Stein kernels for $q$-moment measures and new bounds for the rate of convergence in the central limit theorem

HUYNH KHANH\textsuperscript{1}  FILIPPO SANTAMBROGIO\textsuperscript{2}  DOAN THAI SON\textsuperscript{3}

Abstract. Given an isotropic probability measure $\mu$ on $\mathbb{R}^d$ with $d\mu(x) = (\varphi(x))^{-\alpha} dx$, where $\alpha > d + 1$ and $\varphi: \mathbb{R}^d \to (0, +\infty)$ is a continuous function and uniformly convex ($\nabla^2 \varphi \geq \varepsilon_0 \mathbf{I}$). By using Stein kernels for $(\alpha - d)$-moment measures, we prove that the rates of convergence in the central limit theorem with sequence of i.i.d. random variables $X_1, X_2, ..., X_n$ of the law $\mu$, to be of form $c \varepsilon_0 \sqrt{d/n}$. The general case (i.e., $\varphi$ is only convex and continuous) remains open.

Contents

1 Introduction ................................................................. 1
2 Stein kernels for $(\alpha - d)$-Moment Measures .................. 4
3 Proof of Proposition 1.1 .................................................. 5
References ................................................................. 14

1 Introduction

Let $\mu$ be an isotropic probability measure on $\mathbb{R}^d$ and $X_1, X_2, ..., X_n$ are i.i.d random variables in $\mathbb{R}^d$ with law $\mu$. Under mild conditions, the central limit theorem (CLT) in probability theory is well-known that, the law $\mu_n$ of $n^{-1/2} \sum_{k=1}^n X_k$, as $n \to \infty$, converges to the standard Gaussian measures $\gamma$ on $\mathbb{R}^d$. However, “how fast does $\mu_n$ converge to $\gamma$?” is still hot problem in many fact cases until now. In recent years, many papers concerned with estimation of the rate of convergence in the central limit theorem in $\mathbb{R}^d$ have appeared, see for examples [8, 7, 10]... They have significantly extended our knowledge in this area.

The fact that one can quantify convergence ($\mu_n \to \gamma$) by Wasserstein distances of order $p \geq 1$ from optimal transport theory, defined as (see also [16, 19])

$$W_p(\mu, \nu) := \left( \min \left\{ \int |x - y|^p d\pi(x, y) : \pi \in \Pi(\mu, \nu) \right\} \right)^{1/p}. \quad (1)$$

\begin{flushright}
\textsuperscript{1}Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi, Vietnam (khanh.edu02@gmail.com /or hhkhanh@math.ac.vn).
\textsuperscript{2}Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex, France, and Institut Universitaire de France (santambrogio@math.univ-lyon1.fr).
\textsuperscript{3}Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi, Vietnam (dtson@math.ac.vn).
\end{flushright}
As a case we are interested in its ideas, M. Fathi [8], applying the Stein’s approximation method [14, 15] for estimating distances between probability measures, and obtained new rates of convergence in the CLT with explicit polynomial dependence on the dimension when considering the log-concave situations, i.e., in the case the measure \( \mu \) is supported on an open, convex set \( K \subset \mathbb{R}^d \) and its density takes the form \( e^{-W(x)} \) where the function \( W : K \to \mathbb{R} \) is convex.

We recall here that a measurable matrix-valued map \( \tau_\mu \) on \( \mathbb{R}^d \) is said to be a Stein kernel for the centered probability \( \mu \) (i.e. has mean zero) if for every smooth test function \( \zeta : \mathbb{R}^d \to \mathbb{R} \),

\[
\int_{\mathbb{R}^d} x.\nabla \zeta \, d\mu = \int_{\mathbb{R}^d} \langle \tau_\mu, \nabla^2 \zeta \rangle_{\text{HS}} \, d\mu
\]

where \( \nabla^2 \zeta \) stands for the Hessian of \( \zeta \), and \( \langle \cdot, \cdot \rangle_{\text{HS}}, \| \cdot \|_{\text{HS}} \) denote the usual Hilbert-Schmidt scalar product and norm, respectively. In applications, it generally suffices to consider the integrals in (2) are well-defined as soon as \( \tau_\mu \in L^2(\mu) \), provided \( \mu \) has finite second moments.

For the standard Gaussian measure \( d\gamma(x) = (2\pi)^{-d/2} e^{-|x|^2/2} dx \) on \( \mathbb{R}^d \), one may directly verify (by using integration by parts formula) that \( \tau_\gamma = \text{Id} \), the identity matrix in \( \mathbb{R}^d \), the proximity of \( \tau_\mu \) with \( \text{Id} \) indicates that \( \mu \) should be close to the Gaussian distribution \( \gamma \). Therefore, whenever such a Stein kernel \( \tau_\mu \) exists, the quantity, called Stein discrepancy of order \( p \geq 2 \) (of \( \mu \) with respect to \( \gamma \))

\[
S_p(\mu | \gamma) := \left( \int_{\mathbb{R}^d} \| \tau_\mu - \text{Id} \|_{\text{HS}}^p \, d\mu \right)^{1/p}
\]

becomes relevant as a measure of the proximity of \( \mu \) and \( \gamma \). As a result, Proposition 3.1 in [8] asserts that: if \( \tau_\mu \) is a Stein kernel for the probability measure \( \mu \) on \( \mathbb{R}^d \) then, for any \( p \geq 2 \)

\[
\mathcal{W}_p(\mu, \gamma) \leq c_p S_p(\mu | \gamma)
\]

(4)

with \( c_p = \sqrt[p]{\int |x|^p \, d\gamma} \). Furthermore, when \( \mu \) is a isotropic probability measure on \( \mathbb{R}^d \), if \( X_1, X_2, \ldots, X_n \) are i.i.d random variables in \( \mathbb{R}^d \) with law \( \mu \) and \( \mu_n \), be law of \( n^{-1/2} \sum_{k=1}^n X_k \), then argument from proof of the Theorem 3.3 in [8] also shows that

\[
S_p(\mu_n | \gamma) \leq 2K_p n^{-1/2} \left( d^{p/2} + \int \left( \sum_{j=1}^d \left| \tau_\mu \right|_j \right)^{p/2} \, d\mu \right)^{1/p}
\]

(5)

where \( K_p = O(p) \) is the best constant in the Rosenthal inequality and \( \{\theta_j\} \) is an orthonormal basis of \( \mathbb{R}^d \).

Observing that both the estimate (4) and the bound in (5) don’t need the log-concavity of the measure \( \mu \). Therefore, this implies that if the Stein kernels exist for an isotropic probability measure \( \mu \) on \( \mathbb{R}^d \), then a potential way for estimating distance between probability measures \( \mu_n \) and \( \gamma \) in the CLT is to find a upper bound for the Stein kernel \( \tau_\mu \).

When \( \mu \) is log-concave and uniformly convex (i.e., \( d\mu = e^{-\varepsilon_0|x|} dx \) and \( \nabla^2 \mu \geq \varepsilon_0 \text{Id} \) for some \( \varepsilon_0 > 0 \)), Klartag [11] gave an upper bound of corresponding \( \tau_\mu \) by \( 1/\varepsilon_0 \). The author in [8]
controlled the rate of convergence in the CLT by using estimates (4), (5) above and Klartag’s result.

Inspired by recent studies \[9, 12\] on the existence of \(q\)-moment maps (i.e., convex functions such that \(\mu\) is their \(q\)-moment measures\(^4\)), the main goal of present paper is to establish new results on rates of convergence in the CLT in the context of power probability distributions, i.e., we shall replace probability distributions of the form \(d\mu(x) = e^{-W(x)}dx\) in \[8\] by power distributions \(d\mu(x) = [g(x)]^{-\mu}dx\) and consider possible results in some particular cases. Namely, our main result is following.

**Theorem 1.1.** Let \(\alpha > d + 1\) and let \(\mu\) be an isotropic probability measure on \(\mathbb{R}^d\) with \(d\mu(x) = (g(x))^{-\alpha}dx\), where \(g : \mathbb{R}^d \to (0, +\infty)\) is continuous function and the function \(x \mapsto g(x) - \varepsilon_0 \frac{|x|^2}{2}\) is convex for some \(\varepsilon_0 > 0\). Suppose that \(\mu_n\) is the law of \(\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k\), where \((X_k)_{k \geq 1}\) is a sequence of i.i.d. random variables with law \(\mu\). Then, we have

\[
W_p(\mu_n, \gamma) \leq C(p, \varepsilon_0) \sqrt{\frac{d}{n}},
\]

for any \(p \geq 2\) and \(C(p, \varepsilon_0)\) is a constant, depending only on \(p, \varepsilon_0\) (which grows like \(p^4\)) and does not depend on \(\mu\).

The typical examples for the probability distributions in Theorem 1.1 are the generalized Cauchy distributions given by (see also \[3\] for more information about this distribution) \(d\mu_\beta(x) := \frac{1}{Z_\beta} (1 + |x|^2)^{-\beta}dx\), for \(\beta > \max \left\{ \frac{d}{2} + 2, d \right\} \)

where \(Z_\beta\) is a normalizing constant \(Z_\beta = \int (1 + |x|^2)^{-\beta}dx = \pi^{d/2} \frac{1}{\Gamma(\beta)} \left( \beta - \frac{d}{2} \right)\). According to Corollary 2.6 in \[4\], Stein kernels exist for the generalized Cauchy distributions on \(\mathbb{R}^d\).

Denote by \(\|\cdot\|_\text{op}\) the operator norm on the \(d \times d\) matrices. The proof of Theorem 1.1 in section 2 is based the estimations from (4) – (5) and the following proposition, as a key challenge in our study.

**Proposition 1.1.** Suppose that hypothesis for the estimate (6) in Theorem 1.1 is satisfied. Then, for each \(x \in \mathbb{R}^d\), the Stein kernel \(\tau_\mu(x)\) for \(\mu\) is positive symmetric matrix, and moreover, there exists a constant \(c(\varepsilon_0)\) depending only on \(\varepsilon_0\) such that \(\|\tau_\mu\|_\text{op} \leq c(\varepsilon_0)\).

The sequel of this paper shall show that the Stein kernels exist for the measure \(\mu\) in Theorem 1.1, and to be of form

\[
\tau_\mu := \frac{1}{\alpha - 1} \psi(\nabla \varphi) \cdot (\nabla^2 \varphi)^{-1}
\]

where convex function \(\psi : \mathbb{R}^d \to (0, +\infty)\) whose \((\alpha - d)\)-moment measure is \(\mu\) and \(\varphi := \psi^*\) is Legendre transform of \(\psi\), defined by \(\varphi(y) = \sup_x (x \cdot y - \psi(x))\).

\(^4\)Given a convex function \(\psi : \mathbb{R}^d \to (0, +\infty)\) such that \(\psi\) goes to \(+\infty\) at infinity and a positive real number \(q > 0\), a Borel probability measure \(\mu\) on \(\mathbb{R}^d\) is said to be the \(q\)-moment measure of \(\psi\) if the differential of \(\psi\) pushes forward the measure \(\psi^{-(d+q)}dx\) towards \(\mu\), i.e., \(\mu = (\nabla \psi)_\# \pi\) with \(d\pi = \psi^{-(d+q)}dx\).
2 Stein kernels for \((\alpha - d)\)-Moment Measures

Let \(q > 0\) be a positive real number, and \(\psi : \mathbb{R}^d \to (0, +\infty)\) is a convex function satisfying \(\lim_{|x| \to +\infty} \psi(x) = +\infty\), we recall that (see also [9]) a probability measures \(\mu\) on \(\mathbb{R}^d\) is said to be the \(q\)-moment measures of \(\psi\) if \(\mu = (\nabla \psi)_x \psi^{-(n+q)}\). The following result provide an existence of the convex function \(\psi\) in case \(q > 1\). In the present context, we consider \(q = \alpha - d\).

**Theorem 2.1.** (Klartag [12] and Santambrogio-Huynh [9]) If the probability measure \(\mu\) is a centered measure, with finite first moment and that is not supported on a hyperplane, then there exists a convex function \(\psi : \mathbb{R}^d \to (0, +\infty)\) whose \(q\)-moment measure is \(\mu\). Moreover, this convex \(\psi\) is uniquely determined up to translation and also, \(\psi\) is essentially continuous, i.e., \(\lim_{x \to x_0} \psi(x) = +\infty\) for \(H^{d-1}\)-a.e. \(x_0 \in \partial \{\psi < +\infty\}\).

The convex function \(\psi\) in Theorem 2.1 is called the \(q\)-moment map of \(\mu\). Caffarelli’s regularity theory for optimal transportation also shows that \(\psi\) is \(C^\infty\)-smooth in \(\mathbb{R}^d\). So, with \(\mu\) being in Theorem 1.1 and \(\psi\) being in Proposition 1.1, the transport equation

\[
\left[ q \left( \nabla \psi (x) \right) \right]^{-\alpha} \det \left( \nabla^2 \psi (x) \right) = \left( \psi (x) \right)^{-\alpha}
\]

holds everywhere in \(\mathbb{R}^d\), where \(\nabla^2 \psi (x)\) is the Hessian matrix of \(\psi\) at \(x\).

Equation (9) is quite similar to the moment measures equation of Berman and Berndtsson [2] in their work on Kähler-Einstein metrics in toric manifolds; Cordero-Erausquin and Klartag extended the study of [2] in [5] presenting a functional version of the classical Minkowski problem or the logarithmic Minkowski problem and providing a variational characterization. More recently, the second author provided in [17] a dual counter-part of the results obtained in [5] with ideas coming from the theory of optimal transport, by considering the minimization of an entropy and a transport cost among probability measures.

The next Theorem give a connection between \((\alpha - d)\)-moment maps and Stein kernels, as well as describes a construction of Stein kernels via \((\alpha - d)\)-moment maps.

**Theorem 2.2.** With the measures \(\mu\) being in Theorem 1.1 and convex function \(\psi\) is solution to the equation (9), then the Stein kernels exist for \(\mu\) and is defined by the formula (8).

The idea for proving the theorem relies on Stein’s equation of the distribution \(\mu\), a few tools from optimal transport and convex analysis. As an example (see [18]), the Student’s \(t\)-distribution \(t_m\) on \(\mathbb{R}\), with \(m = 1, 2, \ldots\) degrees of freedom, has a density function

\[
\rho (x, m) = \frac{\Gamma \left( (m + 1)/2 \right)}{\sqrt{m\pi} \Gamma \left( m/2 \right)} \left( 1 + \frac{x^2}{m} \right)^{-(m+1)/2}, \quad x \in \mathbb{R}
\]

Then, the Stein operator in this case is given by

\[
\mathcal{A} f (x) = \left( 1 + \frac{x^2}{m} \right) f' (x) - \frac{m - 1}{m} x f (x)
\]
and the Stein equation for the $t_m$ distribution is $\mathbb{E}[Af(x)] = 0$, where $\mathbb{E}$ denotes for mathematical expectation. On hypothesis of Theorem 2.2, for any smooth test function $f$ taking values in $\mathbb{R}^d$, we have

$$\int \nabla \psi, f \psi^{-\alpha} \, dx = \int f, \nabla \left( - \frac{1}{\alpha - 1} \psi^{-(\alpha - 1)} \right) \, dx = \frac{1}{\alpha - 1} \int \text{div}(f) \cdot \psi^{-(\alpha - 1)} \, dx$$  \hspace{1cm} (12)

is Stein equation for the distribution $\mu$.

**Proof of Theorem 2.2.** Taking $f(x) = g(\nabla \psi(x))$ in the equation (12), we obtain

$$\int \nabla \psi, g(\nabla \psi) \psi^{-\alpha} \, dx = \frac{1}{\alpha - 1} \int \langle \nabla^2 \psi, \nabla g(\nabla \psi) \rangle_{\text{HS}} \psi^{-(\alpha - 1)} \, dx.$$  \hspace{1cm} (13)

We know that the map $x \mapsto \nabla \psi(x)$ sends $\psi^{-\alpha}$ into $\mu$ and the map $x \mapsto \nabla \varphi(x)$ sends $\mu$ into $\psi^{-\alpha}$, $\nabla \psi(\nabla \varphi(x)) = x$, using the change of variable $y = \nabla \varphi(x)$, we get

$$\int x.g(x) \, d\mu(x) = \frac{1}{\alpha - 1} \int \langle \psi(\nabla \varphi), (\nabla^2 \varphi)^{-1}, \nabla g \rangle_{\text{HS}} \, d\mu.$$  \hspace{1cm} (14)

This implies that the Stein kernel exists for $\mu$ and the formula (8) holds true.

**Proof of Theorem 1.1.** Without loss of generality, one may assume that the law $\mu$ has a compact support and a density bounded away from zero. This fact is deduced from, if the support of $\mu$ is not compact, then one can take a sequence of compact sets $F_k$ that converge to the support of $\mu$ and apply results to the restriction of $\mu$ to $F_k$. The estimate on the Wasserstein distance does not depend on $F_k$, so that one can let $k$ go to infinity and obtain results in general cases.

According to Theorem 2.2, the Stein kernel $\tau_\mu = \left( (\tau_\mu)_{ij} \right)_{i,j=1,\ldots,d}$ for $\mu$ is given by the formula (8), and as is standard, $\tau_n(x) := \mathbb{E} \left[ \frac{1}{n} \sum_{j=1}^n \tau_\mu(X_j) \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j = x \right]$ is a Stein kernel for $\mu_n$. Using the estimates (4) and (5), we obtain

$$\mathcal{W}_p(\mu_n, \gamma) \leq 2K_p n^{-1/2} c_p \left( d^{p/2} + \int \left( \sum_{j=1}^d \left| (\tau_\mu)_{ij} \right|^2 \right)^{p/2} \, d\mu \right)^{1/p}.$$  \hspace{1cm} (15)

So, the estimate (6) is deduced directly from (15) and Proposition 1.1. This finishes proof of Theorem 1.1.

**3 Proof of Proposition 1.1**

The goal of this section is to provide a detailed proof of Proposition 1.1. Recall that $\psi$ is solution to the transport equation $\left( \varrho(\nabla \psi(x)) \right)^{-\alpha} \text{det}(\nabla^2 \psi(x)) = (\psi(x))^{-\alpha}$, $\varphi := \psi^*$ and
Stein kernels for $(\alpha-d)$-Moment Measures and Applications

\[ \tau_\mu(x) = \frac{1}{\alpha-1} \psi(h^2 \nabla^2 \psi(x)) \nabla^2 \psi(x) \nabla^2 \psi(x) \]. Therefore, what we need to prove is

\[
\sup_{x \in \nabla \phi(Supp(\mu))} \left\{ \psi(x) \cdot \left\| \nabla^2 \psi(x) \right\| \right\} < +\infty. \quad (16)
\]

For $\varepsilon > 0$, $h \in \mathbb{R}^d$ and a function $f : \mathbb{R}^d \to \mathbb{R}$ denote

\[
\partial_{hh}^\varepsilon f(x) = f(x + \varepsilon h) + f(x - \varepsilon h) - 2f(x), \quad x \in \mathbb{R}^d. \quad (17)
\]

If $f$ is smooth and $\varepsilon$ is enough small, then the quantity $\partial_{hh}^\varepsilon f(x)$ approximates the pure second derivative $f_{hh}(x)$. The proof of Proposition 1.1 is similar to Proposition 3.1 in [11] and the Caffarelli’s contraction theorem in [6], up to some modifications and other techniques. The main difficult in new setting is, we shall work on the following equation which is obtained by taking logarithm of the equation (9)

\[
\log \left[ \det \left( \nabla^2 \psi(x) \right) \right] = \alpha \left[ -\log \left( \psi(x) \right) + \log \left( \theta \left( \nabla \psi(x) \right) \right) \right], \quad x \in \mathbb{R}^d. \quad (18)
\]

According to properties of the $q$-moment maps (see [12], Proposition 4.4), the smooth map $\nabla \psi : \nabla \varphi(Supp(\mu)) \to Supp(\mu)$ is one-to-one. On the other hand, we have $0 < m_1 \leq \frac{1}{\psi(x)} \leq m_2$ for any $x \in \nabla \varphi(Supp(\mu))$ and some $m_1, m_2 \in \mathbb{R}$. So, to prove (16), it is sufficient to show that

\[
(log \psi)_{hh}(x) = \lim_{\varepsilon \to 0^+} \frac{\partial_{hh}^\varepsilon \log \psi(x)}{\varepsilon^2} \leq c(\varepsilon_0). \quad (19)
\]

Denote by $(\psi^{ij}(x))_{i,j=1,...,d} = (\nabla^2 \psi(x))^{-1}$ the inverse matrix of $\nabla^2 \psi$ at $x$. For a smooth function $u : \mathbb{R}^d \to \mathbb{R}$,

\[
Au(x) := \text{Tr} \left[ (\nabla^2 \psi(x))^{-1} \nabla^2 u(x) \right] = \psi^{ij}(x) u_{ij}(x). \quad (20)
\]

Arguments coming from matrix inequalities\footnote{we are applying, that is:}

\[
\log \det \nabla^2 \psi(x + h) \leq \log \det \nabla^2 \psi(x) + \psi^{ij}(x) \psi_{ij}(x + h) - d \quad (21)
\]

with an equality for $h = 0$. In general, when $A$ and $B$ are symmetric, positive-definite $d \times d$ matrices, then

\[
\log (\det B) \leq \log (\det A) + \text{Tr} \left( A^{-1} B \right) - n. \quad (22)
\]

\[\text{Proof of Proposition 1.1.} \] The proof will be conducted in 5 steps.
Step 1. Existence of maximum point of function \((x, h) \mapsto \partial_{hh}^x \log \psi\) over \(\mathbb{R}^d \times S^{d-1}\).

We first have that the function the function \((x, h) \mapsto \partial_{hh}^x \log \psi\) is continuous on \(\mathbb{R}^d \times S^{d-1}\) and

\[
\partial_{hh}^x \log \psi (x) = \log \psi (x + \varepsilon h) + \log \psi (x - \varepsilon h) - 2 \log \psi (x)
\]

\[
= \log \frac{\psi (x + \varepsilon h) \cdot \psi (x - \varepsilon h)}{(\psi (x))^2}.
\]

On the other hand, \(\psi\) is convex function satisfying \(\lim_{|x| \to +\infty} \psi (x) = +\infty\) and \(x\) is middle point of \(x + \varepsilon h\) and \(x - \varepsilon h\) for any \((x, h) \in \mathbb{R}^d \times S^{d-1}\), so that there exists at least a pair \((\overline{x}, \overline{h}) \in \mathbb{R}^d \times S^{d-1}\) such that \(\psi (\overline{x} + \varepsilon \overline{h}) \geq \psi (\overline{x})\) and \(\psi (\overline{x} - \varepsilon \overline{h}) > \psi (\overline{x})\), it follows that

\[
\partial_{hh}^x \log \psi (\overline{x}) > 0. \quad (24)
\]

Moreover,

\[
2(\psi (x))^2 \cdot \partial_{hh}^x \log \psi (x) = 2 \psi (x + \varepsilon h) \cdot \psi (x - \varepsilon h)
\]

\[
= [\psi (x + \varepsilon h) + \psi (x - \varepsilon h)]^2 - [\psi^2 (x + \varepsilon h) + \psi^2 (x - \varepsilon h)]
\]

\[
= [\partial_{hh}^x \psi (x) + 2 \psi (x)]^2 - [\psi^2 (x + \varepsilon h) + \psi^2 (x - \varepsilon h)]
\]

and hence

\[
e^{\partial_{hh}^x \log \psi (x)} = \left( \frac{\partial_{hh}^x \psi (x)}{2 \psi^2 (x)} \right)^2 + \frac{2 \partial_{hh}^x \psi (x)}{\psi (x)} + 2 - \frac{\psi^2 (x + \varepsilon h) + \psi^2 (x - \varepsilon h)}{2 \psi^2 (x)}. \quad (25)
\]

Setting \(m_1 (\varepsilon) = \frac{2}{2 \psi^2 (x)} [\psi^2 (x + \varepsilon h) + \psi^2 (x - \varepsilon h)]\), then \(m_1 (\varepsilon) > 0\) for \(\varepsilon\) small enough and \(\lim_{|x| \to +\infty} m_1 (\varepsilon) = 1\). The equality \((25)\) becomes

\[
e^{\partial_{hh}^x \log \psi (x)} = m_1 (\varepsilon) + \frac{(\partial_{hh}^x \psi (x))^2}{2 \psi^2 (x)} + \frac{2 \partial_{hh}^x \psi (x)}{\psi (x)}. \quad (26)
\]

Dividing \((26)\) by \(m_1 (\varepsilon)\), note that \(\frac{1}{m_1 (\varepsilon)} e^{\partial_{hh}^x \log \psi (x)} = e^{\log \frac{1}{m_1 (\varepsilon)} + \partial_{hh}^x \log \psi (x)}\) and \(\frac{1}{m_1 (\varepsilon)} \leq \frac{4}{(m_1 (\varepsilon))^4} =: (m_2 (\varepsilon))^2\) when \(m_1 (\varepsilon) \sim 1\), we have

\[
e^{\log \frac{1}{m_1 (\varepsilon)} + \partial_{hh}^x \log \psi (x)} = 1 + m_2 (\varepsilon) \frac{\partial_{hh}^x \psi (x)}{\psi (x)} + \frac{1}{2} \frac{1}{m_1 (\varepsilon)} \left( \frac{\partial_{hh}^x \psi (x)}{\psi (x)} \right)^2
\]

\[
\leq 1 + m_2 (\varepsilon) \frac{\partial_{hh}^x \psi (x)}{\psi (x)} + \frac{1}{2} \left( m_2 (\varepsilon) \frac{\partial_{hh}^x \psi (x)}{\psi (x)} \right)^2
\]

\[
\leq 1 + m_2 (\varepsilon) \frac{\partial_{hh}^x \psi (x)}{\psi (x)} + \frac{1}{2} \left( m_2 (\varepsilon) \frac{\partial_{hh}^x \psi (x)}{\psi (x)} \right)^2 + \sum_{k=3}^{+\infty} \frac{1}{k!} \left( m_2 (\varepsilon) \frac{\partial_{hh}^x \psi (x)}{\psi (x)} \right)^k. \quad (27)
\]

Using Taylor expansion of function \(y \mapsto e^y\), the \((27)\) yields

\[
\exp \left( \log \frac{1}{m_1 (\varepsilon)} + \partial_{hh}^x \log \psi (x) \right) \leq \exp \left( m_2 (\varepsilon) \frac{\partial_{hh}^x \psi (x)}{\psi (x)} \right) \quad (28)
\]
The monotonicity of function $x \mapsto e^x$ on $\mathbb{R}$ give us
\[
\log \frac{1}{m_1(\varepsilon)} + \partial_{hh} \log \psi(x) \leq \frac{m_2(\varepsilon) \partial^2_{hh} \psi(x)}{\psi(x)} \leq m_3(\varepsilon) \partial^2_{hh} \psi(x) \tag{29}
\]
for $|x|$ big enough and $m_3(\varepsilon) > 0$. By using (29) and proof of Corollary 3.4 in [11], we obtain
\[
\lim_{R \to \infty} \left\{ \sup_{|x| \geq R, h \in \mathbb{S}^{d-1}} \partial_{hh} \log \psi(x) \right\} \leq \lim_{R \to \infty} \left\{ \sup_{|x| \geq R, h \in \mathbb{S}^{d-1}} \left( m_3(\varepsilon) \partial^2_{hh} \psi(x) + \log m_1(\varepsilon) \right) \right\} = 0. \tag{30}
\]
Here we applied similar argument as in the proof of Corollary 3.4 in [11] in order to conclude
\[
\lim_{R \to +\infty} \left( \sup_{|x| \geq R, h \in \mathbb{S}^{d-1}} \partial_{hh} \psi(x) \right) = 0, \quad \text{and the rest } \lim_{|x| \to +\infty} (\log m_1(\varepsilon)) = 0. \quad \text{Combining with (24), the (30) implies that the maximum of the function (x, h) \mapsto \partial_{hh} (\log \psi) over } \mathbb{R}^d \times \mathbb{S}^{d-1} \text{ is indeed attained at some } (x_0, e) \in \mathbb{R}^d \times \mathbb{S}^{d-1}.
\]

Step 2. Deducing that $A(\partial_{ee} \psi)(x_0) = o(\varepsilon^2)$ with $\lim_{\varepsilon \to 0} \frac{o(\varepsilon^2)}{\varepsilon^2} < +\infty$.\footnote{In this case, we can prove that $\lim_{\varepsilon \to 0} \frac{o(\varepsilon^2)}{\varepsilon^2} = 0.$}

According to the definition of the operator $A$, we have
\[
A(\partial_{ee} \psi)(x_0) = \text{Tr} \left[ (\nabla^2 \psi(x_0))^{-1} \nabla^2 U(x_0) \right] \tag{31}
\]
where
\[
\nabla^2 U(x_0) = \nabla^2 (\partial_{ee} \psi)(x_0) = \nabla^2 \psi(x_0 + \varepsilon e) + \nabla^2 \psi(x_0 - \varepsilon e) - 2 \nabla^2 \psi(x_0) = [\nabla^2 \psi(x_0 + \varepsilon e) - \nabla^2 \psi(x_0)] + [\nabla^2 \psi(x_0 - \varepsilon e) - \nabla^2 \psi(x_0)]
\]
\[
= [\psi_{ij}(x_0 + \varepsilon e) - \psi_{ij}(x_0)]_{ij} + [\psi_{ij}(x_0 - \varepsilon e) - \psi_{ij}(x_0)]_{ij}
\]
\[
= \left[ \nabla \psi_{ij} \left( x_0 + \lambda_{ij}^{(1)} \varepsilon e \right) \right]_{ij} + \left[ \nabla \psi_{ij} \left( x_0 - \lambda_{ij}^{(2)} \varepsilon e \right) \right]_{ij}
\]
\[
= \varepsilon. \left\{ \left[ \nabla \psi_{ij} \left( x_0 + \lambda_{ij}^{(1)} \varepsilon e \right) \right]_{ij} - \left[ \nabla \psi_{ij} \left( x_0 - \lambda_{ij}^{(2)} \varepsilon e \right) \right]_{ij} \right\}
\]
\[
= \varepsilon. \left[ \left( \nabla \psi_{ij} \left( x_0 + \lambda_{ij}^{(1)} \varepsilon e \right) \right)_{ij} - \left( \nabla \psi_{ij} \left( x_0 - \lambda_{ij}^{(2)} \varepsilon e \right) \right)_{ij} \right] \tag{32}
\]
with $\lambda_{ij}^{(1)}, \lambda_{ij}^{(2)} \in (0, 1)$ and their existences are from the mean value theorem. And additionally, for any $i, j \in \{1, \ldots, d\}$,
\[
\nabla \psi_{ij} \left( x_0 + \lambda_{ij}^{(1)} \varepsilon e \right) - \nabla \psi_{ij} \left( x_0 - \lambda_{ij}^{(2)} \varepsilon e \right) = \begin{pmatrix}
\psi_{ij1}(x_0 + \lambda_{ij}^{(1)} \varepsilon e) - \psi_{ij1}(x_0 - \lambda_{ij}^{(2)} \varepsilon e) \\
\psi_{ij2}(x_0 + \lambda_{ij}^{(1)} \varepsilon e) - \psi_{ij2}(x_0 - \lambda_{ij}^{(2)} \varepsilon e) \\
\psi_{ijd}(x_0 + \lambda_{ij}^{(1)} \varepsilon e) - \psi_{ijd}(x_0 - \lambda_{ij}^{(2)} \varepsilon e)
\end{pmatrix} \tag{33}
\]
Once again using the mean value theorem, there exist \( \lambda_{ij} \in (0, 1) \) such that
\[
\psi_{ijk} \left( x_0 + \lambda_{ij}^{(1)} \varepsilon e \right) - \psi_{ijk} \left( x_0 - \lambda_{ij}^{(2)} \varepsilon e \right) = \varepsilon \left( \lambda_{ij}^{(1)} + \lambda_{ij}^{(2)} \right) \left( e \cdot \nabla \psi_{ijk} \left( x_0 + \lambda_{ij}^{(1)} \varepsilon e - (1 - \lambda_{ij}) \cdot \lambda_{ij}^{(2)} \varepsilon e \right) \right)
\]
for each \( k = 1, 2, \ldots, d \). From (32), (33) and (34), the matrix \( \nabla^2 U (x_0) \) could be rewrite as following
\[
\nabla^2 U (x_0) = o \left( \varepsilon^2 \right) \cdot \nabla^2 V (x_0)
\]
in this case, it’s easy to see that the matrix \( \nabla^2 V (x_0) \) tends to zero matrix as \( \varepsilon \to 0^+ \). Combining (31) and (35), we get
\[
A \left( \partial_{ee}^\varphi \psi \right) (x_0) = o \left( \varepsilon^2 \right) \cdot \text{Tr} \left[ \left( \nabla^2 \psi (x_0) \right)^{-1} \nabla^2 V (x_0) \right] = o \left( \varepsilon^2 \right).
\]
As a corollary (from the equation (18), (23) and (36)),
\[
\partial_{ee}^\varphi \left( \log \varphi \left( \nabla \psi \right) \right) (x_0) = \partial_{ee}^\varphi \left( \log \psi \right) (x_0) + \frac{1}{\alpha} \partial_{ee}^\varphi \left( \log \det \nabla^2 \psi \right) (x_0)
\]
\[
\leq \partial_{ee}^\varphi \left( \log \psi \right) (x_0) + \frac{1}{\alpha} A \left( \partial_{ee}^\varphi \psi \right) (x_0)
\]
\[
= \partial_{ee}^\varphi \left( \log \psi \right) (x_0) + o \left( \varepsilon^2 \right).
\]

**Step 3.** Finding a lower bound of \( e \partial_{ee}^\varphi \left( \log \varphi \left( \nabla \psi \right) \right) (x_0) \) in terms of \(|u|\) and \( \varepsilon \).

The function \( (x, h) \mapsto \partial_{hh}^\varphi \left( \log \psi \right) \) is attained maximum at \((x_0, e)\), the first optimality condition gives us
\[
0 = \nabla \left( \partial_{ee}^\varphi \log \psi \right) (x_0)
\]
\[
= \frac{1}{\psi (x_0 + \varepsilon e)} \nabla \psi (x_0 + \varepsilon e) + \frac{1}{\psi (x_0 - \varepsilon e)} \nabla \psi (x_0 - \varepsilon e) - 2 \frac{1}{\psi (x_0)} \nabla \psi (x_0)
\]
deduces the existence of a vector \( u \in \mathbb{R}^d \) such that
\[
\begin{cases}
\frac{1}{\psi (x_0 + \varepsilon e)} \nabla \psi (x_0 + \varepsilon e) = \frac{1}{\psi (x_0)} \left( \nabla \psi (x_0) + u \right) \\
\frac{1}{\psi (x_0 - \varepsilon e)} \nabla \psi (x_0 - \varepsilon e) = \frac{1}{\psi (x_0)} \left( \nabla \psi (x_0) - u \right)
\end{cases}
\]
Setting \( v = \nabla \psi (x_0) \) and
\[
\begin{aligned}
y_1 &= \frac{\psi (x_0 - \varepsilon e)}{\psi (x_0)} (v - u) \\
y_2 &= \frac{\psi (x_0 + \varepsilon e)}{\psi (x_0)} (v + u) \\
y_3 &= \frac{1}{2} (y_1 + y_2)
\end{aligned}
\]
Moreover, the convexity of \( \psi \) and \( 2 \) yields
\[
g(y_1) + g(y_2) - 2g\left(\frac{1}{2}(y_1 + y_2)\right) \geq 0. \tag{41}
\]

It follows that
\[
\begin{align*}
\partial_{ee}^\psi(\nabla \psi)(x_0) &= \psi(\nabla \psi(x_0 + \varepsilon e)) + \psi(\nabla \psi(x_0 - \varepsilon e)) - 2\psi(\nabla \psi(x_0)) \\
&= \psi(y_2) + \psi(y_1) - 2\psi(v) \\
&\geq \varepsilon_0 \left(y_1^2 + y_2^2 - 2y_3^2\right) + 2\left(\psi(y_3) - \psi(v)\right) \\
&= \varepsilon_0 \left(y_1^2 + y_2^2 - \frac{1}{2}(y_1 + y_2)^2\right) + 2\left(\psi(y_3) - \psi(v)\right) \\
&= \frac{\varepsilon_0}{4}(y_1 - y_2)^2 + 2\left(\psi(y_3) - \psi(v)\right). \tag{42}
\end{align*}
\]
Moreover, the convexity of \( \psi \) implies that
\[
\psi(x_0 + \varepsilon e) + \psi(x_0 - \varepsilon e) \geq 2\psi\left(\frac{1}{2}(x_0 + \varepsilon e) + \frac{1}{2}(x_0 - \varepsilon e)\right) = 2\psi(x_0)
\]
and
\[
\psi(x_0 + \varepsilon e) - \psi(x_0 - \varepsilon e) \geq 2\epsilon \left(\nabla \psi(x_0 - \varepsilon e) \cdot (e)\right).
\]

Therefore, we have
\[
\frac{\varepsilon_0}{4}(y_1 - y_2)^2 = \frac{\varepsilon_0}{4\psi^2(x_0)} \left\{ \left(\psi(x_0 + \varepsilon e) + \psi(x_0 - \varepsilon e)\right)u + \left(\psi(x_0 + \varepsilon e) - \psi(x_0 - \varepsilon e)\right)v \right\}^2 \\
\geq \varepsilon_0 |u|^2 + \varepsilon a_0(\varepsilon) |u| + o(\varepsilon^2). \tag{43}
\]

On the other hand, by the mean value theorem, there exists a vector \( \xi \in \mathbb{R}^d \), \( \lambda_1 \in (-1, 1) \) and \( \lambda_2, \lambda_3 \in (0, 1) \) such that
\[
\begin{align*}
\psi(y_3) - \psi(v) &= \nabla g(\xi)(y_3 - v) \tag{44} \\
\psi(x_0 + \varepsilon e) - \psi(x_0 - \varepsilon e) &= \nabla \psi(x_0 + \lambda_1 \varepsilon e) \cdot (2\varepsilon e) \tag{45} \\
\psi(x_0 + \varepsilon e) - \psi(x_0) &= \nabla \psi(x_0 + \lambda_2 \varepsilon e) \cdot (\varepsilon e) \tag{46} \\
\psi(x_0 - \varepsilon e) - \psi(x_0) &= \nabla \psi(x_0 - \lambda_3 \varepsilon e) \cdot (-\varepsilon e) \tag{47}
\end{align*}
\]
and this yields
\[
\begin{align*}
2\left(\psi(y_3) - \psi(v)\right) &= 2\nabla g(\xi)(y_3 - v) = \nabla g(\xi)(y_1 + y_2 - 2v) \\
&= \frac{1}{\psi(x_0)} \nabla g(\xi) \left\{ \left(\psi(x_0 + \varepsilon e) - \psi(x_0 - \varepsilon e)\right)u + \left(\psi(x_0 + \varepsilon e) - \psi(x_0)\right)v + \left(\psi(x_0 - \varepsilon e) - \psi(x_0)\right)v \right\} \\
&= \frac{e}{\psi(x_0)} \nabla g(\xi) \left\{ \left(\nabla \psi(x_0 + \lambda_1 \varepsilon e) \cdot (2\varepsilon e)\right)u + \left(\nabla \psi(x_0 + \lambda_2 \varepsilon e) \cdot (\varepsilon e)\right)v + \left(\nabla \psi(x_0 - \lambda_3 \varepsilon e) \cdot (-\varepsilon e)\right)v \right\} \\
&\geq \varepsilon a(\varepsilon) |u| + \frac{e}{\psi(x_0)} \nabla g(\xi) \left\{ \left(\nabla \psi(x_0 + \lambda_2 \varepsilon e) - \nabla \psi(x_0 - \lambda_3 \varepsilon e)\right) \cdot (e)\right\}v. \tag{48}
\end{align*}
\]
Once again, using the mean value theorem, there exists \( \lambda_j, j = 1, ..., d \)

\[
\nabla \psi (x_0 + \lambda_2 \varepsilon e) - \nabla \psi (x_0 - \lambda_3 \varepsilon e) = \left( \begin{array}{l}
\psi_1 (x_0 + \lambda_2 \varepsilon e) - \psi_1 (x_0 - \lambda_3 \varepsilon e) \\
\vdots \\
\psi_k (x_0 + \lambda_2 \varepsilon e) - \psi_k (x_0 - \lambda_3 \varepsilon e) \\
\psi_d (x_0 + \lambda_2 \varepsilon e) - \psi_d (x_0 - \lambda_3 \varepsilon e)
\end{array} \right)
\]

\[
= \varepsilon (\lambda_2 + \lambda_3) \left( \begin{array}{l}
\nabla \psi_1 (x_0 + \lambda_1 \varepsilon e) \cdot (e) \\
\vdots \\
\nabla \psi_k (x_0 + \lambda_k \varepsilon e) \cdot (e) \\
\nabla \psi_d (x_0 + \lambda_d \varepsilon e) \cdot (e)
\end{array} \right)
\]

From (48) and (49),

\[
2 \left( \varrho (y_3) - \varrho (v) \right) \geq \varepsilon a (\varepsilon) |u| + o(\varepsilon^2).
\]

Combining (42), (43) and (50), we obtain

\[
\partial_{ee} \left( \varrho (\nabla \psi) \right) (x_0) \geq |u|^2 - \varepsilon \overline{\omega} \varepsilon u (\varepsilon) |u| + o(\varepsilon^2).
\]

Now we evaluate

\[
\partial_{ee} \left( \log \varrho (\nabla \psi) \right) (x_0) = \frac{\partial_{ee} \left( \varrho (\nabla \psi) \right) (x_0) + 2 \varrho (v) + \varrho^2 (\nabla \psi (x_0 + \varepsilon) + \varrho^2 (\nabla \psi (x_0 - \varepsilon))}{2 \varrho^2 (v)}
\]

\[
= \frac{\partial_{ee} \left( \varrho (\nabla \psi) \right) (x_0)}{2 \varrho^2 (v)} + \frac{2 \varrho (v) \partial_{ee} \left( \varrho (\nabla \psi) \right) (x_0)}{2 \varrho^2 (v)}
\]

\[
+ 1 - \frac{\left( \varrho^2 (y_2) - \varrho^2 (v) \right) + \left( \varrho^2 (y_1) - \varrho^2 (v) \right)}{2 \varrho^2 (v)}
\]

\[
\geq 1 + \frac{2 \varrho (v) \partial_{ee} \left( \varrho (\nabla \psi) \right) (x_0) - \left( \varrho^2 (y_2) - \varrho^2 (v) \right) + \left( \varrho^2 (y_1) - \varrho^2 (v) \right)}{2 \varrho^2 (v)}.
\]

Note that \( \varrho (y_1) \rightarrow \varrho (v), \varrho (y_2) \rightarrow \varrho (v) \) as \( \varepsilon \rightarrow 0^+ \), and the function \( \varrho \) is bounded on \( \text{Supp} \mu \), hence there exists a real number \( k \sim 1 \) (i.e., \( |k - 1| \ll 0 \) when \( \varepsilon \) is enough small) such that

\[
\frac{\left( \varrho^2 (y_2) - \varrho^2 (v) \right) + \left( \varrho^2 (y_1) - \varrho^2 (v) \right)}{2 \varrho^2 (v)} = \frac{\left( \varrho (y_2) + \varrho (v) \right) \left( \varrho (y_2) - \varrho (v) \right) + \varrho (y_1) + \varrho (v) \left( \varrho (y_1) - \varrho (v) \right)}{2 \varrho^2 (v)}
\]

\[
\leq \frac{k}{\varrho (v)} [(\varrho (y_2) - \varrho (v)) + (\varrho (y_1) - \varrho (v))]
\]

\[
= \frac{k}{\varrho (v)} \partial_{ee} \left( \varrho (\nabla \psi) \right) (x_0).
\]
Setting \( m = \frac{2-k}{\varepsilon u} \sim \frac{1}{\varepsilon u} \), from (51), (52) and (53),

\[
\left( \log \varepsilon (\nabla \psi) \right)(x_0) \geq 1 + m.\partial_{ee}(\log \varepsilon(\nabla \psi))(x_0) \geq 1 + m\varepsilon_0 |u|^2 - \varepsilon a_1(\varepsilon) |u| + o(\varepsilon^2). 
\]

(54)

**Step 4. Finding an upper bound in terms of \(|u|\) and \(\varepsilon\) for \(\partial_{ee}(\log \psi)(x_0)\).**

According the definition, we have

\[
\partial_{ee}(\log \psi)(x_0) = \left( \log \psi(x_0 + \varepsilon \epsilon) - \log \psi(x_0) \right) + \left( \log \psi(x_0 - \varepsilon \epsilon) - \log \psi(x_0) \right).
\]

By the theorem of mean values, there exist \(\beta_1, \beta_2 \in (0, 1)\),

\[
\partial_{ee}(\log \psi)(x_0) = \frac{1}{\psi(x_0 + \beta_1 \varepsilon \epsilon)} \nabla \psi(x_0 + \beta_1 \varepsilon \epsilon). (\varepsilon \epsilon) + \frac{1}{\psi(x_0 - \beta_2 \varepsilon \epsilon)} \nabla \psi(x_0 - \beta_2 \varepsilon \epsilon). (-\varepsilon \epsilon)
\]

\[
\epsilon \left\{ \frac{1}{\psi(x_0 + \beta_1 \varepsilon \epsilon)} \nabla \psi(x_0 + \beta_1 \varepsilon \epsilon). e - \frac{1}{\psi(x_0 - \beta_2 \varepsilon \epsilon)} \nabla \psi(x_0 - \beta_2 \varepsilon \epsilon). e \right\}
\]

Since \(\nabla \psi(\mathbb{R}^d) \subseteq \text{Supp}(\mu)\), which is a convex, compact set, this deduces that there exists a real number \(K\) such that

\[
\partial_{ee}(\log \psi)(x_0) \leq \epsilon K \left\{ \frac{1}{\psi(x_0 + \varepsilon \epsilon)} \nabla \psi(x_0 + \varepsilon \epsilon). e - \frac{1}{\psi(x_0 - \varepsilon \epsilon)} \nabla \psi(x_0 - \varepsilon \epsilon). e \right\}
\]

\[
= \epsilon K \left\{ \frac{\psi(x_0 + \epsilon) + \psi(x_0 - \epsilon)}{\psi(x_0)} (u.e) + \frac{\psi(x_0 + \epsilon) - \psi(x_0 - \epsilon)}{\psi(x_0)} (v.e) \right\}
\]

\[
\leq \epsilon K \left\{ \frac{\psi(x_0 + \epsilon) + \psi(x_0 - \epsilon)}{\psi(x_0)} (u.e) + \frac{\nabla \psi(x_0 + \epsilon) (2\epsilon \epsilon)}{\psi(x_0)} (v.e) \right\}
\]

It follows that

\[
\partial_{ee}(\log \psi)(x_0) \leq \epsilon a_2(\varepsilon) |u| + o(\varepsilon^2)
\]

(55)

where \(a_2(\varepsilon) > 0\) and \(\lim_{\varepsilon \to 0^+} a_2(\varepsilon) < +\infty\).

**Step 5. Conclusion**

From (37), (54), (55) and Taylor expansion of function \(y \mapsto e^y\), we get

\[
1 + m\varepsilon_0 |u|^2 - \varepsilon a_1(\varepsilon) |u| + o(\varepsilon^2) \leq e^{\partial_{ee}(\log \varepsilon(\nabla \psi))(x_0)}
\]

\[
\leq e^{\partial_{ee}(\log \psi)(x_0) + o(\varepsilon^2)}
\]

\[
= e^{\epsilon a_2(\varepsilon) |u| + o(\varepsilon^2)}
\]

\[
= \sum_{k=0}^{\infty} \left( \frac{\epsilon a_2(\varepsilon) |u| + o(\varepsilon^2)}{k!} \right)^k
\]

\[
= 1 + \varepsilon a_2(\varepsilon) |u| + o(\varepsilon^2)
\]

(56)
once again, we note that $o(\varepsilon^2)$ means \(\lim_{\varepsilon\to 0^+} \frac{a(\varepsilon^2)}{\varepsilon^2} < +\infty\). Therefore, the (56) yields
\[
m \varepsilon_0 |u|^2 - \varepsilon a_3(\varepsilon) |u| + o(\varepsilon^2) \leq 0.
\]
and it follows that
\[
0 \leq |u| \leq \varepsilon a_4(\varepsilon)
\]
where $a_3(\varepsilon)$ and $a_4(\varepsilon)$ satisfying $\lim_{\varepsilon\to 0^+} a_3(\varepsilon) < +\infty$ and $\lim_{\varepsilon\to 0^+} a_4(\varepsilon) < +\infty$. Combining (55) and (58), we get
\[
\partial_{ee}^x \left( \log \psi \right)(x_0) \leq \varepsilon^2 a_5(\varepsilon) + o(\varepsilon^2).
\]
Finally, we obtain
\[
\left( \log \psi \right)_{hh}(x) = \lim_{\varepsilon \to 0^+} \varepsilon^2 \partial_{hh}^x \left( \log \psi \right)(x)
\leq \lim_{\varepsilon \to 0^+} \varepsilon^2 \partial_{ee}^x \left( \log \psi \right)(x_0)
= \lim_{\varepsilon \to 0^+} \frac{\varepsilon^2 a_5(\varepsilon) + o(\varepsilon^2)}{\varepsilon^2} = c(\varepsilon_0)
\]
for any $x \in \mathbb{R}^d$ and $h \in S^{d-1}$. This completes the proof.

\[\square\]

References

[1] D. Bakry, I. Gentil, M. Ledoux, Analysis and geometry of Markov diffusion operators. Grundlehren der Mathematischen Wissenschaften 348 (2014). Springer.

[2] Robert J. Berman, B. Berndtsson, Real Monge-Ampère equations and Kähler-Ricci solitons on toric log Fano varieties. Ann. Fac. Sci. Toulouse Math. (6), Vol. 22, No. 4, (2013), 649–711.

[3] S. Bobkov and M. Ledoux, Weighted Poincaré-type inequalities for Cauchy and other convex measures. Ann. Probab. 37 (2009) 403–427.

[4] T. A. Courtadea, M. Fathi and A. Pananjady, Existence of Stein kernels under a spectral gap, and discrepancy bounds. Annales de l’Institut Henri Poincaré - Probabilités et Statistiques, 2019, Vol. 55, No. 2, 777–790

[5] D. Cordero-Erausquin, B. Klartag, Moment measures, J. Functional Analysis, Vol. 268, No. 12, (2015), 3834–3866.

[6] L. A. Caffarelli, Monotonicity properties of optimal transportation and the FKG and related inequalities. Comm. Math. Phys., Vol. 214, No. 3, (2000), 547–563.

[7] R. Eldan, D. Mikulincer, A. Zhai, The CLT in high dimensions: quantitative bounds via martingale embedding. Arxiv preprint, 2020.
Stein kernels for \((\alpha - d)\)-Moment Measures and Applications

[8] M. Fathi, Stein kernels and moment maps. *Annals of Probability*, 47(4):2172–2185, 2019.

[9] H. Khanh, F. Santambrogio, \(q\)-Moment Measures and Applications: a new Approach via Optimal Transport. *Journal of Convex Analysis* (2021).

[10] Y. Koike, Notes on the dimension dependence in high-dimensional central limit theorems for hyperrectangles. *Japanese Journal of Statistics and Data Science* (2020).

[11] B. Klartag, Logarithmically-concave moment measures I. *Geometric Aspects of Functional Analysis*, Lecture Notes in Math. 2116, Springer (2014), 231–260.

[12] B. Klartag, Affine hemispheres of elliptic type, *Algebra i Analiz*, 29:145–188, 2017.

[13] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[14] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* 583–602. Univ. California, Berkeley, CA, 1970/1971. Vol. II: Probability Theory. Univ. California Press, Berkeley, CA, 1972.

[15] C. Stein, *Approximate Computation of Expectations*. Institute of Mathematical Statistics Lecture Notes – Monograph Series 7. Institute of Mathematical Statistics, Hayward, CA, 1986.

[16] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, book, dans Progress in Nonlinear Differential Equations and Their Applications 87, Birkhäuser Basel (2015).

[17] F. Santambrogio, Dealing with moment measures via entropy and optimal transport, *Journal of Functional Analysis*, 271 (2016), no. 2, 418–436.

[18] W. Schoutens, Orthogonal Polynomials in Stein’s Method. *Journal of Mathematical Analysis and Applications*, 2001.

[19] C. Villani, *Topics in Optimal Transportation*, no. 58, American Mathematical Soc., 2003.