-X_t)_{t\geq 0}$, and independent of $(X_t)_{t\geq 0}$. The process $(\widehat{X}_t)_{t\geq 0}$ is said to be the *dual process* of $(X_t)_{t\geq 0}$. Note that $(\widehat{X}_t)_{t\geq 0}$ is a Lévy process with triplet $(0, 0, \nu(-dx))$. We define its semigroup by $\widehat{P}_t f(x) = \mathbb{E}^x[f(\widehat{X}_t)]$. Note that for any Borel function f and g, we have

$$\int_{\mathbb{R}^d} P_t f(x) g(x)\, dx = \int_{\mathbb{R}^d} f(x) \widehat{P}_t g(x)\, dx,$$

which explains why $(\widehat{X}_t)_{t\geq 0}$ is called the dual of $(X_t)_{t\geq 0}$.

Let $\widetilde{X}_t = X_{\frac{t}{2}} + \widehat{X}_{\frac{t}{2}}$ for $t \geq 0$. We define $\widetilde{\psi}(\xi) = \mathrm{Re}(\psi(\xi))$ and $\widetilde{\nu}(B) = \frac{1}{2}(\nu(B) + \nu(-B))$ for any measurable set B in $\mathbb{R}^d$. Since we have

$$\mathbb{E}[e^{i\xi\cdot\widetilde{X}_t}] = \mathbb{E}[e^{i\xi\cdot X_{\frac{t}{2}}}]\mathbb{E}[e^{i\xi\cdot\widehat{X}_{\frac{t}{2}}}] = e^{-\frac{t}{2}\psi(\xi)} e^{-\frac{t}{2}\psi(-\xi)} = e^{-t\widetilde{\psi}(\xi)}$$

and

$$\begin{aligned}
\widetilde{\psi}(\xi) &= \int_{\mathbb{R}^d} (1 - \cos(\xi\cdot y))\, \nu(dy) \\
&= \int_{\mathbb{R}^d} (1 - \cos(\xi\cdot y))\, \widetilde{\nu}(dy) \\
&= \int_{\mathbb{R}^d} (1 - e^{i\xi\cdot y} + i\xi\cdot y \mathbb{1}_{\{|y|\leq 1\}})\, \widetilde{\nu}(dy),
\end{aligned}$$

the process $\widetilde{X}_t$ is a Lévy process with characteristic exponent $\widetilde{\psi}(\xi)$ and the Lévy measure $\widetilde{\nu}$. We say that $\widetilde{X}_t$ is the symmetrization of X_t. Define $\widetilde{P}_t f(x) = \mathbb{E}^x[f(\widetilde{X}_t)]$. The Fourier transform of $\widetilde{P}_t f$ is given by

$$\mathcal{F}(\widetilde{P}_t f)(\xi) = e^{-t\widetilde{\psi}(\xi)}\, \widehat{f}(\xi) = e^{-t\,\mathrm{Re}(\psi(\xi))}\, \widehat{f}(\xi).$$

Since $\widetilde{X}_t$ is a symmetric pure jump Lévy process and the measure $\widetilde{\nu}$ satisfies (HW) condition, it leads us to apply the result of [10] for the symmetrization $\widetilde{X}_t$. In particular, we obtain two side estimates for the square functions of $\widetilde{X}_t$. We define the square functions of the symmetrized process $\widetilde{X}_t$ by

$$\widetilde{G}(f)(x) = \Big(\int_0^\infty \int_{\mathbb{R}^d} \left| \widetilde{P}_t f(x+y) - \widetilde{P}_t f(x) \right|^2 \widetilde{\nu}(dy)dt \Big)^{\frac{1}{2}}, \tag{8.4.2}$$

$$\widetilde{G}_*(f)(x) = \Big(\int_0^\infty \int_{A(t,x,f)} \left| \widetilde{P}_t f(x+y) - \widetilde{P}_t f(x) \right|^2 \widetilde{\nu}(dy)dt \Big)^{\frac{1}{2}} \tag{8.4.3}$$

where $A(t,x,f) = \{y \in \mathbb{R}^d : |\widetilde{P}_t f(x)| > |\widetilde{P}_t f(x+y)|\}$. The following lemma is found in [10, Theorem 4.1, Corollary 4.4].

Lemma 8.4.2. *Let $2 \le p < \infty$ and $f \in L^p(\mathbb{R}^d)$. Then there are constants c_p and C_p depending only on p such that*

$$c_p \|f\|_p \le \|\widetilde{G}(f)\|_p \le C_p \|f\|_p.$$

If $1 < p < \infty$ and $f \in L^p(\mathbb{R}^d)$, then we have

$$d_p \|f\|_p \le \|\widetilde{G}_*(f)\|_p \le D_p \|f\|_p,$$

for some d_p and D_p depending only on p.

For a function f and a measure μ, the essential supremum of f with respect to the measure μ is denoted by $\|f\|_{\infty,\mu}$.

Lemma 8.4.3. *Let $\overline{\nu}(B) = \frac{1}{2}(\nu(B) - \nu(-B))$ for any measurable set $B \subseteq \mathbb{R}^d$. Then, there is a measurable function $r(y)$ such that $\overline{\nu}(dy) = r(y)\widetilde{\nu}(dy)$. Furthermore, the function $r(y)$ is bounded $\widetilde{\nu}$-a.s. with $\|r\|_{\infty,\widetilde{\nu}} \le 1$.*

Proof. Note that ν is σ-finite since $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(1 \wedge |x|^2)\nu(dx) < \infty$. So are $\widetilde{\nu}$ and $\overline{\nu}$. Suppose that $B \subseteq \mathbb{R}^d$ is a measurable set such that $\widetilde{\nu}(B) = 0$. Since ν is a positive measure, we have $\nu(B) = \nu(-B) = 0$, which implies $\overline{\nu}(B) = 0$. Thus $\overline{\nu}$ is absolutely continuous with respect to $\widetilde{\nu}$. By the Radon-Nikodym theorem, we conclude that there is a measurable function $r(y)$ such that $\overline{\nu}(dy) = r(y)\widetilde{\nu}(y)$.

To see $r(y)$ is bounded, we consider the set $B^\varepsilon := \{y \in \mathbb{R}^d : |r(y)| > 1+\varepsilon\}$ for an arbitrary $\varepsilon > 0$. From the relation $\overline{\nu}(dy) = r(y)\widetilde{\nu}(y)$ obtained above, we have $\overline{\nu}(B^\varepsilon) > (1+\varepsilon)\widetilde{\nu}(B^\varepsilon)$. It then yields $\varepsilon\nu(B^\varepsilon) + (2+\varepsilon)\nu(-B^\varepsilon) < 0$ so that $\nu(B^\varepsilon) = \nu(-B^\varepsilon) = 0$. Therefore, $r(y)$ is bounded $\widetilde{\nu}$-a.s. and $\|r\|_{\infty,\widetilde{\nu}} \le 1$. □

Proof of Theorem 8.4.1. The first argument is directly obtained by Theorem 8.3.1. Indeed, since $F(a,b;2) = |a-b|^2$, Theorem 8.3.1 yields that

$$\begin{aligned}\|f\|_2^2 &= \int_{\mathbb{R}^d}\int_0^\infty\int_{\mathbb{R}^d} F(P_tf(x), P_tf(x+y);2)\,\nu(dy)dtdx \\ &= \int_{\mathbb{R}^d}\int_0^\infty\int_{\mathbb{R}^d} |P_tf(x)-P_tf(x+y)|^2\,\nu(dy)dtdx.\end{aligned}$$

It then follows from the Cauchy-Schwartz inequality that

$$\begin{aligned}|\Lambda_\phi(f,g)| &\le \|\phi\|_\infty \int_{\mathbb{R}^d}\int_0^\infty\int_{\mathbb{R}^d} |P_tf(x+y)-P_tf(x)||P_tg(x+y)-P_tg(x)|\,\nu(dy)dtdx \\ &\le \|\phi\|_\infty \|f\|_2\|g\|_2.\end{aligned}$$

Since $f, g \in L^2(\mathbb{R}^d)$, Theorem 8.3.1 implies that $\Lambda_\phi(f,g)$ is absolutely convergent. To see the second assertion, we use Parseval's formula (8.2.1) so that

$$\begin{aligned}&\Lambda_\phi(f,g) \\ &= \frac{1}{(2\pi)^d}\iiint \mathcal{F}(P_tf(\cdot+y)-P_tf(\cdot))(\xi)\overline{\mathcal{F}(P_tg(\cdot+y)-P_tg(\cdot))(\xi)}\phi(t,y)\,d\xi\nu(dy)dt \\ &= \frac{1}{(2\pi)^d}\iiint (e^{i\xi\cdot y}-1)e^{-t\psi(\xi)}\widehat{f}(\xi)\overline{(e^{i\xi\cdot y}-1)e^{-t\psi(\xi)}\widehat{g}(\xi)}\phi(t,y)\,d\xi\nu(dy)dt \\ &= \frac{1}{(2\pi)^d}\iiint |e^{i\xi\cdot y}-1|^2 e^{-2t\,\mathrm{Re}(\psi(\xi))}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\phi(t,y)\,d\xi\nu(dy)dt\end{aligned}$$

where $\psi(\xi)$ is the characteristic exponent of $(X_t)_{t\ge0}$. In the second equality, we have used the fact that

$$\mathcal{F}(P_tf(\cdot+y)-P_tf(\cdot))(\xi) = (e^{i\xi\cdot y}-1)\mathcal{F}(P_tf)(\xi) = (e^{i\xi\cdot y}-1)e^{-t\psi(\xi)}\widehat{f}(\xi).$$

By Lemma 8.4.3, there is a measurable function $r(y)$ such that $\overline{\nu}(dy) = r(y)\widetilde{\nu}(dy)$ with $\|r\|_{\infty,\widetilde{\nu}} \le 1$. Using $\nu = \widetilde{\nu} + \overline{\nu}$, we have

$$\begin{aligned}\Lambda_\phi(f,g) = &\frac{1}{(2\pi)^d}\int_0^\infty\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |e^{i\xi\cdot y}-1|^2 e^{-2t\,\mathrm{Re}(\psi(\xi))}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\phi(t,y)\,d\xi\widetilde{\nu}(dy)dt \\ &+ \frac{1}{(2\pi)^d}\int_0^\infty\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |e^{i\xi\cdot y}-1|^2 e^{-2t\,\mathrm{Re}(\psi(\xi))}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\phi(t,y)\,d\xi\overline{\nu}(dy)dt.\end{aligned}$$

If we define $\eta(t,y)=\phi(t,y)(1+r(y))$, then η is bounded $\widetilde{\nu}$-a.s.; thus, we obtain

$$\Lambda_\phi(f,g)=\frac{1}{(2\pi)^d}\int_0^\infty\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|e^{i\xi\cdot y}-1|^2e^{-2t\,\mathrm{Re}(\psi(\xi))}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\eta(t,y)\,d\xi\widetilde{\nu}(dy)dt.$$

We consider $\widetilde{X}_t$ and $\widetilde{P}_t$, the symmetrization of X_t and P_t. Since the characteristic exponent of $\widetilde{X}_t$ is the real part of $\psi(\xi)$, $\widetilde{\psi}(\xi)=\mathrm{Re}(\psi(\xi))$, and

$$\mathcal{F}(\widetilde{P}_tf(\cdot+y)-\widetilde{P}_tf(\cdot))(\xi)=(e^{i\xi\cdot y}-1)\mathcal{F}(\widetilde{P}_tf)(\xi)=(e^{i\xi\cdot y}-1)e^{-t\,\mathrm{Re}(\psi(\xi))}\widehat{f}(\xi),$$

it follows from Parseval's formula (8.2.1) that

$$\begin{aligned}\Lambda_\phi(f,g)&=\int_{\mathbb{R}^d}\int_0^\infty\int_{\mathbb{R}^d}(\widetilde{P}_tf(x+y)-\widetilde{P}_tf(x))(\widetilde{P}_tg(x+y)-\widetilde{P}_tg(x))\eta(t,y)\,\widetilde{\nu}(dy)dtdx\\&=:\widetilde{\Lambda}_\eta(f,g).\end{aligned}\tag{8.4.4}$$

To show the boundedness of $\Lambda_\phi(f,g)$, we use the square functions defined in (8.4.2). It is enough to show the case $p>2$ and $1<q<2$. Note that $\|\eta\|_{\infty,\widetilde{\nu}}$ is finite and $\|\eta\|_{\infty,\widetilde{\nu}}\le 2\|\phi\|_\infty$. Let $A(t,x,g):=\{y\in\mathbb{R}^d:|\widetilde{P}_tg(x)|>|\widetilde{P}_tg(x+y)|\}$. Note that it follows from the symmetry of $\widetilde{\nu}$ that

$$\begin{aligned}&\int_{\mathbb{R}^d}\int_0^\infty\int_{A(t,x,g)}|\widetilde{P}_tf(x+y)-\widetilde{P}_tf(x)||\widetilde{P}_tg(x+y)-\widetilde{P}_tg(x)|\,\widetilde{\nu}(dy)dtdx\\&\quad=\int_{\mathbb{R}^d}\int_0^\infty\int_{\mathbb{R}^d\setminus A(t,x,g)}|\widetilde{P}_tf(x+y)-\widetilde{P}_tf(x)||\widetilde{P}_tg(x+y)-\widetilde{P}_tg(x)|\,\widetilde{\nu}(dy)dtdx.\end{aligned}$$

Applying Cauchy-Schwartz and Hölder's inequalities, we have

$$\begin{aligned}&|\widetilde{\Lambda}_\eta(f,g)|\\&\le\|\eta\|_{\infty,\widetilde{\nu}}\int_{\mathbb{R}^d}\int_0^\infty\int_{\mathbb{R}^d}|\widetilde{P}_tf(x+y)-\widetilde{P}_tf(x)||\widetilde{P}_tg(x+y)-\widetilde{P}_tg(x)|\,\widetilde{\nu}(dy)dtdx\\&\le 2\|\eta\|_{\infty,\widetilde{\nu}}\int_{\mathbb{R}^d}\int_0^\infty\int_{A(t,x,g)}|\widetilde{P}_tf(x+y)-\widetilde{P}_tf(x)||\widetilde{P}_tg(x+y)-\widetilde{P}_tg(x)|\,\widetilde{\nu}(dy)dtdx\\&\le 2\|\eta\|_{\infty,\widetilde{\nu}}\int_{\mathbb{R}^d}\widetilde{G}(f)(x)\widetilde{G}_*(g)(x)\,dx\\&\le 2\|\eta\|_{\infty,\widetilde{\nu}}\|\widetilde{G}(f)\|_p\|\widetilde{G}_*(g)\|_q.\end{aligned}$$

It follows from Lemma 8.4.2 and (8.4.4) that

$$\Lambda_\phi(f,g)\le 4C_pD_q\|\phi\|_\infty\|f\|_p\|g\|_q.$$

Therefore, the Riesz representation theorem yields that there is a unique linear operator S_ϕ satisfying $\Lambda_\phi(f, g) = \langle S_\phi(f), g\rangle$. $\square$

REFERENCES

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, second ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008. MR 2401600

[2] David Applebaum, *Lévy processes and stochastic calculus*, second ed., Cambridge Studies in Advanced Mathematics, vol. 116, Cambridge University Press, Cambridge, 2009. MR 2512800

[3] David Applebaum and Rodrigo Bañuelos, *Probabilistic approach to fractional integrals and the Hardy-Littlewood-Sobolev inequality*, Analytic methods in interdisciplinary applications, Springer Proc. Math. Stat., vol. 116, Springer, Cham, 2015, pp. 17–40. MR 3329498

[4] K. I. Babenko, *An inequality in the theory of Fourier integrals*, Izv. Akad. Nauk SSSR Ser. Mat. **25** (1961), 531–542. MR 0138939

[5] Rodrigo Bañuelos, *Martingale transforms and related singular integrals*, Trans. Amer. Math. Soc. **293** (1986), no. 2, 547–563. MR 816309

[6] ———, *On an estimate of Cranston and McConnell for elliptic diffusions in uniform domains*, Probab. Theory Related Fields **76** (1987), no. 3, 311–323. MR 912657

[7] ———, *The foundational inequalities of D. L. Burkholder and some of their ramifications*, Illinois J. Math. **54** (2010), no. 3, 789–868 (2012). MR 2928339

[8] Rodrigo Bañuelos, Adam Bielaszewski, and Krzysztof Bogdan, *Fourier multipliers for non-symmetric Lévy processes*, Marcinkiewicz centenary volume, Banach Center Publ., vol. 95, Polish Acad. Sci. Inst. Math., Warsaw, 2011, pp. 9–25. MR 2918086

[9] Rodrigo Bañuelos and Krzysztof Bogdan, *Lévy processes and Fourier multipliers*, J. Funct. Anal. **250** (2007), no. 1, 197–213. MR 2345912

[10] Rodrigo Bañuelos, Krzysztof Bogdan, and Tomasz Luks, *Hardy-Stein identities and square functions for semigroups*, J. Lond. Math. Soc. (2) **94** (2016), no. 2, 462–478. MR 3556449

[11] Rodrigo Bañuelos and Daesung Kim, *Hardy–Stein identity for non-symmetric Lévy processes and Fourier multipliers*, arXiv preprint arXiv:1702.06573v2 (2017).

[12] Rodrigo Bañuelos and Tadeusz Kulczycki, *The Cauchy process and the Steklov problem*, J. Funct. Anal. **211** (2004), no. 2, 355–423. MR 2056835

[13] _____, *Eigenvalue gaps for the Cauchy process and a Poincaré inequality*, J. Funct. Anal. **234** (2006), no. 1, 199–225. MR 2214145

[14] _____, *Spectral gap for the Cauchy process on convex, symmetric domains*, Comm. Partial Differential Equations **31** (2006), no. 10-12, 1841–1878. MR 2273977

[15] Rodrigo Bañuelos and Pedro J. Méndez-Hernández, *Space-time Brownian motion and the Beurling-Ahlfors transform*, Indiana Univ. Math. J. **52** (2003), no. 4, 981–990. MR 2001941

[16] _____, *Symmetrization of Lévy processes and applications*, J. Funct. Anal. **258** (2010), no. 12, 4026–4051. MR 2609537

[17] Rodrigo Bañuelos and Adam Osękowski, *Sharp martingale inequalities and applications to Riesz transforms on manifolds, Lie groups and Gauss space*, J. Funct. Anal. **269** (2015), no. 6, 1652–1713. MR 3373431

[18] Franck Barthe and Alexander V. Kolesnikov, *Mass transport and variants of the logarithmic Sobolev inequality*, J. Geom. Anal. **18** (2008), no. 4, 921–979. MR 2438906

[19] Fabrice Baudoin, *Diffusion processes and stochastic calculus*, EMS Textbooks in Mathematics, European Mathematical Society (EMS), Zürich, 2014. MR 3236242

[20] William Beckner, *Inequalities in Fourier analysis*, Ann. of Math. (2) **102** (1975), no. 1, 159–182. MR 0385456

[21] _____, *Pitt's inequality and the uncertainty principle*, Proc. Amer. Math. Soc. **123** (1995), no. 6, 1897–1905. MR 1254832

[22] Jean Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996. MR 1406564

[23] Gabriele Bianchi and Henrik Egnell, *A note on the Sobolev inequality*, J. Funct. Anal. **100** (1991), no. 1, 18–24. MR 1124290

[24] S. G. Bobkov, G. P. Chistyakov, and F. Götze, *Entropic instability of Cramer's characterization of the normal law*, Selected works of Willem van Zwet, Sel. Works Probab. Stat., Springer, New York, 2012, pp. 231–242. MR 2918082

[25] _____, *Regularized distributions and entropic stability of Cramer's characterization of the normal law*, Stochastic Process. Appl. **126** (2016), no. 12, 3865–3887. MR 3565483

[26] S. G. Bobkov, N. Gozlan, C. Roberto, and P.-M. Samson, *Bounds on the deficit in the logarithmic Sobolev inequality*, J. Funct. Anal. **267** (2014), no. 11, 4110–4138. MR 3269872

[27] S. G. Bobkov and M. Ledoux, *From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities*, Geom. Funct. Anal. **10** (2000), no. 5, 1028–1052. MR 1800062

[28] Salomon Bochner, *Harmonic analysis and the theory of probability*, University of California Press, Berkeley and Los Angeles, 1955. MR 0072370

[29] Krzysztof Bogdan, Tomasz Byczkowski, Tadeusz Kulczycki, Michal Ryznar, Renming Song, and Zoran Vondraček, *Potential analysis of stable processes and its extensions*, Lecture Notes in Mathematics, vol. 1980, Springer-Verlag, Berlin, 2009, Edited by Piotr Graczyk and Andrzej Stos. MR 2569321

[30] Krzysztof Bogdan, Bartłomiej Dyda, and Tomasz Luks, *On Hardy spaces of local and nonlocal operators*, Hiroshima Math. J. **44** (2014), no. 2, 193–215. MR 3251822

[31] H. J. Brascamp, Elliott H. Lieb, and J. M. Luttinger, *A general rearrangement inequality for multiple integrals*, J. Functional Analysis **17** (1974), 227–237. MR 0346109

[32] Lorenzo Brasco, Eleonora Cinti, and Stefano Vita, *A quantitative stability estimate for the fractional Faber-Krahn inequality*, arXiv preprint arXiv:1901.10845 (2019).

[33] Lorenzo Brasco and Guido De Philippis, *Spectral inequalities in quantitative form*, Shape optimization and spectral theory, De Gruyter Open, Warsaw, 2017, pp. 201–281. MR 3681151

[34] Lorenzo Brasco, Guido De Philippis, and Bozhidar Velichkov, *Faber-Krahn inequalities in sharp quantitative form*, Duke Math. J. **164** (2015), no. 9, 1777–1831. MR 3357184

[35] L. Brasco, E. Lindgren, and E. Parini, *The fractional Cheeger problem*, Interfaces Free Bound. **16** (2014), no. 3, 419–458. MR 3264796

[36] Yann Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417. MR 1100809

[37] D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Probab. **12** (1984), no. 3, 647–702. MR 744226

[38] Luis Caffarelli, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys. **214** (2000), no. 3, 547–563. MR 1800860

[39] Luis Caffarelli and Luis Silvestre, *An extension problem related to the fractional Laplacian*, Comm. Partial Differential Equations **32** (2007), no. 7-9, 1245–1260. MR 2354493

[40] Eric A. Carlen, *Superadditivity of Fisher's information and logarithmic Sobolev inequalities*, J. Funct. Anal. **101** (1991), no. 1, 194–211. MR 1132315

[41] _____, *Duality and stability for functional inequalities*, Ann. Fac. Sci. Toulouse Math. (6) **26** (2017), no. 2, 319–350. MR 3640893

[42] Eric A. Carlen, José A. Carrillo, and Michael Loss, *Hardy-Littlewood-Sobolev inequalities via fast diffusion flows*, Proc. Natl. Acad. Sci. USA **107** (2010), no. 46, 19696–19701. MR 2745814

[43] E. Carlen and M. Loss, *Competing symmetries, the logarithmic HLS inequality and Onofri's inequality on S^n*, Geom. Funct. Anal. **2** (1992), no. 1, 90–104. MR 1143664

[44] Shibing Chen, Rupert L. Frank, and Tobias Weth, *Remainder terms in the fractional Sobolev inequality*, Indiana Univ. Math. J. **62** (2013), no. 4, 1381–1397. MR 3179693

[45] Michael Christ, *A sharpened Hausdorff-Young inequality*, arXiv preprint arXiv:1406.1210 (2014).

[46] Michael Christ, Heping Liu, and An Zhang, *Sharp Hardy-Littlewood-Sobolev inequalities on quaternionic Heisenberg groups*, Nonlinear Anal. **130** (2016), 361–395. MR 3424625

[47] ———, *Sharp Hardy-Littlewood-Sobolev inequalities on the octonionic Heisenberg group*, Calc. Var. Partial Differential Equations **55** (2016), no. 1, Art. 11, 18. MR 3449396

[48] A. Cianchi, N. Fusco, F. Maggi, and A. Pratelli, *The sharp Sobolev inequality in quantitative form*, J. Eur. Math. Soc. (JEMS) **11** (2009), no. 5, 1105–1139. MR 2538501

[49] Marco Cicalese and Gian Paolo Leonardi, *A selection principle for the sharp quantitative isoperimetric inequality*, Arch. Ration. Mech. Anal. **206** (2012), no. 2, 617–643. MR 2980529

[50] Dario Cordero-Erausquin, *Some applications of mass transport to Gaussian-type inequalities*, Arch. Ration. Mech. Anal. **161** (2002), no. 3, 257–269. MR 1894593

[51] ———, *Transport Inequalities for Log-concave Measures, Quantitative Forms and Applications*, Canad. J. Math. **69** (2017), 481–501.

[52] E. B. Davies, *Heat kernels and spectral theory*, Cambridge Tracts in Mathematics, vol. 92, Cambridge University Press, Cambridge, 1989. MR 990239

[53] Jean Dolbeault and Giuseppe Toscani, *Stability results for logarithmic Sobolev and Gagliardo-Nirenberg inequalities*, Int. Math. Res. Not. IMRN (2016), no. 2, 473–498. MR 3493423

[54] Jingbo Dou and Meijun Zhu, *Sharp Hardy-Littlewood-Sobolev inequality on the upper half space*, Int. Math. Res. Not. IMRN (2015), no. 3, 651–687. MR 3340332

[55] R. Eldan, J. Lehec, and Y. Shenfeld, *Stability of the logarithmic Sobolev inequality via the Föllmer Process*, arXiv preprint arXiv: 1903.04522 (2019).

[56] László Erdös and Horng-Tzer Yau, *A dynamical approach to random matrix theory*, Courant Lecture Notes in Mathematics, vol. 28, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017. MR 3699468

[57] Max Fathi, Emanuel Indrei, and Michel Ledoux, *Quantitative logarithmic Sobolev inequalities and stability estimates*, Discrete Contin. Dyn. Syst. **36** (2016), no. 12, 6835–6853. MR 3567822

[58] F. Feo, E. Indrei, M. R. Posteraro, and C. Roberto, *Some Remarks on the Stability of the Log-Sobolev Inequality for the Gaussian Measure*, Potential Anal. **47** (2017), no. 1, 37–52. MR 3666798

[59] A. Figalli, F. Maggi, and A. Pratelli, *A mass transportation approach to quantitative isoperimetric inequalities*, Invent. Math. **182** (2010), no. 1, 167–211. MR 2672283

[60] Rupert L. Frank and Elliott H. Lieb, *Inversion positivity and the sharp Hardy-Littlewood-Sobolev inequality*, Calc. Var. Partial Differential Equations **39** (2010), no. 1-2, 85–99. MR 2659680

[61] ———, *Sharp constants in several inequalities on the Heisenberg group*, Ann. of Math. (2) **176** (2012), no. 1, 349–381. MR 2925386

[62] Rupert L. Frank and Robert Seiringer, *Non-linear ground state representations and sharp Hardy inequalities*, J. Funct. Anal. **255** (2008), no. 12, 3407–3430. MR 2469027

[63] Bent Fuglede, *Stability in the isoperimetric problem for convex or nearly spherical domains in* $\mathbf{R}^n$, Trans. Amer. Math. Soc. **314** (1989), no. 2, 619–638. MR 942426

[64] N. Fusco, F. Maggi, and A. Pratelli, *The sharp quantitative isoperimetric inequality*, Ann. of Math. (2) **168** (2008), no. 3, 941–980. MR 2456887

[65] Nicola Fusco, Vincent Millot, and Massimiliano Morini, *A quantitative isoperimetric inequality for fractional perimeters*, J. Funct. Anal. **261** (2011), no. 3, 697–715. MR 2799577

[66] Stefan Geiss, Stephen Montgomery-Smith, and Eero Saksman, *On singular integral and martingale transforms*, Trans. Amer. Math. Soc. **362** (2010), no. 2, 553–575. MR 2551497

[67] Alison L. Gibbs and Francis Edward Su, *On choosing and bounding probability metrics*, International Statistical Review **70** (2002), no. 3, 419–435.

[68] L. B. Golinskii and G. P. Chistyakov, *Order-sharp estimates for the stability of decompositions of the normal distribution in the Lévy metric*, Stability problems for stochastic models (Russian) (Moscow, 1991), Vsesoyuz. Nauchno-Issled. Inst. Sistem. Issled., Moscow, 1991, Translated in J. Math. Sci. **7**2 (1994), no. 1, 2848–2871, pp. 16–40. MR 1296264

[69] Leonard Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR 0420249

[70] Richard F. Gundy, *Sur les transformations de Riesz pour le semi-groupe d'Ornstein-Uhlenbeck*, C. R. Acad. Sci. Paris Sér. I Math. **303** (1986), no. 19, 967–970. MR 877182

[71] ———, *Some topics in probability and analysis*, CBMS Regional Conference Series in Mathematics, vol. 70, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1989. MR 995461

[72] Richard F. Gundy and Martin L. Silverstein, *On a probabilistic interpretation for the Riesz transforms*, Functional analysis in Markov processes (Katata/Kyoto, 1981), Lecture Notes in Math., vol. 923, Springer, Berlin-New York, 1982, pp. 199–203. MR 661625

[73] Richard F. Gundy and Nicolas Th. Varopoulos, *Les transformations de Riesz et les intégrales stochastiques*, C. R. Acad. Sci. Paris Sér. A-B **289** (1979), no. 1, A13–A16. MR 545671

[74] R. R. Hall, *A quantitative isoperimetric inequality in n-dimensional space*, J. Reine Angew. Math. **428** (1992), 161–176. MR 1166511

[75] R. R. Hall, W. K. Hayman, and A. W. Weitsman, *On asymmetry and capacity*, J. Analyse Math. **56** (1991), 87–123. MR 1243100

[76] W. Hansen and N. Nadirashvili, *Isoperimetric inequalities in potential theory*, Proceedings from the International Conference on Potential Theory (Amersfoort, 1991), vol. 3, 1994, pp. 1–14. MR 1266215

[77] G. H. Hardy and J. E. Littlewood, *Some properties of fractional integrals. I*, Math. Z. **27** (1928), no. 1, 565–606. MR 1544927

[78] ———, *On Certain Inequalities Connected with the Calculus of Variations*, J. London Math. Soc. **5** (1930), no. 1, 34–39. MR 1574995

[79] Lars Inge Hedberg, *On certain convolution inequalities*, Proc. Amer. Math. Soc. **36** (1972), 505–510. MR 0312232

[80] I. I. Hirschman, Jr., *A note on entropy*, Amer. J. Math. **79** (1957), 152–156. MR 0089127

[81] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. MR 1011252

[82] Emanuel Indrei and Daesung Kim, *Deficit estimates for the logarithmic Sobolev inequality*, arXiv preprint arXiv:1805.06107 (2018).

[83] Emanuel Indrei and Diego Marcon, *A quantitative log-Sobolev inequality for a two parameter family of functions*, Int. Math. Res. Not. IMRN (2014), no. 20, 5563–5580. MR 3271181

[84] Daesung Kim, *Martingale transforms and the Hardy-Littlewood-Sobolev inequality for semigroups*, Potential Anal. **45** (2016), no. 4, 795–807. MR 3558361

[85] ———, *Instability results for the logarithmic Sobolev inequality and its application to the Beckner–Hirschman inequality*, arXiv preprint arXiv:1805.06272 (2018).

[86] ———, *Quantitative inequalities for the expected lifetime of Brownian motion*, arXiv preprint arXiv:1904.09565 (2019).

[87] Victoria Knopova and René L. Schilling, *A note on the existence of transition probability densities of Lévy processes*, Forum Math. **25** (2013), no. 1, 125–149. MR 3010850

[88] M. Ledoux, *On an integral criterion for hypercontractivity of diffusion semigroups and extremal functions*, J. Funct. Anal. **105** (1992), no. 2, 444–465. MR 1160084

[89] Xiang-Dong Li, *Martingale transforms and L^p-norm estimates of Riesz transforms on complete Riemannian manifolds*, Probab. Theory Related Fields **141** (2008), no. 1-2, 247–281. MR 2372971

[90] Elliott H. Lieb, *Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities*, Ann. of Math. (2) **118** (1983), no. 2, 349–374. MR 717827

[91] ———, *Gaussian kernels have only Gaussian maximizers*, Invent. Math. **102** (1990), no. 1, 179–208. MR 1069246

[92] Elliott H. Lieb and Michael Loss, *Analysis*, second ed., Graduate Studies in Mathematics, vol. 14, American Mathematical Society, Providence, RI, 2001. MR 1817225

[93] Vladimir Maz'ya, *Sobolev spaces with applications to elliptic partial differential equations*, augmented ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 342, Springer, Heidelberg, 2011. MR 2777530

[94] Robert J. McCann, *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J. **80** (1995), no. 2, 309–323. MR 1369395

[95] Quoc Anh Ngô and Van Hoang Nguyen, *Sharp reversed Hardy-Littlewood-Sobolev inequality on* $\mathbf{R}^n$, Israel J. Math. **220** (2017), no. 1, 189–223. MR 3666824

[96] ———, *Sharp reversed Hardy-Littlewood-Sobolev inequality on the half space* $\mathbf{R}^n_+$, Int. Math. Res. Not. IMRN (2017), no. 20, 6187–6230. MR 3712196

[97] Adam Osękowski, *Sharp logarithmic inequalities for Riesz transforms*, J. Funct. Anal. **263** (2012), no. 1, 89–108. MR 2920841

[98] ———, *Sharp martingale and semimartingale inequalities*, Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series) [Mathematics Institute of the Polish Academy of Sciences. Mathematical Monographs (New Series)], vol. 72, Birkhäuser/Springer Basel AG, Basel, 2012. MR 2964297

[99] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. MR 1760620

[100] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR 1725357

[101] Gian-Carlo Rota, *An "Alternierende Verfahren" for general positive operators*, Bull. Amer. Math. Soc. **68** (1962), 95–102. MR 0133847

[102] Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 2013, Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. MR 3185174

[103] V. V. Senatov, *Refinement of estimates of stability for a theorem of H. Cramér*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **61** (1976), 125–134, 138, Continuity and stability in problems of probability theory and mathematical statistics. MR 0428383

[104] Ichiro Shigekawa, *Littlewood-Paley inequality for a diffusion satisfying the logarithmic Sobolev inequality and for the Brownian motion on a Riemannian manifold with boundary*, Osaka J. Math. **39** (2002), no. 4, 897–930. MR 1951521

[105] S. L. Sobolev, *On a theorem of functional analysis*, Mat. Sbornik **4** (1938), 471–497.

[106] Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 0290095

[107] ———, *Topics in harmonic analysis related to the Littlewood-Paley theory*, Annals of Mathematics Studies, No. 63, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1970. MR 0252961

[108] M. Talagrand, *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal. **6** (1996), no. 3, 587–600. MR 1392331

[109] Giorgio Talenti, *Best constant in Sobolev inequality*, Ann. Mat. Pura Appl. (4) **110** (1976), 353–372. MR 0463908

[110] ———, *Elliptic equations and rearrangements*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **3** (1976), no. 4, 697–718. MR 0601601

[111] ———, *Aspects of probabilistic Littlewood-Paley theory*, J. Funct. Anal. **38** (1980), no. 1, 25–60. MR 583240

[112] N. Th. Varopoulos, *Hardy-Littlewood theory for semigroups*, J. Funct. Anal. **63** (1985), no. 2, 240–260. MR 803094

[113] Cédric Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR 1964483

[114] Kosaku Yosida, *Functional analysis*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the sixth (1980) edition. MR 1336382