Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 261 (2011) 1236–1283

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces

Hsin-Hung Shih [1]

Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung, 811, Taiwan

Received 25 February 2011; accepted 28 April 2011

Available online 11 May 2011

Communicated by L. Gross

Abstract

In this paper, we generalize Stein's method to "*infinite-variate*" normal approximation that is an infinite-dimensional approximation by abstract Wiener measures on a real separable Banach space. We first establish a Stein's identity for abstract Wiener measures and solve the corresponding Stein's equation. Then we will present a Gaussian approximation theorem using exchangeable pairs in an infinite-variate context. As an application, we will derive an explicit error bound of Gaussian approximation to the distribution of a sum of independent and identically distributed Banach space-valued random variables based on a Lindeberg–Lévy type limit theorem. In addition, an analogous of Berry–Esséen type estimate for abstract Wiener measures will be obtained.

Keywords: Abstract Wiener measure; Abstract Wiener space; Gaussian approximation; Stein's method

1. Introduction

In 1972, C. Stein [31] introduced a powerful way of determining the accuracy of normal approximation to the distribution of a sum of dependent random variables, known as Stein's method. So far, the scope of his discovery has expanded rapidly by Chen [6], Barbour [1,2], and many other authors, which showed that Stein's method can be adapted to approximation by a broad class of probability distributions such as the Poisson, compound Poisson, and Gamma

E-mail address: hhshih@nuk.edu.tw.

[1] Research supported by the National Science Council of Taiwan.

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2011.04.016

distributions, as well as extensively applied in a wide range of other fields such as the theory of random graphs, computational molecular biology, etc. For further details, see [1–3,6,29,30,32] and references cited therein.

Recently, many authors extended Stein's method to multivariate (see e.g. the works of Chatterjee and Meckes [5], Meckes [20], Goldstein and Rinott [8], Götze [10], and Nourdin et al. [22]) and functional (see e.g. the works of Barbour [2], Nourdin and Peccati [21], Nourdin et al. [23]) settings for normal approximation, inclusive of the Gaussian approximation of vectors of random variables defined on a fixed Wiener chaos (see [22]). A problem naturally arises: How to extend Stein's method to "*infinite-variate*" normal approximation, by which we mean an approximation by Gaussian measures on infinite-dimensional Banach spaces?

Since, by applying Kuelbs's theorem in [15], every non-degenerate Gaussian measure on a real separable Banach space B with zero mean can be regarded as an abstract Wiener measure, our study will focus on establishing an infinite-dimensional version of Stein's method of exchangeable pairs for approximation by abstract Wiener measures. There the approximation will be performed with respect to a distance in the sense of Wasserstein, which measures the distance between two B-valued random variables X and Y defined by

$$d_\omega(X,Y)=\sup_{\|g\|_{\mathrm{ULip}}\leqslant 1}\bigl|\mathbb{E}\bigl[g(X)\bigr]-\mathbb{E}\bigl[g(Y)\bigr]\bigr|,$$

where g is a real-valued functions on B and $\|g\|_{\mathrm{ULip}}\equiv\sup\{\frac{|g(x)-g(y)|}{\|x-y\|};\ x\neq y\in B\}$. See also Corollary 4.11.

We briefly describe Stein's approach to univariate normal approximation, which was laid out in his famous monograph [32], as follows. First of all, as a simple application of integration by parts formula, it was verified that the standard univariate normal distribution μ can be characterized by the property that for all sufficiently smooth real-valued functions f,

$$\int_{-\infty}^{\infty}\bigl(f''(t)-tf'(t)\bigr)\,\mu(dt)=0. \tag{1.1}$$

The formula (1.1) and the operator $\mathcal{J}f(t)=f''(t)-tf'(t)$ are known, respectively, as Stein's identity and Stein's operator for μ, where putting $g=f'$ in (1.1) gives Stein's original characterization (see Stein [32, Lemma 1 in p. 21]). Next, for any h in a class $\mathcal{H}$ of test functions, find a solution f_h of the so-called Stein's equation

$$h(t)-\int_{-\infty}^{\infty}h(u)\,\mu(du)=\mathcal{J}f_h(t). \tag{1.2}$$

Hence, for any real-valued random variable W, the distance between W and μ can be computed as differences of expectations of test functions by applying the following formula:

$$d_{\mathcal{H}}(W,\mu):=\sup_{h\in\mathcal{H}}\left|\mathbb{E}\bigl[h(W)\bigr]-\int_{-\infty}^{\infty}h(t)\,\mu(dt)\right|=\sup_{h\in\mathcal{H}}\bigl|\mathbb{E}\bigl[\mathcal{J}f_h(W)\bigr]\bigr|. \tag{1.3}$$

Subsequently, in order to derive an explicit bound of (1.3), Stein introduced a useful method of exchangeable pairs by constructing an auxiliary random variable W' from W on the same probability space as W such that (W, W') is an exchangeable pair, where W' is usually given by making a small random change of W, and is close to W. For example, Stein assumed that W' has the linear regression property $\mathbb{E}[W' - W \mid W] = -\lambda W$ with $0 < \lambda < 1$. Then, applying Taylor's theorem to $f_h(W') - f_h(W)$ around W and taking expectations, Stein proved that

$$d_{\mathcal{H}}(W, \mu) \leqslant \frac{1}{\lambda} \sup_{h \in \mathcal{H}} \mathbb{E}\big[\big|f_h''(W)\big|\big|\big(W' - W\big)^2 - \lambda\big|\big] + \frac{1}{\lambda} \sup_{h \in \mathcal{H}} \mathbb{E}\big[\big|R_h\big(W', W\big)\big|\big], \tag{1.4}$$

where $R_h(W', W)$ is the second-order remainder term. See [32, Lecture III] for thorough treatment of the above estimation.

In this paper, we will devote ourself to establish an infinite-dimensional counterpart of Stein's method of exchangeable pairs for Gaussian approximation by abstract Wiener measures along the line of Stein's approach from (1.1) to (1.4). We organize this paper as follows. In Section 2, it will be divided into two parts. In the one part, we will review the background concerning the concepts of abstract Wiener space. For details, see Gross [11,12] and Kuo [17]. In the other part, we will discuss the relationship between the notions of Gaussian measure on a Banach space and abstract Wiener space. As a biproduct, we will show that there is associated a setting of a sequence of countably Hilbert spaces on every abstract Wiener space (i, H, B). Such a framework will be necessary in our investigation. In Section 3, we present a characterization of abstract Wiener measures, which is an infinite-dimensional version of Stein's identity in (1.1) (see Theorem 3.1). For the associated Stein's equation as one in (1.2), we will solve it in the space of scalar-valued uniformly Lip-1 functions and discuss the related estimations concerning the first and second Gross derivatives of the solution of Stein's equation in Section 4 (see Theorems 4.8–4.10). An infinite-dimensional version of Stein's method of exchangeable pairs will be constructed in Section 5 (see Theorem 5.2). As an application, we will derive an explicit error bound of Gaussian approximation to the distribution of a sum of independent and identically distributed Banach space-valued random variables based on a Lindeberg–Lévy type limit theorem (see Theorem 5.5). In addition, an analogous of Berry–Esséen type estimate for abstract Wiener measures will be obtained in Theorem 5.8.

Notations. Throughout this paper, we adopt the following notations:

- For a real Banach space U, $\mathcal{B}(U)$ denotes the Borel σ-algebra of U.
- For any bounded linear mapping T from U into another Banach space V, we use the symbol $\|T\|_{U,V}$ for the operator norm of T. If E is a subspace of U, the restriction of T to E is denoted by $T|_E$. Denote by U^* the dual space of U with the $\|\cdot\|_{U^*}$-norm, where $\|\cdot\|_{U^*} = \|\cdot\|_{U,\mathbb{R}}$. For any $f \in U^*$ and $x \in U$, $(x, f)_{U,U^*}$ means the U–U^* dual pairing. If U is separable, a well-known fact is that $\mathcal{B}(U)$ coincides with the σ-algebra generated by all cylinder sets in U, which are sets of the form $\{x \in U;\ ((x, \eta_1)_{U,U^*}, (x, \eta_2)_{U,U^*}, \ldots, (x, \eta_n)_{U,U^*}) \in D\}$ for any $\eta_1, \eta_2, \ldots, \eta_n \in U^*$, $D \in \mathcal{B}(\mathbb{R}^n)$, and $n \in \mathbb{N}$ (see [17]).
- Let Σ be a real separable Hilbert space. Then $\mathcal{TR}(\Sigma)$ stands for the Banach space of all trace-class operators from Σ to itself with $\|\cdot\|_{\mathrm{tr}(\Sigma)}$-norm, as well as $\mathcal{HS}(\Sigma)$ is the Hilbert space of all Hilbert–Schmidt operators from Σ to itself with the inner product $\langle\langle A, B\rangle\rangle_{\mathrm{HS}(\Sigma)} = \mathrm{Tr}_\Sigma(B^*A)$ and the corresponding Hilbert–Schmidt norm $\|\cdot\|_{\mathrm{HS}(\Sigma)}$, where $\mathrm{Tr}_\Sigma(S)$ is the trace of a trace-class operator S on Σ.

2. Gaussian measure on a Banach space and abstract Wiener space

In this section, we first review some basic notions of abstract Wiener space, introduced by L. Gross [11], and the related results which will be used in the subsequent study. For further details, we refer the reader to [1,12,17]. Secondly, combining [27, Theorem 2.7] with the idea of Kuelbs in [15], we will present a relationship among the notions of Gaussian measure on a Banach space, reproducing kernel Hilbert space, and abstract Wiener space.

2.1. Abstract Wiener space

Let H be a given real separable Hilbert space with $|\cdot|_0$-norm induced by the inner product $\langle\cdot,\cdot\rangle_0$, and $\|\cdot\|$ be another norm defined on H which is weaker than the $|\cdot|_0$-norm. If $\|\cdot\|$-norm is measurable on H, introduced by L. Gross (see [11]), then the triple (i, H, B) is called an abstract Wiener space (AWS, in short), where B is the completion of H with respect to $\|\cdot\|$-norm and i is the canonical embedding of H into B. As H is identified as a dense subspace of B, we identify B^* as a dense subspace of H^* under the adjoint operator i^* of i by the following way: For any $x \in H$ and $\eta \in B^*$, $\langle x, i^*(\eta)\rangle_0 = (i(x), \eta)_{B,B^*}$. Applying the Riesz representation theorem to identify H^* with H (denoted by $H^* \approx H$), we have the continuous inclusion maps $B^* \subset H \subset B$. In addition, B carries a probability measure p_t, known as the abstract Wiener measure with variance parameter $t > 0$, which is characterized as the Borel measure on B such that for any $\eta \in B^*$,

$$\int_B e^{\mathrm{i}(x,\eta)_{B,B^*}}\, p_t(dx) = e^{-\frac{t}{2}|\eta|_0^2}. \tag{2.1}$$

Remark 2.1. Let $\{\eta_n\} \subset B^*$ be a countable dense subset of B. Obviously, for any $r > 0$, $B = \bigcup_n N(\eta_n, r)$, where $N(x, r)$ is the open ball of radius $r > 0$ and center $x \in B$. A well-known fact is that for any $t, r > 0$ and $\phi_1, \phi_2 \in B^*$,

$$p_t\big(N(\phi_1, r)\big) = \int_{N(\phi_2,r)} \exp\left\{-\frac{|\phi_1 - \phi_2|_0^2}{2t} - \frac{1}{t}(x, \phi_1 - \phi_2)_{B,B^*}\right\} p_t(dx).$$

Hence p_t's are all non-degenerate, that is, every nonempty open set of B has positive p_t-measure.

From (2.1), $(\cdot, \eta)_{B,B^*}$ is a random variable on $(B, \mathcal{B}(B), p_t)$ with mean zero and variance $t|\eta|_0^2$. For any $h \in H$, let $\{\eta_n\}$ be a sequence in B^* such that $|\eta_n - h|_0 \to 0$ as $n \to \infty$. Then $\{(\cdot, \eta_n)_{B,B^*}\}$ forms a Cauchy sequence in $L^2(B, p_t)$, the $L^2(B, p_t)$-limit of which is denoted by $\langle\cdot, h\rangle_0$. One notes that $\langle\cdot, h\rangle_0$ is independent of the choice of $\{\eta_n\}$ and distributed by the law of $N(0, t|h|_0^2)$.

Lemma 2.2. *(See Kuo [17].) There exist another AWS (i_0, H, B_0) and an increasing sequence of orthogonal projections $\{P_n\}$ converging strongly to the identity in H such that*

(i) $\|h\| \leqslant \|h\|_0$ *for any $h \in H$, where $\|\cdot\|_0$ denotes the B_0-norm,*

(ii) *each* P_n *extends by continuity to a projection* $\tilde{P}_n$ *of* B_0 *such that* $\|\tilde{P}_n(x)\|_0 \leqslant \|x\|_0$ *for any* $x \in B_0$*, and*
(iii) $\tilde{P}_n$ *converges strongly to the identity in* B_0 *with respect to* $\|\cdot\|_0$*-norm.*

Let $\tilde{p}_t$ be the abstract Wiener measure on B_0 with variance parameter $t > 0$. For any $x \in B_0$, define $\rho(x) = \lim_{n\to\infty} i(\eta_n)$ with respect to $\|\cdot\|$-norm, where $\{\eta_n\}$ is a sequence in H such that $\|i_0(\eta_n) - x\|_0 \to 0$, as $n \to \infty$. Note that the definition of ρ is independent of the choice of $\{\eta_n\}$. Then ρ is a bounded linear operator from B_0 into B with $\|\rho\|_{B_0,B}$ not greater than 1. Clearly, $i(h) = (\rho \circ i_0)(h)$ for any $h \in H$, where "$\circ$" means the composition of functions. Moreover, for any $h \in H$ and $\eta \in B^*$,

$$\langle h, i^*(\eta)\rangle_0 = \big(i(h), \eta\big)_{B,B^*} = \big((\rho \circ i_0)(h), \eta\big)_{B,B^*} = \big(i_0(h), \eta \circ \rho\big)_{B_0,B_0^*} = \langle h, i_0^*(\eta \circ \rho)\rangle_0.$$

We thus conclude that $i^*(\eta) = i_0^*(\eta \circ \rho)$ for any $\eta \in B^*$, which implies that for all nonnegative measurable or integrable functions f on B,

$$\int_B f(x)\, p_t(dx) = \int_{B_0} (f \circ \rho)(x)\, \tilde{p}_t(dx). \tag{2.2}$$

Remark 2.3.

(1) The norms $\|\cdot\|$ and $\|\cdot\|_0$ are comparable. However, from Kuo's construction of $\|\cdot\|_0$-norm in the proof of [17, Chapter I, Corollary 4.2], it is not clear to us whether $\|\cdot\|_0$-norm and $\|\cdot\|$-norm are compatible or not. In other words, we cannot exclude this possibility that ρ is not one-to-one. Certainly, if these two norms are compatible, $\rho(B_0)$ is a Borel dense subset of B, since ρ is one-to-one and $(B_0, \|\cdot\|_0)$ and $(B, \|\cdot\|)$ are both standard measurable spaces (see [14,34]). In this case, B_0 is identified with $\rho(B_0)$, and hence it follows from (2.2) that $p_t(B_0) = 1$.
(2) For those orthogonal projections P_n in Lemma 2.2, $P_n(H)$ is contained in the dual space B_0^* of B_0. In fact, since $P_n(H)$ is finite-dimensional, there is a constant $c > 0$ such that, for any $h \in P_n(H)$, $|h|_0 \leqslant c\|h\|_0$. Then, for any $x, h \in H$, we have

$$\begin{aligned} \big|\langle x, P_n(h)\rangle_0\big| &= \big|\langle P_n(x), h\rangle_0\big| = \big|\langle \tilde{P}_n(x), h\rangle_0\big| \\ &\leqslant \big|\tilde{P}_n(x)\big|_0 |h|_0 \leqslant c\big\|\tilde{P}_n(x)\big\|_0 |h|_0 \leqslant c\|x\|_0 |h|_0, \end{aligned}$$

implying $P_n(h) \in B_0^*$ and $\|P_n(h)\|_{B_0^*} \leqslant c|h|_0$.

A test operator T on B is a bounded operator of finite rank from B to B with range in B^* (see [12]). Then $T|_H$ can be regarded as a bounded linear operator from H into itself. The following properties concerning test operators are useful in estimating solutions of Stein's equation for abstract Wiener measures.

Proposition 2.4. *(See Gross [12].)*

(i) *The set* $\mathcal{T}(H)$ *of restrictions of test operators to* H *is dense in the space* $\mathcal{K}(H)$ *of compact operators on* H*.*

(ii) *Let T be a test operator on B. Then, for any $r \geqslant 1$,*

$$\int_B \|Tx\|^r \, p_t(dx) \leqslant \|T|_H\|_{H,H}^r \int_B \|x\|^r \, p_t(dx).$$

Remark 2.5.

(1) For any test operator T on B, $T^* = (T|_H)^*$ on B^*, where $T^* : B^* \to B^*$ is the adjoint operator of T, and $(T|_H)^*$ is the Hilbert adjoint operator of $T|_H$. In fact, for any $\eta \in B^*$ and $h \in H$,

$$\langle h, (T|_H)^*(\eta)\rangle_0 = \langle T|_H(h), \eta\rangle_0 = \big(T(h), \eta\big)_{B,B^*} = \big(h, T^*(\eta)\big)_{B,B^*} = \langle h, T^*(\eta)\rangle_0.$$

(2) $\mathcal{T}(H)$ is dense in $\mathcal{HS}(H)$. In fact, let $\{e_1, e_2, \ldots\} \subset B^*$ be an orthonormal basis of H. Take a sequence of test operators on B, say $\{T_n\}$, where $T_n(x) = \sum_{j=1}^n (x, e_j)_{B,B^*} e_j$, $x \in B$. For any $A \in \mathcal{HS}(H)$, set $Q_n = T_n \circ A \circ T_n$. Then Q_n's are all test operators and, for any $n \in \mathbb{N}$,

$$\begin{aligned} \|A - Q_n|_H\|_{\mathrm{HS}(H)} &\leqslant \|A - T_n \circ A\|_{\mathrm{HS}(H)} + \|T_n \circ A - Q_n|_H\|_{\mathrm{HS}(H)} \\ &\leqslant \|A - T_n \circ A\|_{\mathrm{HS}(H)} + \|T_n|_H\|_{H,H} \|A - A \circ T_n|_H\|_{\mathrm{HS}(H)} \\ &\leqslant \left\{ \sum_{j=n+1}^\infty \left|A^*(e_j)\right|_0^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{j=n+1}^\infty \left|A(e_j)\right|_0^2 \right\}^{\frac{1}{2}}, \end{aligned}$$

which approaches to zero as n tends to infinity.

Gross differentiation. In [12], L. Gross introduced the notion of H-differentiation in H-direction as follows. Let f be a function defined from an open set U of B into a Banach space W. Then f is said to be H-differentiable at a point $x \in U$ if the mapping $\phi(h) = f(x+h)$, $h \in H$, regarded as a function defined in a neighborhood of the origin of H is Fréchet differentiable at 0. The Fréchet derivative $\phi'(0)$ at $0 \in H$ is called the H-derivative of f at $x \in B$. In notation, we denote the H-derivative of f at x in the direction $h \in H$ by $\langle Df(x), h\rangle$. The k-th order H-derivatives of f at x are defined inductively and denoted by $D^k f(x)$ for $k \geqslant 2$ if they exist. One notes that $D^k f(x)$ is a bounded k-linear mapping from the Cartesian product $H \times \cdots \times H$ of k copies of H into W for any $k \in \mathbb{N}$. In particular, as f is scalar-valued, $Df(x) \in H^* \approx H$ and $D^2 f(x)$ is regarded as a bounded linear operator from H into $H^* \approx H$ for any $x \in U$, where the notation $\langle D^2 f(x)h, k\rangle$ means $D^2 f(x)(h,k)$, $h, k \in H$. Further, if $D^2 f(x)$ is a trace-class operator on H, we define the Gross Laplacian $\Delta_G f(x)$ of f at x by $\Delta_G f(x) = \mathrm{Tr}_H(D^2 f(x))$.

For an H-derivative $Df(x)$ at $x \in B$, we say $Df(x)$ determines an element in B^* (abbreviated in form $Df(x) \in B^*$) if there is a constant $C > 0$ such that $|\langle Df(x), h\rangle| \leqslant C\|h\|$ for any $h \in H$. In this case, $Df(x)$ defines an element of B^* by continuity, and we will still denote this linear functional by $Df(x)$. Further, if f is twice Fréchet differentiable on B, then $Df(x)$ equals the first-order Fréchet derivative $f'(x)$ at $x \in B$ and is automatically in B^*. And, $D^2 f(x)$ equals the restriction of the second-order Fréchet derivative $f''(x)$ to $H \times H$ at $x \in B$. In this circumstance, since, for any $x \in B$, $f''(x)$ is a bounded linear operator from B into B^*, the following Goodman's theorem implies that $D^2 f(x)$ is a trace-class operator on H and then $\Delta_G f(x)$ is immediately obtained.

Theorem 2.6 *(Goodman). (See [17, Chapter I, Theorem 4.6].) Let A be a bounded linear operator from B into itself with range in B^*. Then A is a trace-class operator on H. Moreover,*

$$\|A\|_{\mathrm{tr}(H)} \leqslant \|A\|_{B,B^*} \int_B \|x\|^2 \, p_1(dx).$$

For each $x \in B$ and $t > 0$, define $p_t(x, E) = p_t(E - x)$, $E \in \mathcal{B}(B)$. Then, for each $E \in \mathcal{B}(B)$, the mapping $x \in B \mapsto p_t(x, E)$ is measurable, and the family $\{p_t(x, \cdot);\ t > 0,\ x \in B\}$ forms a semigroup under the convolution $*$ of measures. In fact,

$$p_s(x, \cdot) * p_t(y, \cdot) = p_{s+t}(x + y, \cdot) \quad \text{for any } x, y \in B \text{ and } s, t > 0.$$

Proposition 2.7. *(See Lee [19].) Let $x \in B$ be fixed and f a real-valued function defined on B. Assume that $f \in L^\alpha(B, p_t(x, dy))$ for some $\alpha > 1$. Then $p_t f(x) \equiv \int_B f(y) p_t(x, dy)$ is infinitely H-differentiable at x.*

2.2. Gaussian measure on a Banach space

Let B be a real separable Banach space with $\|\cdot\|$-norm. A Borel measure ν on B is said to be Gaussian with mean zero if the law of arbitrary $\eta \in B^*$ considered as a random variable on $(B, \mathcal{B}(B), \nu)$ has a normal distribution in $\mathbb{R}$ with mean zero.

Let B_s be the intersection of all closed subspaces of B with ν-measure 1. By the separability of B, $\nu(B_s) = 1$. Hereafter, we restrict the measure ν to $(B_s, \|\cdot\|)$. Let $|\cdot|_{-1}$ be another norm on B_s, which is induced by an inner product $\langle\cdot,\cdot\rangle_{-1}$ and weaker than $\|\cdot\|$-norm. Upon completing B_s with respect to $|\cdot|_{-1}$-norm, we obtain a separable Hilbert space K with the inner product $\langle\cdot,\cdot\rangle_{-1}$. Since $(B_s, \|\cdot\|)$ and $(K, |\cdot|_{-1})$ are both standard measurable spaces and the canonical embedding $i_{B_s,K}$ from B_s into K is continuous, $\mathcal{B}(B_s) = \mathcal{B}(K) \cap B_s$ (see [14,34]).

Set $\nu^K(E) = \nu(E \cap B_s)$ for any $E \in \mathcal{B}(K)$. For any $k \in K$, there is a unique $\phi_k \in B_s^*$ such that $(x, \phi_k)_{B_s,B_s^*}$ equals $\langle x, k\rangle_{-1}$ for any $x \in B_s$, from which it follows that

$$\int_K e^{\mathrm{i}\langle x,k\rangle_{-1}} \, \nu^K(dx) = \int_{B_s} e^{\mathrm{i}\langle x,k\rangle_{-1}} \, \nu(dx) = \int_{B_s} e^{\mathrm{i}(x,\phi_k)_{B_s,B_s^*}} \, \nu(dx),$$

whence $\langle\cdot, k\rangle_{-1}$ is normally distributed with mean zero on $(K, \mathcal{B}(K), \nu^K)$. So, ν^K is a Gaussian measure on K with mean zero. By the Prohorov theorem (see e.g. [17]), the covariance operator S_{ν^K} of ν^K is a nonnegative self-adjoint trace-class operator on K. Moreover, S_{ν^K} is one-to-one. In the proof of the Itô theorem in [13], it was shown that the topological support of ν^K, denoted by $\mathrm{supp}(\nu^K)$, is the orthogonal complement of the kernel of S_{ν^K}, whence $\nu(B_s \cap \mathrm{supp}(\nu^K)) = 1$. Since B_s is the least closed subspace of B with ν-measure 1, $B_s \cap \mathrm{supp}(\nu^K) = B_s$, by which we see that $\mathrm{supp}(\nu^K) = K$ and the injectivity of S_{ν^K} immediately follows.

Let $H = \sqrt{S_{\nu^K}}(K)$ endowed with $|\cdot|_0$-norm induced by the inner product $\langle\cdot,\cdot\rangle_0$, where $\langle\sqrt{S_{\nu^K}}(x), \sqrt{S_{\nu^K}}(y)\rangle_0 = \langle x, y\rangle_{-1}$ for any $x, y \in K$. Note that S_{ν^K} has the spectral decomposition $\sum_{j=1}^{\infty} \lambda_j \langle\cdot, k_j\rangle_{-1} k_j$, where λ_j's > 0, $\sum \lambda_j < +\infty$, and $\{k_j;\ j \in \mathbb{N}\}$ is an orthonormal basis of K, the set $\{\lambda^{\frac{1}{2}} k_j;\ j \in \mathbb{N}\}$ is an orthonormal basis of H, whence H is dense in K and the restriction of $\sqrt{S_{\nu^K}}$ to H is a Hilbert–Schmidt operator from H into itself. Accordingly, we can conclude that the triple $(i_{H,K}, H, K)$ forms an AWS, where $i_{H,K}$ is the canonical embedding of H into K (see [17]).

Let $i^*_{H,K}$ be the adjoint operator of $i_{H,K}$. For any $f \in K^*$, there is a unique $k_f \in K$ such that $(x, f)_{K,K^*} = \langle x, k_f\rangle_{-1} = \langle\sqrt{S_{\nu^K}}(x), \sqrt{S_{\nu^K}}(k_f)\rangle_0 = \langle x, S_{\nu^K}(k_f)\rangle_0$ for any $x \in H$, from which it follows that $i^*_{H,K}(f) = S_{\nu^K}(k_f)$ and

$$\int_K e^{\mathrm{i}(x,f)_{K,K^*}}\,\nu^K(dx) = e^{-\frac{1}{2}\langle S_{\nu^K}(k_f),k_f\rangle_{-1}} = e^{-\frac{1}{2}|S_{\nu^K}(k_f)|_0^2} = e^{-\frac{1}{2}|i^*_{H,K}(f)|_0^2}. \tag{2.3}$$

So, ν^K is the associated abstract Wiener measure of $(i_{H,K}, H, K)$ with variance parameter 1. Utilizing the fact that the translation measure of ν^K by $x \in K$ is equivalent to ν^K if and only if $x \in H$ (see [17]), together with $\nu^K(B_s) = 1$, we can deduce the following inclusion maps: $H \subset B_s \subset K$.

Remark 2.8. Such a Hilbert space K always exists. See the proof of Lemma 2.1 in [15], where a general method to construct K from B was presented by J. Kuelbs.

On the other hand, let $\mathcal{L}$ be the closure of the linear manifold $\{(\cdot,\eta);\ \eta \in B_s^*\}$ in $L^2(B_s,\nu)$ and H_ν the linear subspace $\{\int_{B_s} x\varphi(x)\,\nu(dx);\ \varphi \in \mathcal{L}\}$ of B_s, where the integrals inside the brace exist as B_s-valued Bochner integrals by using the Fernique theorem [27, Theorem 2.6] and the Cauchy–Schwarz inequality. Define an inner product on H_ν by

$$\left\langle \int_{B_s} x\varphi(x)\,\nu(dx), \int_{B_s} x\psi(x)\,\nu(dx)\right\rangle_\nu := \int_{B_s} \varphi(x)\psi(x)\,\nu(dx), \quad \forall \varphi,\psi \in \mathcal{L}.$$

We remark that such an inner product is meaningful since, for $\varphi \in \mathcal{L}$, $\int_{B_s} x\varphi(x)\,\nu(dx) = 0$ if and only if $\varphi = 0$ $[\nu]$-a.e. on B_s. In particular, for $\eta \in B_s^*$, $\int_{B_s} x(x,\eta)_{B_s,B_s^*}\,\nu(dx) = 0$ if and only if $(x,\eta)_{B_s,B_s^*} = 0$ everywhere in $x \in B_s$, since every proper closed subspace of B_s has ν-measure less than 1. Then $(H_\nu, \langle\cdot,\cdot\rangle_\nu)$ is a Hilbert space, and B_s^* can be regarded as a dense subspace of H_ν by identifying arbitrary $\eta \in B_s^*$ with $\int_{B_s} x(x,\eta)_{B_s,B_s^*}\,\nu(dx)$.

Denote by $|\cdot|_\nu$ the norm of H_ν. Observe that for any $\varphi \in \mathcal{L}$,

$$\left\| \int_{B_s} x\varphi(x)\,\nu(dx)\right\| \leqslant \left\{\int_{B_s} \|x\|^2\,\nu(dx)\right\}^{\frac{1}{2}} \left|\int_{B_s} \varphi(x)\,\nu(dx)\right|_\nu,$$

and, for any $\eta \in B_s^*$, $|\eta|_\nu \leqslant \{\int_{B_s} \|x\|^2\,\nu(dx)\}^{1/2}\|\eta\|_{B_s^*}$. Therefore, we have the continuous inclusion maps:

$$K^* \subset B_s^* \subset H_\nu \subset B_s \subset K,$$

where K^* is regarded as a dense subspace of B_s^* by identifying arbitrary $f \in K^*$ with $f|_{B_s}$. Moreover, for each $\eta \in B_s^*$,

$$\int_{B_s} e^{\mathrm{i}(x,\eta)_{B_s,B_s^*}}\,\nu(dx) = e^{-\frac{1}{2}\int_{B_s}(x,\eta)^2_{B_s,B_s^*}\,\nu(dx)} = e^{-\frac{1}{2}|\eta|_\nu^2}.$$

Consequently, H_ν is exactly the unique reproducing kernel Hilbert space (RKHS, in short) for ν in the sense of [27, Theorem 2.7], and ν is the σ-additive extension of the canonical Gaussian cylinder set measure ν_{H_ν} to $\mathcal{B}(B_s)$, where ν_{H_ν} is a finitely additive nonnegative set function on $(H_\nu, \mathcal{B}(H_\nu))$ such that

$$\nu_{H_\nu}\left(\left\{x \in H_\nu;\ \langle x, h\rangle_\nu \leqslant a\right\}\right) = \frac{1}{\sqrt{2\pi}|h|_\nu} \int_{-\infty}^{a} \exp\left\{-\frac{u^2}{2|h|_\nu}\right\} du, \quad \forall h \in H_\nu.$$

Applying the result of Dudley, Feldman and LeCam in [7], $(i_{H_\nu,B_s}, H_\nu, B_s)$ forms an AWS, and ν is the associated abstract Wiener measure with variance parameter 1, where i_{H_ν,B_s} is the canonical embedding from H_ν into B_s. By Remark 2.1, ν is non-degenerate on B_s, implying that B_s is exactly the topological support of ν on B.

From (2.3) and [27, Theorem 2.7] it follows that H is the unique RKHS for ν^K. On the other hand, since H_ν is also continuously embedded in K and for any $f \in K^*$,

$$\begin{aligned}\int_K e^{\mathrm{i}(x,f)_{K,K^*}}\, \nu^K(dx) &= \int_{B_s} e^{\mathrm{i}(x,f|_{B_s})_{B_s,B_s^*}}\, \nu(dx) \\ &= \exp\left\{-\frac{1}{2}\int_{B_s} (x, f|_{B_s})^2_{B_s,B_s^*}\, \nu(dx)\right\} \\ &= e^{-\frac{1}{2}|f|_{B_s}|_\nu^2},\end{aligned} \tag{2.4}$$

H_ν is also a RKHS for ν^K. Consequently, by uniqueness $H = H_\nu$.

Summing up the above arguments together with [27, Theorem 2.7], we can conclude the following theorem.

Theorem 2.9.

(i) *Assume that a real separable Banach space B with $\|\cdot\|$-norm and ν is a Gaussian measure on B with mean zero. Then the topological support of ν is the least closed subspace of B with ν-measure* 1.

(ii) *Assume that a real separable Banach space B with $\|\cdot\|$-norm is continuously and as a Borel subset embedded in a real Hilbert space K with $|\cdot|_{-1}$-norm. Let ν be a non-degenerate Gaussian measure on B with mean zero and ν^K a Gaussian measure on K given by $\nu^K(E) = \nu(E \cap B)$ for any $E \in \mathcal{B}(K)$.*

  (a) *There is a unique Hilbert space H densely embedded in B as a Borel subset such that the triple (i, H, B) forms an AWS with the associated abstract Wiener measure ν with variance parameter* 1, *where i is the canonical embedding of H into B. More precisely, H is the Hilbert space $\{\int_B x\varphi(x)\,\nu(dx);\ \varphi \in \mathcal{L}\}$ with the inner product $\langle \int_B x\varphi(x)\,\nu(dx), \int_B x\psi(x)\,\nu(dx)\rangle_\nu := \int_B \varphi(x)\psi(x)\,\nu(dx)$ for any $\varphi, \psi \in \mathcal{L}$, where $\mathcal{L}$ is the closure of the linear manifold $\{(\cdot, \eta);\ \eta \in B^*\}$ in $L^2(B, \nu)$.*

  (b) *Let S_{ν^K} be the covariance operator of ν^K. Then S_{ν^K} is one-to-one, and the space H in* (a) *is the same as the Hilbert space $\sqrt{S_{\nu^K}}(K)$ with the usual inner product*

$\langle\sqrt{S_{\nu^K}}(x), \sqrt{S_{\nu^K}}(y)\rangle_0 := \langle x, y\rangle_{-1}$ for any $x, y \in K$, where the triple $(i_{H,K}, H, K)$ forms an AWS with the associated abstract Wiener measure ν^K with variance parameter 1, $i_{H,K}$ being the canonical embedding of H into K. Moreover, the space K^* is isometrically isomorphic to $S_{\nu^K}(K)$ as a Borel dense subset of H through the adjoint operator $i^*_{H,K}$ of $i_{H,K}$, where $S_{\nu^K}(K)$ is a Hilbert space with the inner product $\langle S_{\nu^K}(x), S_{\nu^K}(y)\rangle_1 := \langle x, y\rangle_{-1}$, $x, y \in K$, and for any $f \in K^*$ such that $f = \langle\cdot, k_f\rangle_{-1}$, $i^*_{H,K}(f) = S_{\nu^K}(k_f)$.

Remark 2.10. Let $\mathcal{A} = \sqrt{S_{\nu^K}}^{-1}$. Then $\mathcal{A}$ is a densely defined, self-adjoint linear operator in $(H, |\cdot|_0)$. Moreover, under the identification in the above theorem, $K^* = \{h \in H;\ |\mathcal{A}(h)|_0 < +\infty\}$ and $\langle f, g\rangle_1 = \langle\mathcal{A}(f), \mathcal{A}(g)\rangle_0$ for any $f, g \in K^*$, as well as K is the completion of H with respect to $|h|_{-1} = |\mathcal{A}^{-1}(h)|_0$ for any $h \in H$. In fact, by the standard construction of countably Hilbert spaces from $(H, \mathcal{A})$ (see e.g. [24]), we can get a sequence of compatible Hilbertian norms and then obtain the following chain of dense, continuous embeddings:

$$\cdots \subset H_n \subset \cdots \subset H_1 \subset B^* \subset H \subset B \subset H_{-1}\,(= K) \subset \cdots \subset H_{-n} \subset \cdots,$$

where, for any $n \in \mathbb{N}$, $H_n \equiv \{h \in H;\ |\mathcal{A}^n(h)|_0 < +\infty\}\ (= \sqrt{S_{\nu^K}}^{\,n}(H))$ is a Hilbert space with the inner product $\langle\cdot,\cdot\rangle_n$ and the induced norm $|\cdot|_n$ defined by $|h|_n := |\mathcal{A}^n(h)|_0$, as well as H_{-n} is obtained by completing H with respect to $|\cdot|_{-n}$-norm associated with the inner product $\langle\cdot,\cdot\rangle_{-n}$ defined by $\langle h, k\rangle_{-n} = \langle\mathcal{A}^{-n}(h), \mathcal{A}^{-n}(k)\rangle_0$, $h, k \in H$. Note that $\{e_j \equiv \lambda^{\frac{1}{2}} k_j;\ j \in \mathbb{N}\}$ is an orthonormal basis of H, whence $\{\lambda_j^{\frac{r}{2}} e_j;\ j \in \mathbb{N}\}$ is an orthonormal basis of H_r for any $r \in \mathbb{Z}$. According to the spectral theory, we can extend S_{ν^K} to the whole H_{-n} for all $n \in \mathbb{N}$ by the way that if $\sum_j a_j^2 < +\infty$,

$$S_{\nu^K}\left(\sum_j a_j \cdot \lambda_j^{-\frac{n}{2}} e_j\right) := \sum_j (a_j\lambda_j) \cdot \lambda_j^{-\frac{n}{2}} e_j.$$

Consequently, for any $n \in \mathbb{N}$, $H_{-n+1} = \sqrt{S_{\nu^K}}(H_{-n})$ $(H_0 \equiv H)$, as well as H^*_{-n} is isometrically isomorphic to H_n as a Borel dense subset of H through the adjoint operator $i^*_{H,H_{-n}}$ of the canonical embedding $i_{H,H_{-n}}$ from H into H_{-n}, where, for any $f \in H^*_{-n}$ with $f = \langle\cdot, k_f\rangle_{-n}$ and $h \in H$,

$$\langle h, i^*_{H,H_{-n}}(f)\rangle_0 = \big(i_{H,H_{-n}}(h), f\big)_{H_{-n}, H^*_{-n}} = \langle h, k_f\rangle_{-n} = \langle h, S^n_{\nu^K}(k_f)\rangle_0,$$

implying $i^*_{H,H_{-n}}(f) = S^n_{\nu^K}(k_f)$, and moreover, $\|f\|_{H^*_{-n}} = |k_f|_{-n} = |S^n_{\nu^K}(k_f)|_n$.

Corollary 2.11. *Let B be a real separable Banach space with $\|\cdot\|$-norm and $(H, |\cdot|_0)$ a Hilbert space densely embedded in B such that the triple (i, H, B) is an AWS, where i is the canonical embedding of H into B. Then*

$$\|i\|_{H,B} \leqslant \left\{\int_B \|x\|^2\, p_1(dx)\right\}^{\frac{1}{2}}.$$

Proof. For any $h \in H$, we see by Theorem 2.9 that there is a unique $\varphi_h \in \mathcal{L}$ such that

$$h = \int_B x\varphi_h(x)\, p_1(dx) \quad \text{with } |h|_0^2 = \int_B \left|\varphi_h(x)\right|^2 p_1(dx). \tag{2.5}$$

Then the desired inequality is immediately obtained by applying the Cauchy–Schwarz inequality to (2.5). □

3. Characterization of abstract Wiener measures

Everywhere below B will be assumed to be a real separable Banach space with $\|\cdot\|$-norm, and Z a fixed B-valued random variable on some probability space $(\Omega, \mathcal{F}, \mathcal{P})$ such that the distribution μ_Z of Z is a Gaussian measure in B with mean zero.

We may assume that μ_Z is non-degenerate; otherwise, by Theorem 2.9, we will replace B by the topological support of μ_Z. Let H be the unique RKHS for μ_Z with the inner product $\langle\cdot,\cdot\rangle_0$ and the induced norm $|\cdot|_0$. Then it follows by Theorem 2.9 that the triple (i, H, B) forms an AWS and μ_Z is the associated abstract Wiener measure on B with variance parameter 1.

In order to establish an infinite-dimensional version of Stein's method for Gaussian approximation, as mentioned in Section 1, the first step should be to look for a suitable characterizing operator $\mathcal{A}_Z$ for μ_Z. Such an operator is defined on a sufficiently large class $\mathcal{D}$ of complex-valued functions on B such that (1) a B-valued random variable Y has the same distribution as Z if and only if $\mathbb{E}[\mathcal{A}_Z(f(Y))] = 0$ for all f belonging to $\mathcal{D}$; (2) for each test function h, there is a function f_h belonging to $\mathcal{D}$ solving the equation $\mathcal{A}_Z f = h - \mathbb{E}[h(Z)]$ with unknown function f. When B is the Euclidean space $\mathbb{R}^n$, $\mathcal{A}_Z$ is known as a Stein operator for the standard (multivariate) normal distribution. Among many of approaches to find a Stein operator for one-dimensional distributions, the generator approach developed by Barbour [1,2] seems to be readily extended to our infinite-variate distribution μ_Z. Based on Barbour's idea, we proceed with finding the characterizing operator $\mathcal{A}_Z$ as follows.

For each $t \geqslant 0$, let $\mathcal{O}_t$ be the mapping from $B \times \mathcal{B}(B)$ into $[0, 1]$ given by

$$\mathcal{O}_t(x, E) \equiv p_{1-e^{-2t}}\big(e^{-t}x, E\big) = \int_B \mathbf{1}_E\big(e^{-t}x + \sqrt{1-e^{-2t}}\,y\big)\, \mu_Z(dy).$$

Then, for each $E \in \mathcal{B}(B)$, the mapping $x \in B \mapsto \mathcal{O}_t(x, E)$ is $\mathcal{B}(B)$-measurable, and, for each $x \in B$, $\{\mathcal{O}_t(x, \cdot);\ t \geqslant 0\}$ is a family of probability measures on $\mathcal{B}(B)$ satisfying the Chapman–Kolmogorov equations:

$$\int_B \mathcal{O}_s(y, E)\mathcal{O}_t(x, dy) = \mathcal{O}_{s+t}(x, E), \quad \forall s, t \geqslant 0.$$

Thus $\{\mathcal{O}_t(\cdot,\cdot);\ t \geqslant 0\}$ forms a temporally homogeneous Markov transition family. For $a \in B$, it associates a family of probability measures $\{\mathcal{O}_{t_1,t_2,\ldots,t_n};\ 0 \leqslant t_1 < t_2 < \cdots < t_n,\ n \in \mathbb{N}\}$, where $\mathcal{O}_{t_1,\ldots,t_n}$ is a probability measures on the product space of n copies of $(B, \mathcal{B}(B))$ given by, for any $E_1, \ldots, E_n \in \mathcal{B}(B)$,

$$\mathcal{O}_{t_1,\ldots,t_n}(E_1 \times \cdots \times E_n) = \int_{E_1} \cdots \int_{E_n} \mathcal{O}_{t_n - t_{n-1}}(y_{n-1}, dy_n) \cdots \mathcal{O}_{t_2 - t_1}(x_1, dx_2)\, \mathcal{O}_{t_1}(a, dx_1).$$

It is easy to verify that such a family satisfies Kolmogorov's consistency condition, and, by the Kolmogorov existence theorem (see [33, Theorem 7.11]), there exists a probability measure Λ_a on $(\Omega, \mathcal{F})$ such that $\Theta = \{\Theta(t);\ t \geqslant 0\}$ on that space is a B-valued temporally homogeneous Markov process having $\mathcal{O}_{t_1,\ldots,t_n}$ as finite-dimensional distributions, where we take Ω to be the set of all mappings from $[0, \infty)$ to B, $\mathcal{F}$ to be the σ-field generated by cylinder sets, and $\Theta(t; \omega) = \omega(t)$ for any $t \geqslant 0$, $\omega \in \Omega$. We call such a process Θ a canonical B-valued Ornstein–Uhlenbeck process starting at the point a. Notice that $\Lambda_a(\{\Theta(t) \in dy\}) = \mathcal{O}_t(a, dy)$ and the transition probability $\Lambda_a(\Theta(t) \in dy \mid \Theta(s))$ is $\mathcal{O}_{t-s}(\Theta(s; \cdot), dy)$ for any $0 \leqslant s \leqslant t$.

For any $t \geqslant 0$, let $\mathcal{T}_t$ be the transition operator of Θ. So, for each $\mathcal{B}(B)$-measurable function f and $x \in B$, $\mathcal{T}_t f(x) = \mathbb{E}[f(\Theta(t)) \mid \Theta(0) = x]$ with respect to Λ_a. In fact,

$$\mathcal{T}_t f(x) \equiv \int_B f(y)\mathcal{O}_t(x, dy) = \int_B f\big(e^{-t}x + \sqrt{1 - e^{-2t}}\, y\big)\, \mu_Z(dy), \quad t \geqslant 0,$$

provided that such an integral exists. The family $\{\mathcal{T}_t;\ t \geqslant 0\}$ yields a strongly continuous contraction semigroup on $L_c^\alpha(B, \mu_Z)$ for any $1 \leqslant \alpha \leqslant \infty$ (see [4, Theorem 2.9.1]). Let $\mathcal{L}_\alpha$ be the infinitesimal generator of $\{\mathcal{T}_t;\ t \geqslant 0\}$ on $L_c^\alpha(B, \mu_Z)$. Since μ_Z is a unique invariant measure for $\{\mathcal{O}_t(\cdot, \cdot);\ t \geqslant 0\}$ (see [4]), we see that $\mathbb{E}[\mathcal{T}_t f(Z)] = \mathbb{E}[f(Z)]$ for any $f \in L_c^\alpha(B, \mu_Z)$, $t \geqslant 0$, which implies that

$$\mathbb{E}\big[\mathcal{L}_\alpha\big(h(Z)\big)\big] = 0, \tag{3.1}$$

for any h belonging to certain dense domain $\mathrm{Dom}(\mathcal{L}_\alpha)$ of $\mathcal{L}_\alpha$ in $L_c^\alpha(B, \mu_Z)$.

By using standard results about strongly continuous contraction semigroups, the Bochner integral $\int_0^t \mathcal{T}_u h\, du$ exists. In fact, it is in $\mathrm{Dom}(\mathcal{L}_\alpha)$, and satisfies the following equality

$$\mathcal{L}_\alpha\left(\int_0^t \mathcal{T}_u h\, du\right) = \mathcal{T}_t h - h, \quad h \in \mathrm{Dom}(\mathcal{L}_\alpha). \tag{3.2}$$

Formally, $\lim_{u \to \infty} \mathcal{T}_u h(x) = \mathbb{E}[h(Z)]$ for any $x \in B$, and so letting t approach to infinity on both sides of (3.2), $f_h(x) \equiv -\int_0^\infty (\mathcal{T}_u h(x) - \mathbb{E}[h(Z)])\, du$, $x \in B$, may serve as a solution of the following equation (with unknown function f)

$$\mathcal{L}_\alpha f = h - \mathbb{E}\big[h(Z)\big]. \tag{3.3}$$

See e.g. [2,29].

In view of Barbour's approach, the roles of Stein's lemma and Stein's equation for μ_Z should be played respectively by (3.1) and (3.2) as well as $\mathcal{L}_\alpha$ is a candidate for $\mathcal{A}_Z$. As $\alpha = 2$, $-\mathcal{L}_\alpha$ is known as the number operator. In [26], Piech proved that if $f \in L_c^2(B, \mu_Z)$ such that $|Df(x)|_0$

exists for μ_Z-a.e. $x \in B$ and is in $L_c^2(B, \mu_Z)$ and the Hilbert–Schmidt norm of $D^2 f(x)$ is finite for μ_Z-a.e. $x \in B$ and is in $L_c^2(B, \mu_Z)$, then f belongs to $\mathrm{Dom}(\mathcal{L}_2)$, as well as

$$-\mathcal{L}_2 f(x) = \big(x, Df(x)\big)_{B,B^*} - \Delta_G f(x), \quad x \in B, \tag{3.4}$$

provided that $Df(x) \in B^*$ and $D^2 f(x)$ is a trace-class operator on H. It is worth noting that Barbour [2] also derived a representation of $-\mathcal{L}_2$ of the form analogous to (3.4). In that paper, Barbour considered B as the Banach space $C[0,1]$ with the supremum norm $\|\cdot\|_\infty$, consisting of all real-valued functions $f(t)$ in $[0,1]$ with $f(0) = 0$, and take $L \subset C[0,1]$ which is the Banach space of those continuous functions $f : C[0,1] \to \mathbb{R}$ for which the norm defined by $\sup_{x \in C[0,1]} \frac{|f(x)|}{1+\|x\|_\infty^3}$ is finite. Then $\{\mathcal{T}_t;\ t \geqslant 0\}$ is a strongly continuous semigroup on L and (3.4) are fulfilled with those twice Fréchet differentiable functions $f \in L$ such that the second-order derivative f'' satisfies a uniformly Lipschitz condition.

Being inspired by (3.1) and (3.4), we are ready to present a characterization of the abstract Wiener measure μ_Z in the following theorem, which also refines the result in [18, Theorem 2.7].

Theorem 3.1. *Let X be a B-valued random variable with the distribution μ_X.*

(i) *If B is finite-dimensional, then $\mu_X = \mu_Z$ if and only if the following identity holds*:

$$\mathbb{E}\big[\big(X, Df(X)\big)_{B,B^*} - \Delta_G f(X)\big] = 0, \tag{3.5}$$

for any twice differentiable function f on B such that $\mathbb{E}[\|D^2 f(Z)\|_{\mathrm{tr}(H)}] < +\infty$.

(ii) *If B is infinite-dimensional, then $\mu_X = \mu_Z$ if and only if the identity* (3.5) *holds for any twice H-differentiable function f on B such that $Df(x) \in B^*$ for any $x \in B$, $\mathbb{E}[\|D^2 f(Z)\|_{\mathrm{tr}(H)}] < +\infty$, and $\mathbb{E}[\|Df(Z)\|_{B^*}^\alpha] < +\infty$ for some $1 < \alpha < +\infty$.*

Proof. (*The "only if" part for statements* (i) *and* (ii).) Let (i_0, H, B_0) and $\{P_n\}$ be respectively the AWS and an increasing sequence of orthogonal projections as given in Lemma 2.2, $\tilde{P}_n$'s be the extension of P_n's to B_0, and $\tilde{p}_1$ be the abstract Wiener measure on B_0 with variance parameter 1. By Remark 2.3, $P_n(H) \subset B_0^*$ for any $n \in \mathbb{N}$, and so there exists an orthonormal basis $\{e_1, e_2, \ldots\} \subset B_0^*$ for H such that

$$\tilde{P}_n(x) = \sum_{j=1}^{k_n} (x, e_j)_{B_0, B_0^*} e_j, \quad \forall x \in B_0,$$

where $k_1 < k_2 < \cdots < k_n < \cdots \nearrow +\infty$.

Let $\rho : B_0 \to B$ be the mapping as given in (2.2). One notes that ρ is Fréchet differentiable on B_0. Assume that f is a fixed twice H-differentiable function on B with the conditions that $Df(x) \in B^*$ for any $x \in B$ and $\mathbb{E}[\|D^2 f(Z)\|_{\mathrm{tr}(H)}] < +\infty$. Since $i = \rho \circ i_0$, we have for any $x \in B_0$ and $h, k \in H$, $\langle Df(\rho(x)), h\rangle = \langle D(f \circ \rho)(x), h\rangle$ and $\langle D^2 f(\rho(x))h, k\rangle = \langle D^2(f \circ \rho)(x)h, k\rangle$. Moreover,

$$\big|\langle D(f \circ \rho)(x), h\rangle\big| = \big|\langle Df\big(\rho(x)\big), h\rangle\big| \leqslant \big\|Df\big(\rho(x)\big)\big\|_{B^*} \|\rho(h)\| \leqslant \big\|Df\big(\rho(x)\big)\big\|_{B^*} \|h\|_0.$$

Regarding $Df(\rho(x))$ and $D(f\circ\rho)(x)$ as elements in B_0^*, they are the same and

$$\big\|Df\big(\rho(x)\big)\big\|_{B_0^*} = \big\|D(f\circ\rho)(x)\big\|_{B_0^*} \leqslant \big\|Df\big(\rho(x)\big)\big\|_{B^*}, \quad \forall x\in B_0. \tag{3.6}$$

In addition, we see by (2.2) that

$$\int_{B_0} \big\|D^2(f\circ\rho)(x)\big\|_{\mathrm{tr}(H)}\,\tilde p_1(dx) = \int_{B_0} \big\|D^2 f\big(\rho(x)\big)\big\|_{\mathrm{tr}(H)}\,\tilde p_1(dx) = \mathbb{E}\big[\big\|D^2 f(Z)\big\|_{\mathrm{tr}(H)}\big] < +\infty. \tag{3.7}$$

Next, it follows from [12, Remark 2.2] that B_0 can be expressed as the direct sum of $\tilde P_n(B_0)$ and $(I_0-\tilde P_n)(B_0)$, where I_0 is the identity map on B_0. Further, the triple $(i_0|_{K_n}, K_n, (I_0-\tilde P_n)(B_0))$ forms an AWS, where K_n is the orthogonal complement of $\tilde P_n(B_0)$ in H. Moreover, $\tilde p_1 = \tilde p_1^{(n)}\times\nu_n$, where $\tilde p_1^{(n)}$ is the abstract Wiener measure on $\tilde K_n \equiv (I_0-\tilde P_n)(B_0)$ with variance parameter 1 and ν_n is the standard Gauss measure on $\tilde P_n(B_0)$. Observe that

$$\begin{aligned}
&\int_{B_0} \mathrm{Tr}_H\big(\tilde P_n D^2(f\circ\rho)(x)\big)\,\tilde p_1(dx)\\
&= \int_{\tilde K_n}\int_{\tilde P_n(B_0)} \mathrm{Tr}_H\big(\tilde P_n D^2(f\circ\rho)(x_1+x_2)\big)\,\nu_n(dx_1)\,\tilde p_1^{(n)}(dx_2)\\
&= \sum_{j=1}^{k_n}\int_{\tilde K_n}\int_{\mathbb{R}^n}\left\langle D^2(f\circ\rho)\left(x_2+\sum_{i=1}^{k_n} t_i e_i\right)e_j, e_j\right\rangle \tilde\nu_n\big(d(t_1,\ldots,t_{k_n})\big)\,\tilde p_1^{(n)}(dx_2),
\end{aligned} \tag{3.8}$$

where $\tilde\nu_n$ is the standard multivariate normal distribution in $\mathbb{R}^{k_n}$. Fix $x_2\in\tilde K_n$, and set

$$\phi_{x_2}(\vec t\,) = (f\circ\rho)\left(x_2+\sum_{i=1}^{k_n} t_i e_i\right), \quad \forall \vec t = (t_1,\ldots,t_{k_n})\in\mathbb{R}^{k_n}.$$

Note that, for any $j=1,2,\ldots,k_n$,

$$\frac{\partial^2}{\partial t_j^2}\phi_{x_2}(\vec t\,) = \left\langle D^2(f\circ\rho)\left(x_2+\sum_{i=1}^{k_n} t_i e_i\right)e_j, e_j\right\rangle.$$

Therefore, it follows by applying Fubini's theorem to (3.8) that for any $j=1,2,\ldots,k_n$,

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left|\frac{\partial^2}{\partial t_j^2}\phi_{x_2}(\vec t\,)\right|\cdot e^{-\frac12 t_j^2}\,dt_j < +\infty,$$

for $\tilde{p}_1^{(n)}$-a.e. $x_2 \in \tilde{K}_n$ and a.e. $(t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_{k_n}) \in \mathbb{R}^{k_n-1}$ with respect to Lebesgue measure. By using Stein's lemma for univariate normal distributions (see [32]),

$$\begin{aligned}
(3.8) &= \sum_{j=1}^{k_n} \int_{\tilde{K}_n} \int_{\mathbb{R}^{k_n}} t_j \frac{\partial}{\partial t_j} \phi_{x_2}(\vec{t}\,)\, \tilde{\nu}_n(d\vec{t}\,)\, \tilde{p}_1^{(n)}(dx_2) \\
&= \sum_{j=1}^{k_n} \int_{\tilde{K}_n} \int_{\tilde{P}_n(B_0)} \langle x_1, e_j\rangle_0 \big\langle D(f\circ\rho)(x_1+x_2), e_j\big\rangle\, \nu_n(dx_1)\, \tilde{p}_1^{(n)}(dx_2) \\
&= \sum_{j=1}^{k_n} \int_{\tilde{K}_n} \int_{\tilde{P}_n(B_0)} (x_1+x_2, e_j)_{B_0,B_0^*} \big\langle D(f\circ\rho)(x_1+x_2), e_j\big\rangle\, \nu_n(dx_1)\, \tilde{p}_1^{(n)}(dx_2), \\
&\qquad \text{since } (x_2, e_j)_{B_0,B_0^*} = 0 \text{ for any } j = 1, 2, \ldots, k_n, \\
&= \sum_{j=1}^{k_n} \int_{B_0} (x, e_j)_{B_0,B_0^*} \big\langle D(f\circ\rho)(x), e_j\big\rangle\, \tilde{p}_1(dx) \\
&= \int_{B_0} \big(\tilde{P}_n(x), D(f\circ\rho)(x)\big)_{B_0,B_0^*}\, \tilde{p}_1(dx) \\
&= \int_{B_0} \big(\rho\big(\tilde{P}_n(x)\big), Df\big(\rho(x)\big)\big)_{B,B^*}\, \tilde{p}_1(dx). \qquad (3.9)
\end{aligned}$$

If B is finite-dimensional, then $H = B = B_0$ and there is a sufficiently large n such that $\tilde{P}_n = P_n = I_0 = \rho$. As a result of (2.2) and (3.9), $\mathbb{E}[|(Z, Df(Z))_{B,B^*}|] < +\infty$, and then f satisfies the identity (3.5) if $\mu_X = \mu_Z$.

If B is infinite-dimensional and f satisfies an extra-hypothesis that $\mathbb{E}[\|Df(Z)\|_{B^*}^{\alpha}]$ is finite for some $1 < \alpha < +\infty$, then, for any $x \in B_0$ and $n \in \mathbb{N}$, it follows from (3.6) that

$$\big|\big(\rho\big(\tilde{P}_n(x)\big), Df\big(\rho(x)\big)\big)_{B,B^*}\big| \leqslant \|x\|_0 \big\|Df\big(\rho(x)\big)\big\|_{B^*}.$$

By using the Hölder inequality, the Fernique theorem [27, Theorem 2.6], (2.2), and the above extra-hypothesis of f, we see that the mapping $x \in B_0 \mapsto \|x\|_0 \|Df(\rho(x))\|_{B^*}$ is integrable with respect to $\tilde{p}_1$. In addition,

$$\big|\mathrm{Tr}_H\big(\tilde{P}_n D^2(f\circ\rho)(x)\big)\big| \leqslant \big\|\tilde{P}_n D^2(f\circ\rho)(x)\big\|_{\mathrm{tr}(H)} \leqslant \big\|D^2 f\big(\rho(x)\big)\big\|_{\mathrm{tr}(H)}, \quad \forall x \in B_0,$$

where we see by (3.7) that the mapping $x \in B_0 \mapsto \|D^2 f(\rho(x))\|_{\mathrm{tr}(H)}$ is integrable with respect to $\tilde{p}_1$. One notes that $\tilde{P}_n(x) \to x$ and $\mathrm{Tr}_H(\tilde{P}_n D^2(f\circ\rho)(x)) \to \mathrm{Tr}_H(D^2 f(\rho(x)))$ as $n \to \infty$ for any $x \in B_0$. Applying the dominated convergence argument to both sides of (3.9) and then by (2.2), we see that

$$\begin{aligned}\mathbb{E}\big[\Delta_G f(Z)\big] &= \int_{B_0} \mathrm{Tr}_H\big(D^2 f\big(\rho(x)\big)\big)\,\tilde{p}_1(dx)\\ &= \lim_{n\to\infty}\int_{B_0} \mathrm{Tr}_H\big(\tilde{P}_n D^2(f\circ\rho)(x)\big)\,\tilde{p}_1(dx)\\ &= \lim_{n\to\infty}\int_{B_0} \big(\rho\big(\tilde{P}_n(x)\big), Df\big(\rho(x)\big)\big)_{B,B^*}\,\tilde{p}_1(dx)\\ &= \int_{B_0} \big(\rho(x), Df\big(\rho(x)\big)\big)_{B,B^*}\,\tilde{p}_1(dx)\\ &= \mathbb{E}\big[\big(Z, Df(Z)\big)_{B,B^*}\big].\end{aligned}$$

Therefore, if $\mu_X = \mu_Z$, the identity (3.5) is immediately obtained.

(*The "if" part for statements* (i) *and* (ii).) First, it is easy to see that for $f \equiv e^{\mathrm{i}(\cdot,\eta)}$ with $\eta \in B^*$, f is twice Fréchet differentiable on B, $Df(x) = \mathrm{i}\,e^{\mathrm{i}(x,\eta)}\eta \in B^*$, and $\|D^2 f(x)\|_{B,B^*} \leqslant \|\eta\|_{B^*}^2$ for any $x \in B$. Then, for any $1 \leqslant \alpha < +\infty$, $\mathbb{E}[\|Df(Z)\|_{B^*}^\alpha] = \|\eta\|^\alpha < +\infty$. And, by Theorem 2.6 we have

$$\mathbb{E}\big[\big\|D^2 f(Z)\big\|_{\mathrm{tr}(H)}\big] \leqslant \|\eta\|_{B^*}^2 \mathbb{E}\big[\|Z\|^2\big] < +\infty.$$

Thus, by the assumption, f satisfies the identity (3.5) and we obtain the following equality:

$$\mathbb{E}\big[(X,\eta)_{B,B^*} e^{\mathrm{i}(X,\eta)_{B,B^*}}\big] = \mathrm{i}|\eta|_0^2 \mathbb{E}\big[e^{\mathrm{i}(X,\eta)_{B,B^*}}\big], \quad \forall \eta \in B^*. \tag{3.10}$$

Now, for any $\eta \in B^*$, there associates a characteristic function h_η given by $h_\eta(r) = \mathbb{E}[e^{\mathrm{i}r(X,\eta)_{B,B^*}}]$, $r \in \mathbb{R}$. Since, for any $s \neq r \in \mathbb{R}$, by the mean value theorem for differentiation there are two real numbers p_{sr}, q_{sr} between s and r such that

$$e^{\mathrm{i}s(x,\eta)_{B,B^*}} - e^{\mathrm{i}r(x,\eta)_{B,B^*}} = \frac{\mathrm{i}(s-r)}{2}(x,\eta)_{B,B^*}\Psi_{s,r;\eta}(x), \quad x \in B,$$

where $\Psi_{s,r;\eta}(x) = e^{\mathrm{i}\,p_{sr}(x,\eta)_{B,B^*}} - e^{-\mathrm{i}\,p_{sr}(x,\eta)_{B,B^*}} + e^{\mathrm{i}\,q_{sr}(x,\eta)_{B,B^*}} + e^{-\mathrm{i}\,q_{sr}(x,\eta)_{B,B^*}}$, we see that

$$\frac{h_\eta(s) - h_\eta(r)}{s-r} = \frac{\mathrm{i}}{2}\mathbb{E}\big[(X,\eta)_{B,B^*}\Psi_{s,r;\eta}(X)\big], \tag{3.11}$$

which converges to $-r|\eta|_0^2 h_\eta(r)$ by applying the equality (3.10) to (3.11) and then letting s approach to r. So $h_\eta(r)$ satisfies the differential equation: $h'_\eta(r) = -r|\eta|_0^2 h_\eta(r)$ with the initial condition $h_\eta(0) = 1$. Therefore, $h_\eta(r) = e^{-\frac{1}{2}r^2|\eta|_0^2}$, $r \in \mathbb{R}$, from which we conclude that $\mu_X = \mu_Z$. The proof is complete. □

Remark 3.2.

(1) Applying the fact that $p_t(E) = \mu_Z(\frac{1}{\sqrt{t}}E)$ for any $E \in \mathcal{B}(B)$ to Theorem 3.1, we immediately obtain a characterization of p_t, in which the identity (3.5) becomes

$$\mathbb{E}\big[\langle X, Df(X)\rangle - t\Delta_G f(X)\big] = 0.$$

(2) Let $B = \mathbb{R}^n$ with the Euclidean inner product $\langle\cdot,\cdot\rangle$. Then μ_Z is a multivariate normal distribution on $\mathbb{R}^n$ with the covariance matrix $\mathbf{A} = [\mathbf{A}_{i,j}]$. In this case,

$$\Delta_G f(X) = \sum_{i,j=1}^{n} \mathbf{A}_{i,j} \frac{\partial^2}{\partial t_i \partial t_j} f(X),$$

and thus the identity (3.5) can be reformulated as

$$\mathbb{E}\left[\langle X, \nabla f(X)\rangle - \sum_{i,j=1}^{n} \mathbf{A}_{i,j} \frac{\partial^2}{\partial t_i \partial t_j} f(X)\right] = 0, \tag{3.12}$$

where $\nabla f(\vec{t})$ is the gradient of f at $\vec{t} \in \mathbb{R}^n$. See also [5,20,22], where the same Stein identity as in (3.12) was derived.

4. Stein's equation and its solutions for abstract Wiener measures

From (3.3), (3.4), and Theorem 3.1, the role of the Stein equation for the abstract Wiener measure μ_Z should be played by the following differential equation (with unknown functional f):

$$\Delta_G f(x) - \big(x, Df(x)\big)_{B,B^*} = h(x) - \mathbb{E}\big[h(Z)\big], \quad x \in B, \tag{4.1}$$

where h is given in some class of test functionals.

In the case $B = C[0,1]$, Barbour [2] showed that

$$\begin{aligned} f_h(x) &\equiv -\int_0^\infty \big(\mathcal{T}_t h(x) - \mathbb{E}\big[h(Z)\big]\big)\,dt \\ &= -\int_0^\infty \int_B \big(h\big(e^{-t}x + \sqrt{1-e^{-2t}}\,y\big) - \mathbb{E}\big[h(Z)\big]\big)\,\mu_Z(dy)\,dt \end{aligned} \tag{4.2}$$

solves Eq. (4.1), where h is given in the space L defined before Theorem 3.1. Recently, Chatterjee and Meckes [5] proved that Barbour's results also hold for standard multivariate normal distribution on $\mathbb{R}^k$, where, by (3.12), the associated Stein equation is

$$\Delta f(\vec{t}) - \langle \vec{t}, \nabla f(\vec{t})\rangle = h(\vec{t}) - \mathbb{E}\big[h(Z)\big], \quad \vec{t} \in \mathbb{R}^k,$$

in which $\Delta f(\vec{t})$ is the Laplacian of f at $\vec{t}$ and $h : \mathbb{R}^k \to \mathbb{R}$ is twice continuously differentiable having bounded first and second derivatives. See also [22]. In the following, we will show that

Eq. (4.1) can be solved by f_h whenever h is given in the Banach space $\mathcal{U}Lip\text{-}1(B)$ of those scalar-valued uniformly Lip-1 functions h on B with the norm $|||h||| = \|h\|_{\mathrm{ULip}} + |h(0)|$, where

$$\|h\|_{\mathrm{ULip}} \equiv \sup_{x \neq y \in B} \frac{|h(x) - h(y)|}{\|x - y\|} < +\infty.$$

Remark 4.1.

1. By the Fernique theorem, $\mathcal{U}Lip\text{-}1(B) \subset L^\alpha(B, p_t(x, dy))$ for any $1 \leqslant \alpha < +\infty$, $x \in B$, and $t > 0$, implying that $p_t f(x)$ is well defined and so is $\mathcal{T}_t f(x)$ for any $f \in \mathcal{U}Lip\text{-}1(B)$.
2. One notes that the right-hand double integral in (4.2) exists for any $x \in B$, because

$$\begin{aligned}
&\int_0^\infty \left| \int_B \bigl(h\bigl(e^{-t}x + \sqrt{1 - e^{-2t}}\,y\bigr) - \mathbb{E}\bigl[h(Z)\bigr]\bigr)\, \mu_Z(dy) \right| dt \\
&\quad \leqslant \int_0^\infty \int_B \bigl|h\bigl(e^{-t}x + \sqrt{1 - e^{-2t}}\,y\bigr) - h(y)\bigr|\, \mu_Z(dy)\, dt \\
&\quad \leqslant \|h\|_{\mathrm{ULip}} \int_0^\infty \int_B e^{-t} \bigl\|x + \bigl(\sqrt{e^{2t} - 1} - e^t\bigr) y\bigr\|\, \mu_Z(dy)\, dt < +\infty.
\end{aligned}$$

3. For any $t > 0$, $\mathcal{T}_t(\mathcal{U}Lip\text{-}1(B))$ is contained in $\mathcal{U}Lip\text{-}1(B)$ and, for any $f \in \mathcal{U}Lip\text{-}1(B)$, $\|\mathcal{T}_t f\|_{\mathrm{ULip}} \leqslant e^{-t} \|f\|_{\mathrm{ULip}}$.
4. In [25], Piech studied properties of the solution for the Cauchy problem:

$$\frac{\partial}{\partial t} u(x, t) = \bigl(x, Du(x, t)\bigr)_{B, B^*} - \Delta_G u(x, t) \quad (x \in B,\ t > 0),$$
$$\lim_{t \to 0} u(x, t) = f(x), \quad \text{uniformly for } x \in B.$$

   She assumed that f is bounded and uniformly Lip-1 on B, then proved that $u(x, t) = \mathcal{T}_t f(x)$ is the unique solution. So, if h is bounded and uniformly Lip-1 on B, it immediately follows from Piech's result that f_h in (4.2) is a solution of Eq. (4.1). We will justify in Theorem 4.8 that Piech's result is still valid without the hypothesis of boundedness on h.

In order to achieve our goals as mentioned in the beginning of this section, we need a series of lemmas concerning f_h and its H-derivatives.

Lemma 4.2. *Let $f \in \mathcal{U}Lip\text{-}1(B)$ and $x \in B$.*

(i) *$\mathcal{T}_t f(x)$ is infinitely H-differentiable, and $D(\mathcal{T}_t f)(x)$ determines an element $\lambda_{t,x}(f)$ in B^*, that is, for any $z \in H$,*

$$\bigl|\bigl(z, \lambda_{t,x}(f)\bigr)_{B, B^*}\bigr| = \bigl|\bigl\langle D(\mathcal{T}_t f)(x), z\bigr\rangle\bigr| \leqslant e^{-t} \|f\|_{\mathrm{ULip}} \|z\|; \tag{4.3}$$

(ii) *$\mathcal{T}_t f(x)$ is Gâteaux differentiable at x in the direction y with the Gâteaux derivative $\lambda_{t,x}(f)$ for any $y \in B$, that is,*

$$\big(y, \lambda_{t,x}(f)\big)_{B,B^*} = \lim_{s\to 0} s^{-1}\big(\mathcal{T}_t f(x+sy) - \mathcal{T}_t f(x)\big). \tag{4.4}$$

Proof. For the assertion (i), it is straightforward by Proposition 2.7 and the following observation: For any $z \in H$,

$$\begin{aligned}\big|\big\langle D(\mathcal{T}_t f)(x), z\big\rangle\big| &\leqslant \lim_{s\to 0} s^{-1}\int_B \big|f\big(y+e^{-t}x+se^{-t}z\big) - f\big(y+e^{-t}x\big)\big|\, p_{1-e^{-2t}}(dy) \\ &\leqslant e^{-t}\|f\|_{\mathrm{ULip}}\|z\|.\end{aligned}$$

To show the assertion (ii), fix $y \in B$ and take an approximating sequence $\{z_n\} \subset H$ such that $\|z_n - y\| \to 0$ as $n \to +\infty$. Then, for any $n \in \mathbb{N}$,

$$\begin{aligned}&\big|s^{-1}\big(\mathcal{T}_t f(x+sy) - \mathcal{T}_t f(x)\big) - \big(y, \lambda_{t,x}(f)\big)_{B,B^*}\big| \\ &\quad\leqslant \big|s^{-1}\big(\mathcal{T}_t f(x+sy) - \mathcal{T}_t f(x+sz_n)\big)\big| + \big|\big\langle z_n, \lambda_{t,x}(f)\big\rangle_0 - \big(y, \lambda_{t,x}(f)\big)_{B,B^*}\big| \\ &\qquad + \big|s^{-1}\big(\mathcal{T}_t f(x+sz_n) - \mathcal{T}_t f(x)\big) - \big\langle z_n, \lambda_{t,x}(f)\big\rangle_0\big| \\ &\quad\leqslant 2e^{-t}\|f\|_{\mathrm{ULip}}\|y-z_n\| + \big|s^{-1}\big(\mathcal{T}_t f(x+sz_n) - \mathcal{T}_t f(x)\big) - \big\langle D(\mathcal{T}_t f)(x), z_n\big\rangle\big|.\end{aligned}$$

Letting $s \to 0$ yields that for any $n \in \mathbb{N}$,

$$\limsup_{s\to 0}\big|s^{-1}\big(\mathcal{T}_t f(x+sy) - \mathcal{T}_t f(x)\big) - \big(y, \lambda_{t,x}(f)\big)_{B,B^*}\big| \leqslant 2e^{-t}\|f\|_{\mathrm{ULip}}\|y-z_n\|.$$

Then we get the desired equality (4.4) as n goes to infinity. □

As mentioned in Section 2, we still use the notation $D(\mathcal{T}_t f)(x)$ in place of $\lambda_{t,x}(f)$.

Remark 4.3. Let $f \in \mathcal{U}Lip\text{-}1(B)$.

1. In general, $\mathcal{T}_t f(x)$ is not necessarily Fréchet differentiable at x on B. See [9] and Refs. [1,8] cited therein.
2. For any $h \in H$,

$$\begin{aligned}\big\langle D(\mathcal{T}_t f)(x), h\big\rangle &= \lim_{s\to 0} s^{-1}\int_B f\big(e^{-t}x+y\big)\big(e^{\frac{e^{-t}s}{\sqrt{1-e^{-2t}}}\langle y,h\rangle_0 - \frac{1}{2}\frac{e^{-2t}s^2}{1-e^{-2t}}|h|_0^2} - 1\big)\, p_{1-e^{-2t}}(dy) \\ &= \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\int_B \langle y,h\rangle_0 f\big(e^{-t}x+\sqrt{1-e^{-2t}}\,y\big)\,\mu_Z(dy),\end{aligned} \tag{4.5}$$

from which it follows that for any $x_1, x_2 \in B$,

$$\left|D(\mathcal{T}_t f)(x_1) - D(\mathcal{T}_t f)(x_2)\right|_0 \leqslant \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\|f\|_{\mathrm{ULip}}\|x_1 - x_2\|. \tag{4.6}$$

Hence the mapping from B into H by $x \mapsto D(\mathcal{T}_t f)(x)$ is uniformly continuous.

3. By the same argument as in the proof of Lemma 4.2, $D(p_t f)(x)$ also defines an element in B^* for any $x \in B$ and $t > 0$, where, for any $y \in B$,

$$\big(y, D(p_t f)(x)\big)_{B,B^*} = \lim_{s\to 0} s^{-1}\big(p_t f(x+sy) - p_t f(x)\big),$$

implying $\|D(p_t f)(x)\|_{B^*} \leqslant \|f\|_{\mathrm{ULip}}$.

Lemma 4.4. *For $f \in \mathcal{U}Lip\text{-}1(B)$, $y \in B$, and $t > 0$, the mapping*

$$x \in B \mapsto \big(y, D(\mathcal{T}_t f)(x)\big)_{B,B^*}$$

is continuous.

Proof. Let $x_1, x_2 \in B$ and $k \in H$. By Lemma 4.2 and (4.5) we see that

$$\begin{aligned}
&\left|\big(y, D(\mathcal{T}_t f)(x_1) - D(\mathcal{T}_t f)(x_2)\big)_{B,B^*}\right| \\
&\quad\leqslant \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\|f\|_{\mathrm{ULip}}\|x_1 - x_2\| \int_B \left|\langle w, k\rangle_0\right| \mu_Z(dw) + 2e^{-t}\|f\|_{\mathrm{ULip}}\|y - k\| \\
&\quad\leqslant \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\|f\|_{\mathrm{ULip}}|k|_0\|x_1 - x_2\| + e^{-t}\|f\|_{\mathrm{ULip}}\|y - k\|,
\end{aligned}$$

where the last inequality is obtained by using the Cauchy–Schwarz inequality and the property that $\langle\cdot, k\rangle_0 \sim N(0, |k|_0^2)$ with respect to μ_Z. Then

$$\limsup_{x_1\to x_2}\left|\big(y, D(\mathcal{T}_t f)(x_1) - D(\mathcal{T}_t f)(x_2)\big)_{B,B^*}\right| \leqslant 2e^{-t}\|f\|_{\mathrm{ULip}}\|y - k\|, \quad \forall k \in H.$$

Since H is dense in B, this lemma immediately follows. □

Lemma 4.5. *Let $f \in \mathcal{U}Lip\text{-}1(B)$ and $x \in B$. Then $D^2(\mathcal{T}_t f)(x)$ is a trace-class operator on H. Moreover,*

$$\left\|D^2(\mathcal{T}_t f)(x)\right\|_{\mathrm{tr}(H)} \leqslant \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-2t}}}\|f\|_{\mathrm{ULip}} \int_B \|y\|\mu_Z(dy),$$

and

$$\left\|D^2(\mathcal{T}_t f)(x)\right\|_{\mathrm{HS}(H)} \leqslant \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-2t}}}\|f\|_{\mathrm{ULip}} \left\{\int_B \|y\|^2 \mu_Z(dy)\right\}^{\frac{1}{2}}.$$

Proof. For any $z_1, z_2 \in H$, we have by the semigroup property of $\{\mathcal{T}_t;\ t \geqslant 0\}$ that

$$
\begin{aligned}
&\big\langle D^2(\mathcal{T}_t f)(x)z_1, z_2\big\rangle \\
&\quad = \int_B \frac{\partial^2}{\partial r_1 \partial r_2}\bigg|_{r_1=r_2=0} \mathcal{T}_{\frac{t}{2}} f\bigg(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\bigg(y + \frac{e^{-\frac{t}{2}}r_2}{\sqrt{1-e^{-t}}}z_2\bigg) + r_1 e^{-\frac{t}{2}} z_1\bigg)\,\mu_Z(dy) \\
&\quad = \int_B \frac{\partial^2}{\partial r_1 \partial r_2}\bigg|_{r_1=r_2=0} \Big\{\mathcal{T}_{\frac{t}{2}} f\big(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\,y + r_1 e^{-\frac{t}{2}} z_1\big) e^{\frac{e^{-\frac{t}{2}}r_2}{\sqrt{1-e^{-t}}}\langle y, z_2\rangle_0 - \frac{1}{2}\frac{e^{-t}r_2^2}{1-e^{-t}}|z_2|_0^2}\Big\}\,\mu_Z(dy) \\
&\quad = \frac{e^{-t}}{\sqrt{1-e^{-t}}} \int_B \langle y, z_2\rangle_0 \big\langle D(\mathcal{T}_{\frac{t}{2}} f)\big(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\,y\big), z_1\big\rangle\,\mu_Z(dy).
\end{aligned}
\tag{4.7}
$$

For an arbitrarily given test operator T on B, we take an orthonormal basis $\{e_1, e_2, \ldots\} \subset B^*$ for H such that the range $T(B)$ of T is spanned by $\{e_1, e_2, \ldots, e_n\}$. Since, for any $x \in B$, $T|_H \circ D^2(\mathcal{T}_t f)(x)$ is of finite rank, it is a trace-class operator on H. Then we see by (4.7) and Remark 2.5(1) that for any $x \in B$,

$$
\begin{aligned}
&\mathrm{Tr}_H\big(T|_H \circ D^2(\mathcal{T}_t f)(x)\big) \\
&\quad = \sum_{j=1}^{\infty} \big\langle D^2(\mathcal{T}_t f)(x)e_j, (T|_H)^*(e_j)\big\rangle \\
&\quad = \frac{e^{-t}}{\sqrt{1-e^{-t}}} \sum_{j=1}^{\infty} \int_B \big\langle y, (T|_H)^*(e_j)\big\rangle_0 \big\langle D(\mathcal{T}_{\frac{t}{2}} f)\big(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\,y\big), e_j\big\rangle\,\mu_Z(dy) \\
&\quad = \frac{e^{-t}}{\sqrt{1-e^{-t}}} \sum_{j=1}^{n} \int_B (Ty, e_j)_{B,B^*} \big\langle D(\mathcal{T}_{\frac{t}{2}} f)\big(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\,y\big), e_j\big\rangle\,\mu_Z(dy) \\
&\quad = \frac{e^{-t}}{\sqrt{1-e^{-t}}} \int_B \big\langle D(\mathcal{T}_{\frac{t}{2}} f)\big(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\,y\big), T(y)\big\rangle\,\mu_Z(dy).
\end{aligned}
\tag{4.8}
$$

Also, by Proposition 2.4(ii) and Lemma 4.2,

$$
\begin{aligned}
&\int_B \Big|\big\langle D(\mathcal{T}_{\frac{t}{2}} f)\big(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\,y\big), T(y)\big\rangle\Big|\,\mu_Z(dy) \\
&\qquad \leqslant e^{-\frac{t}{2}}\|f\|_{\mathrm{ULip}}\|T|_H\|_{H,H} \int_B \|y\|\,\mu_Z(dy).
\end{aligned}
$$

Hence we have

$$
\Big|\mathrm{Tr}_H\big(T|_H \circ D^2(\mathcal{T}_t f)(x)\big)\Big| \leqslant \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}}\|f\|_{\mathrm{ULip}}\|T|_H\|_{H,H} \int_B \|y\|\,\mu_Z(dy).
$$

Consequently, by Proposition 2.4(i), $D^2(\mathcal{T}_t f)(x)$ determines an element of the dual space $\mathcal{K}(H)^*$ of $\mathcal{K}(H)$. Recall the fact that $\mathcal{K}(H)^*$ is isometrically isomorphic to the Banach space $\mathcal{TR}(H)$, where the pairing between $\mathcal{K}(H)$ and $\mathcal{TR}(H)$ is given by $(A, F) = \mathrm{Tr}_H(AF)$ (see [28]). Therefore, we have shown that $D^2(\mathcal{T}_t f)(x)$ belongs to $\mathcal{TR}(H)$, the trace-class norm of which satisfies the desired inequality.

On the other hand, by (4.8), Lemma 4.2, and Corollary 2.11 we see that for any $x \in B$,

$$
\begin{aligned}
&\bigl|\mathrm{Tr}_H\bigl(T|_H \circ D^2(\mathcal{T}_t f)(x)\bigr)\bigr| \\
&\quad \leqslant \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}}\|f\|_{\mathrm{ULip}} \int_B \bigl\|T(y)\bigr\|\,\mu_Z(dy) \\
&\quad \leqslant \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}}\|f\|_{\mathrm{ULip}} \biggl\{\int_B \|y\|^2\,\mu_Z(dy)\biggr\}^{\frac{1}{2}} \biggl\{\int_B \bigl|T(y)\bigr|_0^2\,\mu_Z(dy)\biggr\}^{\frac{1}{2}},
\end{aligned}
$$

where

$$
\begin{aligned}
\int_B \bigl|T(y)\bigr|_0^2\,\mu_Z(dy) &= \sum_{j=1}^n \int_B \bigl\langle T(y), e_j\bigr\rangle_0^2\,\mu_Z(dy) \\
&= \sum_{j=1}^n \int_B \bigl(y, T^*(e_j)\bigr)_{B,B^*}^2\,\mu_Z(dy) \\
&= \sum_{j=1}^n \bigl|T^*(e_j)\bigr|_0^2 \\
&= \sum_{j=1}^n \bigl|(T|_H)^*(e_j)\bigr|_0^2 \quad \bigl(\text{by Remark 2.5(1)}\bigr) \\
&= \bigl\|(T|_H)^*\bigr\|_{\mathrm{HS}(H)}^2.
\end{aligned}
$$

Consequently, we conclude by Remark 2.5(2) and the above estimation that $D^2(\mathcal{T}_t f)(x)$ determines an element of $\mathcal{HS}(H)$, the Hilbert–Schmidt norm of which satisfies the desired inequality. The proof is complete. □

Remark 4.6.

1. Let $f \in \mathcal{U}Lip\text{-}1(B)$. By the semigroup property of $\{\mathcal{T}_t\}$, (4.6), and (4.8), we see that for any test operator T on B and $x_1, x_2 \in B$,

$$
\begin{aligned}
&\bigl|\mathrm{Tr}_H\bigl(T|_H \circ \bigl(D^2(\mathcal{T}_t f)(x_1) - D^2(\mathcal{T}_t f)(x_2)\bigr)\bigr)\bigr| \\
&\quad \leqslant \frac{e^{-t}}{\sqrt{1-e^{-t}}} \int_B \bigl|\bigl\langle D(\mathcal{T}_{\frac{t}{2}} f)\bigl(e^{-\frac{t}{2}}x_1 + \sqrt{1-e^{-t}}\,y\bigr)
\end{aligned}
$$

$$- D(\mathcal{T}_{\frac{t}{2}} f)\big(e^{-\frac{t}{2}}x_2 + \sqrt{1-e^{-t}}\,y\big), T(y)\rangle\big| \,\mu_Z(dy)$$
$$\leqslant \frac{e^{-\frac{5}{2}t}}{1-e^{-t}} \|f\|_{\mathrm{ULip}} \|x_1 - x_2\| \int_B \big|T(y)\big|_0 \,\mu_Z(dy)$$
$$\leqslant \frac{e^{-\frac{5}{2}t}}{1-e^{-t}} \|f\|_{\mathrm{ULip}} \|x_1 - x_2\| \big\|(T|_H)^*\big\|_{\mathrm{HS}(H)}.$$

Consequently, for any $x_1, x_2 \in B$,

$$\big\|D^2(\mathcal{T}_t f)(x_1) - D^2(\mathcal{T}_t f)(x_2)\big\|_{\mathrm{HS}(H)} \leqslant \frac{e^{-\frac{5}{2}t}}{1-e^{-t}} \|f\|_{\mathrm{ULip}} \|x_1 - x_2\|. \tag{4.9}$$

2. Let $f \in \mathcal{U}Lip\text{-}1(B)$. For any $n \in \mathbb{N}$, $z_1, \dots, z_{n+1} \in H$, and $x \in B$,

$$D^{n+1}(\mathcal{T}_t f)(x)(z_1, \dots, z_{n+1}) = \frac{d}{dr_1}\bigg|_{r_1=0} D^n(\mathcal{T}_t f)(x + r_1 z_1)(z_2, \dots, z_{n+1}).$$

Then, by using the formula in [19, Corollary 3.4(b)], we can establish a generalization of (4.7) for $D^{n+1}(\mathcal{T}_t f)$ as follows:

$$D^{n+1}(\mathcal{T}_t f)(x)(z_1, \dots, z_{n+1}) = \frac{e^{-\frac{n+1}{2}t}}{(1-e^{-t})^{\frac{n}{2}}} \int_B \left\{ \int_B \prod_{j=2}^{n+1} \langle y + \mathrm{i}\,w, z_j\rangle_0 \,\mu_Z(dw) \right\}$$
$$\times \big\langle D(\mathcal{T}_{\frac{t}{2}} f)\big(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\,y\big), z_1\big\rangle \mu_Z(dy). \tag{4.10}$$

Lemma 4.7. *For $f \in \mathcal{U}Lip\text{-}1(B)$ and $t > 0$, the scalar-valued mapping*

$$x \in B \mapsto \mathrm{Tr}_H\big(D^2(\mathcal{T}_t f)(x)\big)$$

is continuous.

Proof. Let (i_0, H, B_0) and $\{P_n\}$ be given as those in Lemma 2.2, $\tilde{P}_n$'s be the extension of P_n's to B_0, $\tilde{p}_1$ be the abstract Wiener measure on B_0 with variance parameter 1, and $\rho: B_0 \to B$ be the mapping as in (2.2). Observe that, for any $x_1, x_2 \in B$ and $n \in \mathbb{N}$,

$$\big|\mathrm{Tr}_H\big(D^2(\mathcal{T}_t f)(x_1)\big) - \mathrm{Tr}_H\big(D^2(\mathcal{T}_t f)(x_2)\big)\big|$$
$$\leqslant \big|\mathrm{Tr}_H\big(P_n \circ D^2(\mathcal{T}_t f)(x_1) - D^2(\mathcal{T}_t f)(x_1)\big)\big|$$
$$+ \big|\mathrm{Tr}_H\big(P_n \circ D^2(\mathcal{T}_t f)(x_2) - D^2(\mathcal{T}_t f)(x_2)\big)\big|$$
$$+ \big|\mathrm{Tr}_H\big(P_n \circ \big[D^2(\mathcal{T}_t f)(x_1) - D^2(\mathcal{T}_t f)(x_2)\big]\big)\big|. \tag{4.11}$$

For any $z_1, z_2 \in H$, it follows from (2.2) and (4.7) that

$$\begin{aligned}
&\left\langle D^2(\mathcal{T}_t f)(x) z_1, z_2\right\rangle \\
&= \int_{B_0} \frac{\partial^2}{\partial r_1 \partial r_2}\bigg|_{r_1=r_2=0} \mathcal{T}_{\frac{t}{2}} f\left(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\rho\left(y + \frac{e^{-\frac{t}{2}} r_2}{\sqrt{1-e^{-t}}} z_2\right) + r_1 e^{-\frac{t}{2}} z_1\right) \tilde{p}_1(dy) \\
&= \frac{e^{-t}}{\sqrt{1-e^{-t}}} \int_{B_0} \langle y, z_2\rangle_0 \left\langle D(\mathcal{T}_{\frac{t}{2}} f)\left(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\rho(y)\right), z_1\right\rangle \tilde{p}_1(dy). \qquad (4.12)
\end{aligned}$$

For any test operator T on B_0, $T \circ \tilde{P}_n$ is also a test operator on B_0. By (4.12) and a similar argument to (4.8) we see that

$$\begin{aligned}
&\mathrm{Tr}_H\left(T|_H \circ P_n \circ D^2(\mathcal{T}_t f)(x)\right) \quad \left((T \circ \tilde{P}_n)|_H = T|_H \circ P_n\right) \\
&\quad = \frac{e^{-t}}{\sqrt{1-e^{-t}}} \int_{B_0} \left\langle D(\mathcal{T}_{\frac{t}{2}} f)\left(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\rho(y)\right), T \circ \tilde{P}_n(y)\right\rangle \tilde{p}_1(dy).
\end{aligned}$$

By using this equality together with Proposition 2.4(ii) and Lemma 4.2 it follows that for any $x \in B$,

$$\begin{aligned}
&\left|\mathrm{Tr}_H\left(T|_H \circ \left[P_n \circ D^2(\mathcal{T}_t f)(x) - D^2(\mathcal{T}_t f)(x)\right]\right)\right| \\
&\quad \leqslant \frac{e^{-t}}{\sqrt{1-e^{-t}}} \int_{B_0} \left|\left\langle D(\mathcal{T}_{\frac{t}{2}} f)\left(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}\rho(y)\right), T\left(\tilde{P}_n(y) - y\right)\right\rangle\right| \tilde{p}_1(dy) \\
&\quad \leqslant \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}} \|f\|_{\mathrm{ULip}} \|T|_H\|_{H,H} \int_{B_0} \left\|\tilde{P}_n(y) - y\right\|_0 \tilde{p}_1(dy).
\end{aligned}$$

Arguing as in the proof of Lemma 4.5 we see that

$$\begin{aligned}
&\left\|P_n \circ D^2(\mathcal{T}_t f)(x) - D^2(\mathcal{T}_t f)(x)\right\|_{\mathrm{tr}(H)} \\
&\quad \leqslant \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}} \|f\|_{\mathrm{ULip}} \int_{B_0} \left\|\tilde{P}_n(y) - y\right\|_0 \tilde{p}_1(dy) \to 0 \quad \text{as } n \to \infty, \text{ uniformly for } x \in B. \qquad (4.13)
\end{aligned}$$

For any $n \in \mathbb{N}$ and $t > 0$, we have by using the Cauchy–Schwarz inequality that

$$\begin{aligned}
&\left|\mathrm{Tr}_H\left(P_n \circ \left[D^2(\mathcal{T}_t f)(x_1) - D^2(\mathcal{T}_t f)(x_2)\right]\right)\right| \\
&\quad \leqslant \sqrt{\dim\left(P_n(H)\right)} \left\|D^2(\mathcal{T}_t f)(x_1) - D^2(\mathcal{T}_t f)(x_2)\right\|_{\mathrm{HS}(H)}, \qquad (4.14)
\end{aligned}$$

which tends to 0 as $x_1 \to x_2$ by (4.9). Finally, by applying (4.13) and (4.14) to (4.11), this lemma immediately follows. □

Now, we are in a position to prove the following Theorem 4.8 to Theorem 4.10, which are the main goals of this section. Let $f \in \mathcal{U}Lip\text{-}1(B)$ be fixed, and set $\alpha_t(s) = 1 - e^{-2(t+s)}$ and $\beta_t(s) = e^{-(t+s)}$, $t, s > 0$. Then, for any $x \in B$,

$$\begin{aligned}\frac{d}{dt}\mathcal{T}_t f(x) &= \lim_{s\to 0} s^{-1}\big\{p_{\alpha_t(s)} f\big(\beta_t(s)x\big) - p_{\alpha_t(0)} f\big(\beta_t(0)x\big)\big\}\\ &= \lim_{s\to 0} s^{-1}\big\{p_{\alpha_t(s)} f\big(\beta_t(s)x\big) - p_{\alpha_t(0)} f\big(\beta_t(s)x\big)\big\}\\ &\quad + \lim_{s\to 0} s^{-1}\big\{p_{\alpha_t(0)} f\big(\beta_t(s)x\big) - p_{\alpha_t(0)} f\big(\beta_t(0)x\big)\big\}\\ &\equiv \lim_{s\to 0}(I) + \lim_{s\to 0}(II).\end{aligned} \tag{4.15}$$

It is easy to see that

$$\begin{aligned}\lim_{s\to 0}(II) &= \lim_{s\to 0} s^{-1}\big\{p_{\alpha_t(0)} f\big(\beta_t(s)x\big) - p_{\alpha_t(0)} f\big(\beta_t(0)x\big)\big\}\\ &= \frac{d}{ds}\bigg|_{s=0} \mathcal{T}_t f\big(e^{-s}x\big)\\ &= -\big(x, D(\mathcal{T}_t f)(x)\big)_{B,B^*}.\end{aligned} \tag{4.16}$$

To evaluate $\lim_{s\to 0}(I)$, we observe that

$$\begin{aligned}\lim_{s\to 0}(I) &= \lim_{s\to 0} s^{-1}\big\{p_{\alpha_t(s)} f\big(\beta_t(s)x\big) - p_{\alpha_t(0)} f\big(\beta_t(s)x\big)\big\}\\ &= \frac{d}{ds}\bigg|_{s=0} p_{\alpha_t(s)} f\big(\beta_t(s)x\big) - \frac{d}{ds}\bigg|_{s=0} p_{\alpha_t(0)} f\big(\beta_t(s)x\big)\\ &= \frac{d}{du}\bigg|_{u=\alpha_t(0)} p_u f\big(\beta_t(0)x\big)\cdot \alpha_t'(0)\\ &= 2e^{-2t} \lim_{s\to 0} s^{-1}\big(p_{\alpha_t(0)+s} f\big(e^{-t}x\big) - p_{\alpha_t(0)} f\big(e^{-t}x\big)\big),\end{aligned} \tag{4.17}$$

where, for any $x \in B$ and $s, t > 0$,

$$\begin{aligned}&p_{\alpha_t(0)+s} f\big(e^{-t}x\big) - p_{\alpha_t(0)} f\big(e^{-t}x\big)\\ &\quad = \int_B \big(p_{\alpha_t(0)} f\big(e^{-t}x + y\big) - p_{\alpha_t(0)} f\big(e^{-t}x\big)\big)\, p_s(dy)\\ &\quad = \int_B \int_0^1 \big(y, D(p_{\alpha_t(0)} f)\big(e^{-t}x + \vartheta y\big)\big)_{B,B^*}\, d\vartheta\, p_s(dy) \quad \text{(by Remark 4.3)}\\ &\quad = \int_0^1 \int_B \big(y, D(p_{\alpha_t(0)} f)\big(e^{-t}x + \vartheta y\big)\big)_{B,B^*}\, p_s(dy)\, d\vartheta.\end{aligned} \tag{4.18}$$

The last equality in (4.18) is obtained by applying Fubini's theorem, since

$$\int_B \int_0^1 \left|\left(y, D(p_{\alpha_t(0)} f)\left(e^{-t}x + \vartheta y\right)\right)_{B,B^*}\right| d\vartheta \; p_s(dy) \leqslant \|f\|_{\mathrm{ULip}} \int_B \|y\| \, p_s(dy) < +\infty,$$

by Remark 4.3. Notice that for any $x, y \in B$, $t > 0$, and $\vartheta \in (0, 1]$,

$$\begin{aligned}\left(y, D(p_{\alpha_t(0)} f)\left(e^{-t}x + \vartheta y\right)\right)_{B,B^*} &= e^t \left(y, D(\mathcal{T}_t f)\left(x + \vartheta e^t y\right)\right)_{B,B^*} \\ &= \vartheta^{-1}\left(y, D\phi_{x,\vartheta,t}(y)\right)_{B,B^*},\end{aligned} \tag{4.19}$$

where $\phi_{x,\vartheta,t}(y) \equiv \mathcal{T}_t f(x + \vartheta e^t y)$. By Lemma 4.2, $\|D\phi_{x,\vartheta,t}(y)\|_{B^*} \leqslant \vartheta \|f\|_{\mathrm{ULip}}$ for any $y \in B$, and by Lemma 4.5,

$$\left\|D^2\phi_{x,\vartheta,t}(y)\right\|_{\mathrm{tr}(H)} \leqslant \frac{\vartheta^2}{\sqrt{1 - e^{-2t}}} \|f\|_{\mathrm{ULip}} \int_B \|u\| \, \mu_Z(du) \quad \text{for any } y \in B.$$

So $\|D\phi_{x,\vartheta,t}(y)\|_{B^*}$ and $\|D^2\phi_{x,\vartheta,t}(y)\|_{\mathrm{tr}(H)}$ are bounded functions considered as functions of y. As a result of Theorem 3.1 and Remark 3.2(1),

$$\begin{aligned}\int_B \left(y, D\phi_{x,\vartheta,t}(y)\right)_{B,B^*} p_s(dy) &= s \int_B \mathrm{Tr}_H\left(D^2\phi_{x,\vartheta,t}(y)\right) p_s(dy) \\ &= s\vartheta^2 e^{2t} \int_B \mathrm{Tr}_H\left(D^2(\mathcal{T}_t f)\left(x + \vartheta e^t y\right)\right) p_s(dy),\end{aligned}$$

from which and (4.19) it follows that

$$(4.18) = s e^{2t} \int_0^1 \int_B \vartheta \, \mathrm{Tr}_H\left(D^2(\mathcal{T}_t f)\left(x + \vartheta e^t y\right)\right) p_s(dy) \, d\vartheta. \tag{4.20}$$

Combining (4.20) with (4.17), we see that

$$\begin{aligned}\lim_{s\to 0}(I) &= 2 \lim_{s\to 0} \int_0^1 \int_B \vartheta \, \mathrm{Tr}_H\left(D^2(\mathcal{T}_t f)\left(x + \vartheta e^t y\right)\right) p_s(dy) \, d\vartheta \\ &= 2 \lim_{s\to 0} \int_0^1 \int_B \vartheta \, \mathrm{Tr}_H\left(D^2(\mathcal{T}_t f)\left(x + \vartheta e^t \sqrt{s} y\right)\right) \mu_Z(dy) \, d\vartheta.\end{aligned} \tag{4.21}$$

By Lemma 4.7,

$$\mathrm{Tr}_H\left(D^2(\mathcal{T}_t f)\left(x + \vartheta e^t \sqrt{s} y\right)\right) \to \mathrm{Tr}_H\left(D^2(\mathcal{T}_t f)(x)\right) \quad \text{as } s \to 0,$$

and by Lemma 4.5,

$$\left|\vartheta\, \mathrm{Tr}_H\big(D^2(\mathcal{T}_t f)\big(x+\vartheta e^t\sqrt{s}y\big)\big)\right| \leqslant \frac{e^{-2t}\vartheta}{\sqrt{1-e^{-2t}}}\|f\|_{\mathrm{ULip}} \int_B \|w\|\,\mu_Z(dw),$$

where the right-hand term is integrable considered as a function of $(\vartheta, y) \in [0,1] \times B$ with respect to the product measure $d\vartheta \times \mu_Z$. Hence we can apply the dominated convergence argument to (4.21) and then obtain that

$$\lim_{s\to 0}(I) = 2\int_0^1\int_B \vartheta\, \mathrm{Tr}_H\big(D^2(\mathcal{T}_t f)(x)\big)\,\mu_Z(dy)\,d\vartheta = \Delta_G(\mathcal{T}_t f)(x). \tag{4.22}$$

Theorem 4.8. *Let $f \in \mathcal{U}Lip\text{-}1(B)$. Then, for any $x \in B$ and $t > 0$,*

$$\Delta_G(\mathcal{T}_t f)(x) - \big(x, D(\mathcal{T}_t f)(x)\big)_{B,B^*} = \frac{d}{dt}\mathcal{T}_t f(x). \tag{4.23}$$

Moreover, for any $x \in B$,

$$\int_0^\infty \Delta_G(\mathcal{T}_t f)(x)\,dt - \int_0^\infty \big(x, D(\mathcal{T}_t f)(x)\big)_{B,B^*}\,dt = \mathbb{E}\big[f(Z)\big] - f(x). \tag{4.24}$$

Proof. The identity (4.23) is just the combination of (4.15), (4.16) and (4.22). So we only need to verify (4.24). One notes that those integrals in (4.24) exist by Lemmas 4.2 and 4.5. Then (4.24) is easily obtained by integrating both sides of (4.23) as functions of t from zero to infinity. □

In the rest of this section, for any $h \in \mathcal{U}Lip\text{-}1(B)$, let f_h be given as one in (4.2). It is obvious that $f_h \in \mathcal{U}Lip\text{-}1(B)$ and $\|f_h\|_{\mathrm{ULip}} \leqslant \|h\|_{\mathrm{ULip}}$.

Theorem 4.9. *Let $h \in \mathcal{U}Lip\text{-}1(B)$. Then we have the following results.*

(i) *f_h is twice H-differentiable at any $x \in B$. Further,*

$$Df_h(x) = -\int_0^\infty D(\mathcal{T}_t h)(x)\,dt \tag{4.25}$$

as a B^*-valued Bochner integral, as well as

$$D^2 f_h(x) = -\int_0^\infty D^2(\mathcal{T}_t h)(x)\,dt \tag{4.26}$$

as an $L(H,H)$-valued Bochner integral, where $L(H,H)$ is the Banach space of all bounded linear operators from H into itself with the operator norm $\|\cdot\|_{H,H}$.

(ii) $\|Df_h(x)\|_{B^*} \leqslant \|h\|_{\mathrm{ULip}}$ *and* $\|D^2 f_h(x)\|_{H,H} \leqslant \frac{\pi}{2}\|h\|_{\mathrm{ULip}}\{\int_B \|y\|^2\,\mu_Z(dy)\}^{\frac{1}{2}}$*, uniformly for* $x \in B$.

(iii) *For any* $x \in B$*,* $D^2 f_h(x)$ *is a trace-class operator on* H*, and*

$$\Delta_G f_h(x) = -\int_0^\infty \Delta_G(\mathcal{T}_t h)(x)\,dt. \tag{4.27}$$

(iv) $\|D^2 f_h(x)\|_{\mathrm{tr}(H)} \leqslant \frac{\pi}{2}\|h\|_{\mathrm{ULip}}\int_B \|y\|\,\mu_Z(dy)$*, uniformly for* $x \in B$.

(v) $\|D^2 f_h(x)\|_{\mathrm{HS}(H)} \leqslant \frac{\pi}{2}\|h\|_{\mathrm{ULip}}\{\int_B \|y\|^2\,\mu_Z(dy)\}^{\frac{1}{2}}$*, uniformly for* $x \in B$.

Proof. First of all, it follows from Lemma 4.2 and (4.7) that for any $x \in B$

$$\int_0^\infty \big\|D(\mathcal{T}_t h)(x)\big\|_{B^*}\,dt \leqslant \|h\|_{\mathrm{ULip}}.$$

And, by (4.7), Corollary 2.11, and Lemma 4.2 we have

$$\begin{aligned}
\int_0^\infty \big\|D^2(\mathcal{T}_t h)(x)\big\|_{H,H}\,dt &= \int_0^\infty \sup_{z_1,z_2\in H\setminus\{0\}} \frac{|\langle D^2(\mathcal{T}_t h)(x)z_1, z_2\rangle|}{|z_1|_0|z_2|_0}\,dt \\
&\leqslant \|h\|_{\mathrm{ULip}}\left\{\int_B \|y\|^2\,\mu_Z(dy)\right\}^{\frac{1}{2}} \int_0^\infty \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}}\,dt \\
&\quad\times \sup_{z_2\in H\setminus\{0\}}\left\{|z_2|_0^{-1}\int_B \big|\langle y, z_2\rangle_0\big|\,\mu_Z(dy)\right\} \\
&\leqslant \frac{\pi}{2}\|h\|_{\mathrm{ULip}}\left\{\int_B \|y\|^2\,\mu_Z(dy)\right\}^{\frac{1}{2}} < +\infty.
\end{aligned}$$

To show H-differentiability of f_h at any $x \in B$, we estimate

$$\begin{aligned}
&|z|_0^{-1}\left|f_h(x+z) - f_h(x) + \int_0^\infty \big\langle D(\mathcal{T}_t h)(x), z\big\rangle\,dt\right| \\
&\leqslant |z|_0^{-1}\left\{\int_0^\infty\int_0^1 \big|\big\langle D(\mathcal{T}_t h)(x+\vartheta z) - D(\mathcal{T}_t h)(x), z\big\rangle\big|\,d\vartheta\,dt\right\} \\
&\leqslant \|h\|_{\mathrm{ULip}}\|z\|\int_0^\infty\int_0^1 \frac{e^{-2t}\vartheta}{\sqrt{1-e^{-2t}}}\,d\vartheta\,dt \quad \big(\text{by (4.6)}\big)
\end{aligned}$$

$$= \frac{1}{2}\|h\|_{\mathrm{ULip}}\|z\| \to 0, \quad \text{as } |z|_0 \to 0.$$

Therefore, f_h is H-differentiable and we have shown (4.25). Next, for the second-order H-differentiability of f_h at $x \in B$, we need to prove

$$\lim_{|z_1|_0 \to 0} |z_1|_0^{-1} \left| Df_h(x+z_1) - Df_h(x) + \int_0^\infty \langle D^2(\mathcal{T}_t h)(x)z_1, \cdot\rangle\, dt \right|_0 = 0. \tag{4.28}$$

Making use of (4.5), we observe that

$$\begin{aligned}
&|z_1|_0^{-1} \left| Df_h(x+z_1) - Df_h(x) + \int_0^\infty \langle D^2(\mathcal{T}_t h)(x)z_1, \cdot\rangle\, dt \right|_0 \\
&\quad \leqslant |z_1|_0^{-1} \int_0^\infty \left| D(\mathcal{T}_t h)(x+z_1) - D(\mathcal{T}_t h)(x) + \langle D^2(\mathcal{T}_t h)(x)z_1, \cdot\rangle \right|_0 dt \\
&\quad = \int_0^\infty \Phi_{z_1}(t)\, dt,
\end{aligned}$$

where

$$\begin{aligned}
\Phi_{z_1}(t) = |z_1|_0^{-1} \sup_{z_2 \in H\setminus\{0\}} \Bigg\{ |z_2|_0^{-1} \Bigg| \int_0^1 \big[&\langle D^2(\mathcal{T}_t h)(x+\vartheta z_1)z_1, z_2\rangle \\
&- \langle D^2(\mathcal{T}_t h)(x)z_1, z_2\rangle \big]\, d\vartheta \Bigg| \Bigg\}\, dt.
\end{aligned}$$

By the infinite H-differentiability of $\mathcal{T}_t h(x)$ at every $x \in B$ we see that

$$\begin{aligned}
0 &\leqslant \Phi_{z_1}(t) \\
&\leqslant \int_0^1 \left\| D^2(\mathcal{T}_t h)(x+\vartheta z_1) - D^2(\mathcal{T}_t h)(x) \right\|_{H,H} d\vartheta \to 0, \quad \text{as } |z_1|_0 \to 0, \text{ for any } t > 0.
\end{aligned}$$

Moreover, by (4.7), Corollary 2.11, and Lemma 4.2,

$$\begin{aligned}
\Phi_{z_1}(t) &\leqslant \text{Const.} \sup_{z_2 \in H\setminus\{0\}} \left\{ \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}} \|h\|_{\mathrm{ULip}} |z_2|_0^{-1} \int_B \left| \langle y, z_2\rangle_0 \right| \mu_Z(dy) \right\} \\
&\leqslant \text{Const.} \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}} \in L^1\big([0,+\infty), dt\big), \quad \text{uniformly for } z_1 \in H.
\end{aligned}$$

Hence the formula (4.28) is an immediate result by the dominated convergence theorem, which implies (4.26). We have shown the assertions (i) and (ii). To show the assertion (iv), let T be an arbitrarily chosen test operator on B. Then, by (i) and a similar argument to the proof of Lemma 4.5, we see that for any $x \in B$,

$$\begin{aligned}\left|\mathrm{Tr}_H\left(T|_H \circ D^2 f_h(x)\right)\right| &\leqslant \int_0^\infty \left|\mathrm{Tr}_H\left(T|_H \circ D^2(\mathcal{T}_t h)(x)\right)\right| dt \\ &\leqslant \|h\|_{\mathrm{ULip}} \|T|_H\|_{H,H} \int_0^\infty \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}}\, dt \int_B \|y\|\, \mu_Z(dy),\end{aligned}$$

implying that $D^2 f_h(x)$ is trace-class whose trace-class norm has the desired bound as given in (iv). Imitating the proof of (iv), we thus verify the assertion (v). In order to prove the assertion (iii), we take an orthonormal basis $\{e_n;\ n = 1, 2, \ldots\}$ for H, consisting of all eigenvectors of $|D^2(\mathcal{T}_t h)(x)|$ for some $x \in B$. Then, by (4.26),

$$\Delta_G f_h(x) = \sum_{n=1}^\infty \left\langle D^2 f_h(x) e_n, e_n \right\rangle = -\lim_{m\to\infty} \int_0^\infty \sum_{n=1}^m \left\langle D^2(\mathcal{T}_t h)(x) e_n, e_n \right\rangle dt.$$

Also, by Lemma 4.5,

$$\left|\sum_{n=1}^m \left\langle D^2(\mathcal{T}_t h)(x) e_n, e_n \right\rangle\right| \leqslant \left\| D^2(\mathcal{T}_t h)(x)\right\|_{\mathrm{tr}(H)} \leqslant \mathrm{Const.} \frac{e^{-\frac{3t}{2}}}{\sqrt{1-e^{-t}}},$$

uniformly for $m \in \mathbb{N}$, where the last term is in $L^1([0, +\infty), dt)$. Therefore (4.27) is immediately obtained by the dominated convergence theorem. □

Finally, we combine Theorem 4.9 with Theorem 4.8 to get the following

Theorem 4.10. *For any $h \in \mathcal{U}Lip\text{-}1(B)$, $f_h(x)$ solves the equation*

$$\Delta_G f(x) - \left(x, Df(x)\right)_{B,B^*} = h(x) - \mathbb{E}\left[h(Z)\right]$$

with unknown function f for any $x \in B$.

Corollary 4.11. *Let W be a B-valued random variable with finite first moment, that is, $\mathbb{E}[\|W\|] < +\infty$. Then we have the following bound on the Wasserstein distance $d_\omega(W, Z)$ between W and Z:*

$$\begin{aligned} d_\omega(W, Z) &\equiv \sup_{\|h\|_{\mathrm{ULip}} \leqslant 1} \left|\mathbb{E}\left[h(W)\right] - \mathbb{E}\left[h(Z)\right]\right| \\ &\leqslant \sup_{\varphi \in \mathcal{F}_{\mu_Z}} \left|\mathbb{E}\left[\Delta_G \varphi(W) - \left(W, D\varphi(W)\right)_{B,B^*}\right]\right|,\end{aligned}$$

where $\mathcal{F}_{\mu_Z}$ is the class of twice H-differentiable functions φ on B so that $D\varphi(x) \in B^*$ with $\|D\varphi(x)\|_{B^*} \leqslant 1$ and $D^2\varphi(x) \in \mathcal{TR}(H)$ with $\|D^2\varphi(x)\|_{\mathrm{tr}(H)} \leqslant \frac{\pi}{2}\int_B \|y\|\,\mu_Z(dy)$ for any $x \in B$.

5. Gaussian approximation in abstract Wiener spaces

As assumed in the beginning of Section 3, let Z be a fixed B-valued random variable on some probability space $(\Omega, \mathcal{F}, \mathcal{P})$ such that the distribution μ_Z of Z is a non-degenerate Gaussian measure on B with mean zero, and H the unique RKHS for μ_Z with the inner product $\langle\cdot,\cdot\rangle_0$ and the induced norm $|\cdot|_0$, where the triple (i, H, B) forms an AWS and μ_Z associated with the abstract Wiener measure μ_Z on B with variance parameter 1. Select another norm $|\cdot|_{-1}$ on B, which is induced by an inner product $\langle\cdot,\cdot\rangle_{-1}$ and weaker than $\|\cdot\|$-norm. If $\|\cdot\|$-norm is Hilbertian, we will set $|\cdot|_{-1} = \|\cdot\|$. Let K be the completion of B with respect to $|\cdot|_{-1}$-norm and μ_Z^K the Gaussian measure on K given by $\mu_Z^K(E) = \mu_Z(E \cap B)$ for any $E \in \mathcal{B}(K)$. Let $S_{\mu_Z^K}$ be the covariance operator of μ_Z^K on K. By Theorem 2.9, $S_{\mu_Z^K}$ is one-to-one and $H = \sqrt{S_{\mu_Z^K}}(K)$. In what follows, let $\{k_j;\ j \in \mathbb{N}\}$ be an orthonormal basis of K, consisting of eigenvectors of $S_{\mu_Z^K}$ with corresponding eigenvalues λ_j, $j \in \mathbb{N}$. Then λ_j's > 0 and $\sum \lambda_j < +\infty$. In addition, by the standard construction of countably Hilbert spaces from $(H, \sqrt{S_{\mu_Z^K}}^{-1})$ as given in Remark 2.10, we have the following chain of dense, continuous embeddings:

$$H_2 \subset H_1 \subset B^* \subset H \subset B \subset H_{-1} = K \subset H_{-2},$$

where H_{-n}^*, $n = 1, 2$, is identified with H_n as a Borel dense subset of H from the viewpoint of Remark 2.10. The inner product and induced norm on H_r, $r = \pm 1, \pm 2$, are denoted respectively by $\langle\cdot,\cdot\rangle_r$ and $|\cdot|_r$.

In this section, we aim to establish an infinite-dimensional version of Stein's method of exchangeable pairs (see e.g. [32, p. 33$\sim$]) along the line of Stein's idea. Let W and W^* be two B-valued random variables with finite first moment on a probability space $(\Omega, \mathcal{P})$. Suppose further that (W, W^*) is an exchangeable pair by which we mean that

$$\mathcal{P}\big(\big\{W \in E_1,\ W^* \in E_2\big\}\big) = \mathcal{P}\big(\big\{W^* \in E_1,\ W \in E_2\big\}\big) \quad \text{for any } E_1, E_2 \in \mathcal{B}(B).$$

In addition, there is a constant $\lambda \in (0, 1)$ such that

$$\mathbb{E}\big[\big(W^* - W, \varphi(W)\big)_{B,B^*}\big] = -\lambda \mathbb{E}\big[\big(W, \varphi(W)\big)_{B,B^*}\big], \tag{5.1}$$

where φ is an arbitrarily given bounded function from B into B^*.

Take a fixed $h \in \mathcal{U}Lip\text{-}1(B)$. Let f_h be given as one in (4.2), the first and second H-derivatives of which satisfy the following uniformly Lipschitz conditions:

$$\begin{cases} \displaystyle\sup_{x \neq y \in B} \sup_{z \in H \setminus \{0\}} \frac{|\langle Df_h(x) - Df_h(y), z\rangle|}{|z|_{-2}|x - y|_{-2}} \equiv \|Df_h\|_{\mathrm{ULip}(H_{-2})} < +\infty; \\ \displaystyle\sup_{x \neq y \in B} \sup_{z_1, z_2 \in H \setminus \{0\}} \frac{|\langle (D^2 f_h(x) - D^2 f_h(y)) z_1, z_2\rangle|}{\|z_1\|\|z_2\|\|x - y\|} \equiv \big\|D^2 f_h\big\|_{\mathrm{ULip}} < +\infty. \end{cases} \tag{5.2}$$

Consider the function $\phi_h : \mathbb{R} \times B \times B \to \mathbb{C}$ defined by

$$\phi_h(r, x, y) = f_h\big(rx + (1 - r)y\big).$$

Lemma 5.1. *Let $h \in \mathcal{U}Lip$-$1(B)$ with the property* (5.2).

(i) *For any $x \in B$, there is a bounded linear operator $T_{h,x}$ from H_{-2} to H^*_{-2} such that the restriction of $T_{h,x}$ to $H \times H$ equals $D^2 f_h(x)$. Moreover,*

$$\|T_{h,x}\|_{H_{-2},H^*_{-2}} \leqslant \|Df_h\|_{\mathrm{ULip}(H_{-2})}.$$

(ii) *Fix $x, y \in B$. Then $\phi_h(r, x, y)$ is twice differentiable with respect to $r \in \mathbb{R}$. Further,*

$$\frac{d}{dr}\phi_h(r, x, y) = \big(x - y, Df_h\big(rx + (1 - r)y\big)\big)_{B,B^*};$$
$$\frac{d^2}{dr^2}\phi_h(r, x, y) = \big(x - y, T_{h,rx+(1-r)y}(x - y)\big)_{B,B^*}.$$

Proof. (i) For any $z_1, z_2 \in H$, we see by the second-order H-differentiability of f_h and (5.2) that

$$\begin{aligned}\big|\big\langle D^2 f_h(x)z_1, z_2\big\rangle\big| &= \lim_{r\to 0} |r|^{-1}\big|\big\langle Df_h(x + rz_1) - Df_h(x), z_2\big\rangle\big| \\ &\leqslant \|Df_h\|_{\mathrm{ULip}(H_{-2})}|z_1|_{-2}|z_2|_{-2}.\end{aligned}$$

Consequently, by the denseness of H in H_{-2}, we can extend by continuity $D^2 f_h(x)$ to the whole Cartesian product $H_{-2} \times H_{-2}$, proving the assertion (i).

(ii) Observe that for any $r \in \mathbb{R}$,

$$\big|s^{-1}\big(\mathcal{T}_t h\big(y + (r + s)(x - y)\big) - \mathcal{T}_t h\big(y + r(x - y)\big)\big)\big| \leqslant \mathrm{Const.}\, e^{-t},$$

where such a constant depends only on h, x, and y, and the function e^{-t}, $t \geqslant 0$, is in $L^1([0, +\infty), dt)$. Then, applying the dominated convergence argument, we see by Lemma 4.2(ii) and Theorem 4.9(i) that

$$\begin{aligned}\frac{d}{dr}\phi_h(r, x, y) &= -\lim_{s\to 0} s^{-1}\int_0^\infty \big(\mathcal{T}_t h\big(y + (r + s)(x - y)\big) - \mathcal{T}_t h\big(y + r(x - y)\big)\big)\, dt \\ &= -\int_0^\infty \big(x - y, D(\mathcal{T}_t h)\big(y + r(x - y)\big)\big)_{B,B^*}\, dt \\ &= \big(x - y, Df_h\big(y + r(x - y)\big)\big)_{B,B^*}.\end{aligned}$$

To show the second-order differentiability of $\phi_h(r, x, y)$ with respect to r, we take an approximating sequence $\{z_n\} \subset H$ such that $\|z_n - (x - y)\| \to 0$ as $n \to +\infty$. Set $a = y + r(x - y)$. Then

$$\begin{aligned}
&\left|s^{-1}\left\{\frac{d}{dr}\phi_h\big(a+s(x-y)\big)-\frac{d}{dr}\phi_h(a)\right\}-\big(x-y,T_{h,a}(x-y)\big)_{B,B^*}\right|\\
&\quad=\left|s^{-1}\big(\big(x-y,Df_h\big(a+s(x-y)\big)-Df_h(a)\big)_{B,B^*}\big)-\big(x-y,T_{h,a}(x-y)\big)_{B,B^*}\right|\\
&\quad\leqslant\left|\big(x-y,T_{h,a}(x-y)-T_{h,a}(z_n)\big)_{B,B^*}\right|\\
&\qquad+\left|s^{-1}\big(\big(x-y,Df_h\big(a+s(x-y)\big)-Df_h(a+sz_n)\big)_{B,B^*}\big)\right|\\
&\qquad+\left|s^{-1}\big(\big(x-y,Df_h(a+sz_n)-Df_h(a)\big)_{B,B^*}\big)-\big(x-y,T_{h,a}(z_n)\big)_{B,B^*}\right|\\
&\quad\equiv I_1+I_2+I_3.
\end{aligned}$$

It is clear that $I_1\to 0$, and by (5.2)

$$I_2\leqslant \text{Const.}\|x-y\|\big|z_n-(x-y)\big|_{-2}\to 0\quad\text{as } n\to+\infty.$$

For I_3, we see by virtue of (i) and (5.2) that

$$\begin{aligned}
I_3&\leqslant\left|s^{-1}\big(\big(x-y-z_n,Df_h(a+sz_n)-Df_h(a)\big)_{B,B^*}\big)-\big(x-y-z_n,T_{h,a}(z_n)\big)_{B,B^*}\right|\\
&\quad+\left|s^{-1}\big(\big(z_n,Df_h(a+sz_n)-Df_h(a)\big)_{B,B^*}\big)-\big(z_n,T_{h,a}(z_n)\big)_{B,B^*}\right|\\
&\leqslant\text{Const.}\cdot\|x-y-z_n\||z_n|_{-2}\\
&\quad+\left|s^{-1}\big(\big(z_n,Df_h(a+sz_n)-Df_h(a)\big)_{B,B^*}\big)-\big(z_n,T_{h,a}(z_n)\big)_{B,B^*}\right|.
\end{aligned}$$

Then, by Theorem 4.9,

$$\limsup_{s\to 0} I_3\leqslant\text{Const.}\|x-y-z_n\||z_n|_{-2},$$

which approaches to zero as $n\to+\infty$. The proof is complete. □

To avoid the notational complexity, we still use the symbol "$D^2 f_h(x)$" in place of "$T_{h,x}$" for any $h\in\mathcal{U}Lip\text{-}1(B)$ with the property (5.2).

By Lemma 5.1, we can apply Taylor's theorem to $\phi_h(r,W^*,W)$ with respect to r, and then obtain that

$$f_h\big(W^*\big)-f_h(W)=\frac{d}{dr}\bigg|_{r=0}\phi_h\big(r,W^*,W\big)+\frac{1}{2}\frac{d^2}{dr^2}\bigg|_{r=0}\phi_h\big(r,W^*,W\big)+R_h\big(W^*,W\big),\quad(5.3)$$

where $R_h(W^*,W)$ is the second-order remainder term. Having been seen in the proof of Lemma 5.1,

$$\frac{d}{dr}\bigg|_{r=0}\phi_h\big(r,W^*,W\big)=\big(W^*-W,Df_h(W)\big)_{B,B^*}.$$

Note that it follows from Theorem 4.9(ii) that the function $\|Df_h(x)\|_{B^*}$, $x \in B$, is bounded by $\|h\|_{\mathrm{ULip}}$ uniformly. Then, by our assumption in (5.1),

$$\mathbb{E}\left[\frac{d}{dr}\bigg|_{r=0}\phi_h\big(r, W^*, W\big)\right] = -\lambda\mathbb{E}\big[\big(W, Df_h(W)\big)_{B,B^*}\big]. \tag{5.4}$$

Next, by Lemma 5.1(ii),

$$\frac{1}{2}\frac{d^2}{dr^2}\bigg|_{r=0}\phi_h\big(r, W^*, W\big) = \frac{1}{2}\big(W^* - W, D^2 f_h(W)\big(W^* - W\big)\big)_{B,B^*}. \tag{5.5}$$

Since Lemma 5.1(i) implies that $D^2 f_h(x)$ is a continuous bilinear map on $K \times K$ for any $x \in B$, the right-hand side of (5.5) can be expanded with respect to $\{k_j\}$ as follows:

$$\begin{aligned}
&\frac{1}{2}\sum_{i,j=1}^{\infty}\big\langle W^* - W, k_i\big\rangle_{-1}\big\langle W^* - W, k_j\big\rangle_{-1}\big(k_i, D^2 f_h(W)k_j\big)_{K,K^*} \\
&\quad = \lambda\sum_{i,j=1}^{\infty}\big\langle S_{\mu_Z^K}(k_i), k_j\big\rangle_{-1}\big(k_i, D^2 f_h(W)k_j\big)_{K,K^*} + \lambda\sum_{i,j=1}^{\infty} E_{i,j}\big(k_i, D^2 f_h(W)k_j\big)_{K,K^*},
\end{aligned} \tag{5.6}$$

where, for any $i, j \in \mathbb{N}$,

$$E_{i,j} = \frac{1}{2\lambda}\big\langle W^* - W, k_i\big\rangle_{-1}\big\langle W^* - W, k_j\big\rangle_{-1} - \big\langle S_{\mu_Z^K}(k_i), k_j\big\rangle_{-1}.$$

Since $\langle S_{\mu_Z^K}(k_i), k_j\rangle_{-1} = \lambda_i\delta_{i,j}$ for any i, j, the first term in the second line of (5.6) equals

$$\begin{aligned}
\lambda\sum_{i=1}^{\infty}\big(\sqrt{\lambda_i}k_i, D^2 f_h(W)\sqrt{\lambda_i}k_i\big)_{K,K^*} &= \lambda\sum_{i=1}^{\infty}\big\langle D^2 f_h(W)\sqrt{\lambda_i}k_i, \sqrt{\lambda_i}k_i\big\rangle \\
&= \lambda\Delta_G f_h(W),
\end{aligned}$$

whence

$$\frac{1}{2}\frac{d^2}{dr^2}\bigg|_{r=0}\phi\big(r, W^*, W\big) = \lambda\Delta_G f_h(W) + \lambda\sum_{i,j=1}^{\infty} E_{i,j}\big(k_i, D^2 f_h(W)k_j\big)_{K,K^*}. \tag{5.7}$$

One notes that $f_h(W^*) - f_h(W)$ is integrable with respect to $\mathcal{P}$, because

$$\mathbb{E}\big[\big|f_h\big(W^*\big) - f_h(W)\big|\big] \leqslant \|h\|_{\mathrm{ULip}}\mathbb{E}\big[\big\|W^* - W\big\|\big] < +\infty.$$

Since (W, W^*) is an exchangeable pair and the mapping $g(x, y) = f_h(x) - f_h(y)$, $(x, y) \in B \times B$, is anti-symmetric,

$$\mathbb{E}\big[f_h\big(W^*\big) - f_h(W)\big] = 0.$$

Taking expectation of Eq. (5.3) together with (5.4) and (5.7),

$$\mathbb{E}\big[\big(W, Df_h(W)\big)_{B,B^*}\big] - \mathbb{E}\big[\Delta_G f_h(W)\big] = \mathbb{E}\Bigg[\sum_{i,j=1}^{\infty} E_{i,j}\big(k_i, D^2 f_h(W)k_j\big)_{K,K^*}\Bigg] + \frac{1}{\lambda}\mathbb{E}\big[R_h\big(W^*, W\big)\big]. \tag{5.8}$$

Therefore, by applying Theorem 4.10 to (5.8), we see that

$$\mathbb{E}\big[h(Z)\big] - \mathbb{E}\big[h(W)\big] = \mathbb{E}\Bigg[\sum_{i,j=1}^{\infty} E_{i,j}\big(k_i, D^2 f_h(W)k_j\big)_{K,K^*}\Bigg] + \frac{1}{\lambda}\mathbb{E}\big[R_h\big(W^*, W\big)\big]. \tag{5.9}$$

For any $a, b \in K$,

$$T_{\lambda,a,b,K}(x, y) \equiv \frac{1}{2\lambda}\langle b-a, x\rangle_{-1}\langle b-a, y\rangle_{-1} - \big\langle S_{\mu_Z^K}(x), y\big\rangle_{-1}, \quad x, y \in K,$$

is a continuous bilinear map on $K \times K$. Let $\mathcal{HS}(K^{\otimes 2})$ denote the collection of Hilbert–Schmidt type bilinear maps on $K \times K$, that is,

$$T \in \mathcal{HS}\big(K^{\otimes 2}\big) \quad \text{if and only if} \quad \sum_{i,j} T(f_i, f_j)^2 < +\infty$$

for any orthonormal basis $\{f_k\}$ of K. Then $\mathcal{HS}(K^{\otimes 2})$ is a Hilbert space with the inner product given by

$$\langle\!\langle S, T\rangle\!\rangle_{\mathrm{HS}(K^{\otimes 2})} = \sum_{i,j} S(f_i, f_j)T(f_i, f_j),$$

and the induced norm denoted by $\|\cdot\|_{\mathrm{HS}(K^{\otimes 2})}$. It is straightforward to see that $T_{\lambda,a,b,K} \in \mathcal{HS}(K^{\otimes 2})$ for any $a, b \in K$, and, in the following Remark 5.3, $D^2 f_h(x)$ is also in $\mathcal{HS}(K^{\otimes 2})$ for any $x \in B$. Therefore, we have

$$\mathbb{E}\Bigg[\sum_{i,j=1}^{\infty} E_{i,j}\big(k_i, D^2 f_h(W)k_j\big)_{K,K^*}\Bigg] = \mathbb{E}\big[\big\langle\!\big\langle T_{\lambda,W,W^*,K}, D^2 f_h(W)\big\rangle\!\big\rangle_{\mathrm{HS}(K^{\otimes 2})}\big]. \tag{5.10}$$

For the remainder term $\frac{1}{\lambda}\mathbb{E}[R_h(W^*, W)]$, we apply the mean-value theorem for differentiation together with Lemma 5.1 and (5.2) to see that there is $\vartheta \in (0, 1)$ such that

$$\begin{aligned}
\frac{1}{\lambda}\big|R_h\big(W^*, W\big)\big| &= \frac{1}{6\vartheta\lambda}\left|\frac{d^2}{dr^2}\bigg|_{r=\vartheta}\phi_h\big(r, W^*, W\big) - \frac{d^2}{dr^2}\bigg|_{r=0}\phi_h\big(r, W^*, W\big)\right| \\
&= \frac{1}{6\vartheta\lambda}\big|\big(W^* - W, \big(D^2 f_h\big(\vartheta W^* + (1-\vartheta)W\big) - D^2 f_h(W)\big)\big(W^* - W\big)\big)_{B,B^*}\big| \\
&\leqslant \frac{1}{6\lambda}\big\|D^2 f_h\big\|_{\mathrm{ULip}}\big\|W^* - W\big\|^3,
\end{aligned}$$

whence

$$\frac{1}{\lambda}\left|\mathbb{E}\left[R_h\left(W^*,W\right)\right]\right| \leqslant \frac{1}{6\lambda}\left\|D^2 f_h\right\|_{\mathrm{ULip}}\mathbb{E}\left[\left\|W^*-W\right\|^3\right]. \tag{5.11}$$

To sum up the above argument together with (5.9), (5.10), and (5.18), we have the following

Theorem 5.2. *Let (W, W^*) be an exchangeable pair of B-valued random variables with finite first moment. Assume that there is a constant $\lambda \in (0, 1)$ such that*

$$\mathbb{E}\left[\left(W^*-W,\varphi(W)\right)_{B,B^*}\right] = -\lambda\mathbb{E}\left[\left(W,\varphi(W)\right)_{B,B^*}\right],$$

where φ is an arbitrarily given bounded function from B into B^*. Let $h \in \mathcal{U}Lip\text{-}1(B)$ associated with f_h having the properties as given in (5.2). Then

$$\begin{aligned}&\left|\mathbb{E}\left[h(W)\right]-\mathbb{E}\left[h(Z)\right]\right|\\&\quad\leqslant\left|\mathbb{E}\left[\left\langle\!\left\langle T_{\lambda,W,W^*,K}, D^2 f_h(W)\right\rangle\!\right\rangle_{\mathrm{HS}(K^{\otimes 2})}\right]\right|+\frac{1}{6\lambda}\left\|D^2 f_h\right\|_{\mathrm{ULip}}\mathbb{E}\left[\left\|W^*-W\right\|^3\right].\end{aligned} \tag{5.12}$$

Remark 5.3. The Hilbert–Schmidt norm $\|D^2 f_h(x)\|^2_{\mathrm{HS}(K^{\otimes 2})}$, $x \in B$, can be estimated as follows. First, since $|k|_{-2} = |\sqrt{S_{\mu_Z^K}}(k)|_{-1}$ for any $k \in K$, one notes that the triple $(i_{K,H_{-2}}, K, H_{-2})$ forms an AWS, where $i_{K,H_{-2}}$ is the canonical embedding of K into H_{-2}. By Lemma 5.1(i) and using Kuo's theorem (see [17, Chapter I, Corollary 4.4]), every $D^2 f_h(x)$, $x \in B$, can be regarded as a Hilbert–Schmidt operator on K satisfying the following inequality

$$\left\|D^2 f_h(x)\right\|_{\mathrm{HS}(K)} \leqslant \left\|D^2 f_h(x)\right\|_{H_{-2},K}\left\{\int_{H_{-2}} |y|^2_{-2}\, p_1^{(K,H_{-2})}(dy)\right\}^{\frac{1}{2}},$$

where $p_1^{(K,H_{-2})}$ is the associated abstract Wiener measure of $(i_{K,H_{-2}}, K, H_{-2})$. Since the covariance operator of $p_1^{(K,H_{-2})}$ is $S_{\mu_Z^K}$,

$$\begin{aligned}\left\{\int_{H_{-2}} |y|^2_{-2}\, p_1^{(K,H_{-2})}(dy)\right\}^{\frac{1}{2}} &= \left\{\sum_{j=1}^{\infty}\int_{H_{-2}}\left\langle y,\lambda_j^{-\frac{1}{2}}k_j\right\rangle^2_{-2} p_1^{(K,H_{-2})}(dy)\right\}^{\frac{1}{2}}\\&=\left\{\sum_{j=1}^{\infty}\left\langle S_{\mu_Z^K}\left(\lambda_j^{-\frac{1}{2}}k_j\right),\lambda_j^{-\frac{1}{2}}k_j\right\rangle_{-2}\right\}^{\frac{1}{2}}\\&=\sqrt{\mathrm{Tr}_K(S_{\mu_Z^K})}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\left\|D^2 f_h(x)\right\|_{H_{-2},K} &= \sup_{z\in H_{-2}\setminus\{0\}} |z|_{-2}^{-1}\left\{\sum_{j=1}^{\infty}\big(k_j, D^2 f_h(x)(z,\cdot)\big)^2_{H_{-2},H^*_{-2}}\right\}^{\frac{1}{2}}\\
&\leqslant \|S_{\mu_Z^K}\|_{K,K} \sup_{z\in H_{-2}\setminus\{0\}} |z|_{-2}^{-1}\left\{\sum_{j=1}^{\infty}\big(\lambda_j^{-\frac{1}{2}}k_j, D^2 f_h(x)(z,\cdot)\big)^2_{H_{-2},H^*_{-2}}\right\}^{\frac{1}{2}}\\
&= \|S_{\mu_Z^K}\|_{K,K}\left\|D^2 f_h(x)\right\|_{H_{-2},H^*_{-2}}\\
&\leqslant \|Df_h\|_{\mathrm{ULip}(H_{-2})}\|S_{\mu_Z^K}\|_{K,K}, \quad \text{uniformly for } x\in B.\end{aligned}$$

From the above estimations, it follows that

$$\left\|D^2 f_h(x)\right\|_{\mathrm{HS}(K^{\otimes 2})} \leqslant \|Df_h\|_{\mathrm{ULip}(H_{-2})}\|S_{\mu_Z^K}\|_{K,K}\sqrt{\mathrm{Tr}_K(S_{\mu_Z^K})}, \tag{5.13}$$

uniformly for $x \in B$.

Application: ***Error bounds in a Lindeberg–Lévy type limit theorem***

As an illustration of Theorem 5.2, we next will derive an explicit error estimate of Gaussian approximation to the distribution of a sum of independent and identically distributed B-valued random variables based on a Lindeberg–Lévy type limit theorem (see [33] for multivariate version) by the abstract Wiener measure μ_Z, where the bound is computed for the difference between the expectations of any one of uniformly Lip-1 functions on B having the condition (5.2).

Let $\{X_1, X_2, \ldots\}$ be a sequence of independent, identically distributed B-valued random variables with finite first moment on the probability space $(\Omega, \mathcal{P})$. Suppose further that they satisfy the following conditions:

5-(a) X_1 has zero mean, that is, the B-valued integral $\mathbb{E}[X_1]$ equals zero.
5-(b) The distribution μ_{X_1} has the same covariance function as μ_Z. That is,

$$\mathbb{E}\big[(X_1,\eta)_{B,B^*}(X_1,\phi)_{B,B^*}\big] = \langle \eta,\phi\rangle_0, \quad \forall \eta,\phi \in B^*.$$

Remark 5.4.

1. For any $f, g \in K^*$, condition 5-(b) implies that

$$\mathbb{E}\big[(X_1, f|_B)_{B,B^*}(X_1, g|_B)_{B,B^*}\big] = \langle f,g\rangle_0 = \big\langle S^{-1}_{\mu_Z^K}(f), g\big\rangle_{-1}.$$

2. By condition 5-(a), we see that for any $f \in K^*$,

$$\big(\mathbb{E}[X_1], f\big)_{K,K^*} = \mathbb{E}\big[(X_1, f|_B)_{B,B^*}\big] = 0$$

Equivalently, $\langle \mathbb{E}[X_1], x\rangle_{-1} = \mathbb{E}[(X_1, S_{\mu_Z^K}(x))_{B,B^*}] = 0$ for any $x \in K$.

Fix $n \in \mathbb{N}$, and set

$$W_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i.$$

Let $\{Y_1, Y_2, \dots, Y_n\}$ be an independent copy of $\{X_1, X_2, \dots, X_n\}$, and I be a random variable which is uniformly distributed over the index set $\{1, 2, \dots, n\}$, and also independent of $\{X_i\}$ and $\{Y_i\}$. Define

$$W_n^* = W_n + \frac{1}{\sqrt{n}}(Y_I - X_I) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(Y_j + \sum_{i \neq j} X_i \right) \mathbf{1}_{\{I=j\}}.$$

It is easy to see that, for any $\eta, \phi \in B^*$ and $s, t \in \mathbb{R}$,

$$\mathbb{E}\left[e^{\mathrm{i}(t(W_n^*,\eta)_{B,B^*} + s(W_n,\phi)_{B,B^*})}\right] = \mathbb{E}\left[e^{\mathrm{i}(t(W_n,\eta)_{B,B^*} + s(W_n^*,\phi)_{B,B^*})}\right],$$

implying that (W_n, W_n^*) is an exchangeable pair of B-valued random variables.

Observe that, for any bounded function φ from B into B^*,

$$\begin{aligned} \mathbb{E}\left[\left(W_n^* - W_n, \varphi(W_n)\right)_{B,B^*}\right] &= \frac{1}{\sqrt{n}} \mathbb{E}\left[\left(Y_I - X_I, \varphi(W_n)\right)_{B,B^*}\right] \\ &= \frac{1}{n\sqrt{n}} \sum_{j=1}^{n} \mathbb{E}\left[\left(Y_j - X_j, \varphi(W_n)\right)_{B,B^*}\right]. \end{aligned} \tag{5.14}$$

By the independence of Y_j's and W_n and condition 5-(a), it follows that

$$\mathbb{E}\left[\left(Y_j, \varphi(W_n)\right)_{B,B^*}\right] = \int_B \mathbb{E}\left[\left(Y_j, \varphi(x)\right)_{B,B^*}\right] \mu_{W_n}(dx) = 0,$$

where μ_{W_n} is the distribution of W_n in B. Combining with (5.14), we see that the assumption (5.1) holds for the pair (W_n, W_n^*) by taking $\lambda = \frac{1}{n}$.

Let $h \in \mathcal{U}Lip\text{-}1(B)$ associated with f_h having the properties as given in (5.2). Then

$$\begin{aligned} &\left|\mathbb{E}\left[\left\langle\!\left\langle T_{\lambda, W_n, W_n^*, K}, D^2 f_h(W_n) \right\rangle\!\right\rangle_{\mathrm{HS}(K^{\otimes 2})}\right]\right| \\ &= \left| \sum_{\ell=1}^{n} \mathbb{E}\left[\mathbf{1}_{\{I=\ell\}} \left\langle\!\left\langle T_{1, X_\ell, Y_\ell, K}, D^2 f_h(W_n) \right\rangle\!\right\rangle_{\mathrm{HS}(K^{\otimes 2})}\right] \right| \\ &= \left| \sum_{\ell=1}^{n} \mathcal{P}\left(\{I=\ell\}\right) \mathbb{E}\left[\left\langle\!\left\langle T_{1, X_\ell, Y_\ell, K}, D^2 f_h(W_n) \right\rangle\!\right\rangle_{\mathrm{HS}(K^{\otimes 2})}\right] \right| \\ &\leqslant \frac{1}{n} \mathbb{E}\left[\left| \left\langle\!\!\left\langle \sum_{\ell=1}^{n} T_{1, X_\ell, Y_\ell, K}, D^2 f_h(W_n) \right\rangle\!\!\right\rangle_{\mathrm{HS}(K^{\otimes 2})} \right|\right] \end{aligned}$$

$$\leqslant \frac{1}{n}\left\{\mathbb{E}\left[\left\|D^2 f_h(W_n)\right\|^2_{\mathrm{HS}(K)}\right]\right\}^{\frac{1}{2}}\left\{\mathbb{E}\left[\left\|\sum_{\ell=1}^{n} T_{1,X_\ell,Y_\ell,K}\right\|^2_{\mathrm{HS}(K^{\otimes 2})}\right]\right\}^{\frac{1}{2}}, \tag{5.15}$$

where the last inequality is obtained by applying the Cauchy–Schwarz inequality. Also,

$$\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{\ell=1}^{n} T_{1,X_\ell,Y_\ell,K}\right\|^2_{\mathrm{HS}(K^{\otimes 2})}\right] \\
&= \sum_{i,j=1}^{\infty}\sum_{\ell=1}^{n}\mathbb{E}\left[\left(\frac{1}{2}\langle Y_\ell - X_\ell, k_i\rangle_{-1}\langle Y_\ell - X_\ell, k_j\rangle_{-1} - \left\langle S_{\mu_Z^K}(k_i), k_j\right\rangle_{-1}\right)^2\right] \\
&= n\sum_{i,j=1}^{\infty}\left\{\frac{1}{4}\mathbb{E}\left[\langle Y_1 - X_1, k_i\rangle_{-1}^2\langle Y_1 - X_1, k_j\rangle_{-1}^2\right] - \left\langle S_{\mu_Z^K}(k_i), k_j\right\rangle_{-1}^2\right\} \\
&= \frac{n}{4}\left(\mathbb{E}\left[|Y_1 - X_1|_{-1}^4\right] - 4\|S_{\mu_Z^K}\|^2_{\mathrm{HS}(K)}\right) \\
&\leqslant n\left(4\mathbb{E}\left[|X_1|_{-1}^4\right] - \|S_{\mu_Z^K}\|^2_{\mathrm{HS}(K)}\right),
\end{aligned} \tag{5.16}$$

where the third and last lines are obtained respectively by the formula

$$\mathbb{E}\left[\frac{1}{2}\langle Y_1 - X_1, k_i\rangle_{-1}\langle Y_1 - X_1, k_j\rangle_{-1} - \left\langle S_{\mu_Z^K}(k_i), k_j\right\rangle_{-1}\right] = 0, \quad \forall i, j \in \mathbb{N}$$

and the inequality

$$\mathbb{E}\left[|Y_1 - X_1|_{-1}^4\right] \leqslant \mathbb{E}\left[\left(|Y_1|_{-1} + |X_1|_{-1}\right)^4\right] \leqslant 16\mathbb{E}\left[|X_1|_{-1}^4\right].$$

Combine (5.15) with (5.13) and (5.16) to get that

$$\begin{aligned}
&\left|\mathbb{E}\left[\langle\!\langle T_{\lambda,W_n,W_n^*,K}, D^2 f_h(W_n)\rangle\!\rangle_{\mathrm{HS}(K^{\otimes 2})}\right]\right| \\
&\leqslant \frac{1}{\sqrt{n}}\|Df_h\|_{\mathrm{ULip}(H_{-2})}\|S_{\mu_Z^K}\|_{K,K}\left\{\mathrm{Tr}_K(S_{\mu_Z^K})\left(4\mathbb{E}\left[|X_1|_{-1}^4\right] - \|S_{\mu_Z^K}\|^2_{\mathrm{HS}(K)}\right)\right\}^{\frac{1}{2}}.
\end{aligned} \tag{5.17}$$

In addition,

$$\begin{aligned}
\mathbb{E}\left[\left\|W_n^* - W_n\right\|^3\right] &= \frac{1}{n\sqrt{n}}\mathbb{E}\left[\|Y_I - X_I\|^3\right] \\
&= \frac{1}{n\sqrt{n}}\mathbb{E}\left[\|Y_1 - X_1\|^3\right] \leqslant \frac{8}{n\sqrt{n}}\mathbb{E}\left[\|X_1\|^3\right].
\end{aligned} \tag{5.18}$$

Consequently, applying (5.17) and (5.18) to (5.12) with $(W, W^*) = (W_n, W_n^*)$ and $\lambda = \frac{1}{n}$, we have an explicit error bound in the following infinite-dimensional version of Lindeberg–Lévy type limit theorem.

Theorem 5.5. *Let $\{X_1, X_2, \ldots\}$ be a sequence of independent, identically distributed B-valued random variables with finite first moment on the same probability space. Suppose that X_1 has zero mean and the distribution μ_{X_1} has the same covariance function as μ_Z. Then, for any $h \in \mathcal{U}Lip\text{-}1(B)$ associated with f_h having the properties as given in* (5.2)*,*

$$\left|\mathbb{E}\big[h(Z)\big] - \mathbb{E}\left[h\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)\right]\right| \leqslant M(Z, X_1, h)\frac{1}{\sqrt{n}},$$

where $M(Z, X_1, h)$ is a constant depending only on Z, X_1, and h. In fact,

$$M(Z, X_1, h) = \|Df_h\|_{\mathrm{ULip}(H_{-2})}\|S_{\mu_Z^K}\|_{K,K}\big\{\mathrm{Tr}_K(S_{\mu_Z^K})\big(4\mathbb{E}\big[|X_1|_{-1}^4\big] - \|S_{\mu_Z^K}\|_{\mathrm{HS}(K)}^2\big)\big\}^{\frac{1}{2}} + \frac{4}{3}\big\|D^2 f_h\big\|_{\mathrm{ULip}}\mathbb{E}\big[\|X_1\|^3\big].$$

Multivariate normal approximation. Consider the finite-dimensional case: $B = \mathbb{R}^k$ with $k \geqslant 2$, where $\|\cdot\|$-norm is the Euclidean norm $|\cdot|$ induced by the Euclidean inner product $\langle\cdot,\cdot\rangle$, as well as μ_Z is a non-degenerate multivariate normal distribution in $\mathbb{R}^k$ with mean zero and the covariance matrix $\mathbf{A}$. Then $\mathbf{A}$ is a non-singular, self-adjoint, and positive-definite $k \times k$ matrix over $\mathbb{R}$. In this case, we set $B = K$ and $S_{\mu_Z^K}$ is the mapping on $\mathbb{R}^k$ by $S_{\mu_Z^K}(x) = \mathbf{A}x$, $x \in \mathbb{R}^k$. Hence $H_r = \mathbb{R}^k$ with $|\cdot|_r$-norm given by $|x|_r = |\sqrt{\mathbf{A}}^{-1-r}x|$ for any $x \in H_r$ with $r = 0, \pm 1, \pm 2, \ldots$.

Let h be a continuously differentiable function on $\mathbb{R}^k$, denoted by $h \in C^1(\mathbb{R}^k)$, with $\|h\|_{\mathrm{ULip}} < +\infty$ and $\|\nabla h\|_{\mathrm{ULip}} < +\infty$, where

$$\|\nabla h\|_{\mathrm{ULip}} \equiv \sup_{x \neq y \in \mathbb{R}^k} \frac{|\nabla h(x) - \nabla h(y)|}{|x - y|}.$$

Then it follows from (4.6) and (4.25) that

$$\big|\nabla f_h(x) - \nabla f_h(y)\big| \leqslant \|h\|_{\mathrm{ULip}}|x - y|, \quad \forall x, y \in \mathbb{R}^k.$$

Since, for any $z \in \mathbb{R}^k$, $\|z\| = |z| = |z|_{-1} \leqslant \|\sqrt{\mathbf{A}}^{-1}\|_{\mathrm{op}}|z|_{-2}$, we see that

$$\begin{aligned}\sup_{z \in \mathbb{R}^k \setminus \{0\}} \frac{|\langle \nabla f_h(x) - \nabla f_h(y), z\rangle|}{|z|_{-2}} &\leqslant \big\|\sqrt{\mathbf{A}}^{-1}\big\|_{\mathrm{op}}\big|\nabla f_h(x) - \nabla f_h(y)\big| \\ &\leqslant \big\|\sqrt{\mathbf{A}}^{-1}\big\|_{\mathrm{op}}\|h\|_{\mathrm{ULip}}|x - y| \\ &\leqslant \big\|\sqrt{\mathbf{A}}^{-1}\big\|_{\mathrm{op}}^2\|h\|_{\mathrm{ULip}}|x - y|_{-2}, \quad \forall x, y \in \mathbb{R}^k,\end{aligned}$$

implying

$$\|Df_h\|_{\mathrm{ULip}(H_{-2})} \leqslant \big\|\sqrt{\mathbf{A}}^{-1}\big\|_{\mathrm{op}}^2\|h\|_{\mathrm{ULip}},$$

where $\|\sqrt{\mathbf{A}}^{-1}\|_{\mathrm{op}} = \sup\{|\langle\sqrt{\mathbf{A}}^{-1}z, w\rangle|;\ |z| = |w| = 1\}$. In addition,

$$\|D^2 f_h\|_{\mathrm{ULip}} \leqslant \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_0^{\infty} \frac{e^{-\frac{t^2}{2}-3s}|t|}{\sqrt{1-e^{-2s}}}\,ds\,dt \cdot \|\nabla h\|_{\mathrm{ULip}} = \frac{\sqrt{2\pi}}{4}\|\nabla h\|_{\mathrm{ULip}},$$

where the above inequality is obtained by using the formula: for any $x, z_1, z_2 \in \mathbb{R}^k$,

$$\langle D^2 f_h(x)z_1, z_2\rangle = \int_0^{\infty} \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\,dt \int_{\mathbb{R}^k} \langle w, z_2\rangle\langle Dh\bigl(e^{-t}x + \sqrt{1-e^{-2t}}w\bigr), z_1\rangle\,\mu_Z(dw).$$

Therefore, by an immediate application of Theorem 5.5, we conclude the following

Theorem 5.6. *(Cf. [5, Theorem 7].) Let $\{X_1, X_2, \ldots\}$ be a sequence of independent, identically distributed random vectors in $\mathbb{R}^k$. Assume that $\mathbb{E}[X_1] = 0$ and $\mathbb{E}[X_1X_1^{\mathbf{t}}]$ is a non-singular, self-adjoint, and positive-definite $k \times k$ matrix over $\mathbb{R}$, denoted by $\mathbf{A}$, where $X_1^{\mathbf{t}}$ means the transpose of X_1. Then, for any $h \in C^1(\mathbb{R}^k)$ with $\|h\|_{\mathrm{ULip}} < +\infty$ and $\|\nabla h\|_{\mathrm{ULip}} < +\infty$, we have*

$$\left|\mathbb{E}\bigl[h(Z)\bigr] - \mathbb{E}\left[h\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\right)\right]\right| \leqslant \frac{C(X_1,\mathbf{A})}{\sqrt{n}} \max\left\{\|\sqrt{\mathbf{A}}^{-1}\|_{\mathrm{op}}^2\|h\|_{\mathrm{ULip}}, \frac{\sqrt{2\pi}}{4}\|\nabla h\|_{\mathrm{ULip}}\right\},$$

where Z is a random vector in $\mathbb{R}^k$, having multivariate normal distribution with zero mean and the covariance matrix $\mathbf{A}$, and

$$C(X_1,\mathbf{A}) = \left\{\|\mathbf{A}\|_{\mathrm{op}}\bigl\{\mathrm{Tr}(\mathbf{A})\bigl(4\mathbb{E}\bigl[|X_1|^4\bigr] - \|\mathbf{A}\|_{\mathrm{HS}}^2\bigr)\bigr\}^{\frac{1}{2}} + \frac{4}{3}\mathbb{E}\bigl[|X_1|^3\bigr]\right\},$$

in which $\mathrm{Tr}(\mathbf{A}) = \sum_{j=1}^k \mathbf{A}_{j,j}$ and $\|\mathbf{A}\|_{\mathrm{HS}}^2 = \sum_{i=1}^k\sum_{j=1}^k \mathbf{A}_{i,j}^2$ for $\mathbf{A} = [\mathbf{A}_{i,j}]$.

A Berry–Esséen type estimate. Let $\{Z_j\}$ be a sequence of independent and identically distributed real-valued random variables with zero mean. If $\mathbb{E}[Z_j^2] = \sigma^2$ and $\mathbb{E}[|Z_j|^3] = \rho < +\infty$, then we can apply the celebrated Berry–Esséen theorem to see that for any $a \in \mathbb{R}$,

$$\left|\mathcal{P}\left(\left\{\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^n Z_j \leqslant a\right\}\right) - \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^a e^{-\frac{t^2}{2\sigma^2}}\,dt\right| \leqslant \frac{0.4784\rho}{\sigma^3}\frac{1}{\sqrt{n}}.$$

As another application of Theorem 5.5, we will establish an analogous of Berry–Esséen type estimate for Gaussian approximation in abstract Wiener spaces. More precisely, we will obtain an explicit uniform error bound for

$$\left|\mathcal{P}\left(\left\{\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\right|_{-2} \leqslant a\right\}\right) - \mu_Z\bigl(\{x \in B;\ |x|_{-2} \leqslant a\}\bigr)\right|.$$

To this end, we proceed along with the idea in proving the Wasserstein bounds imply bounds with respect to the Kolmogorov distance for univariate normal approximation. See e.g. the proof of [3, Theorem 3.1 in p. 13]. Fix a decreasing sequence $\{\delta_n\}$ which approaches zero as n tends to infinity, and will be specified later. Let ω be the cap-shaped function on $\mathbb{R}$, i.e.,

$$\omega(t) = k_\omega e^{-\frac{1}{1-t^2}} \mathbf{1}_{(-1,1)}(t), \quad t \in \mathbb{R},$$

where $k_\omega = 1/\int_{-1}^{1} e^{-\frac{1}{1-s^2}}\, ds$. For any $n \in \mathbb{N}$, set $\omega_n(t) = \frac{3}{\delta_n}\omega(\frac{3t}{\delta_n})$, $t \in \mathbb{R}$, and $g_n = \phi_n * \omega_n$, where $*$ is the convolution of functions, and

$$\phi_n(t) = \begin{cases} 1, & \text{if } t \leqslant \frac{1}{3}\delta_n, \\ 0, & \text{if } t \geqslant \frac{2}{3}\delta_n, \\ -\frac{3}{\delta_n}t + 2, & \text{if } \frac{1}{3}\delta_n \leqslant t \leqslant \frac{2}{3}\delta_n. \end{cases}$$

It is clear that all of g_n are infinitely differentiable functions on $\mathbb{R}$ with values in $[0, 1]$. For any $a \in \mathbb{R}$, set $g_{n,a}(t) = g_n(t - a)$, $t \in \mathbb{R}$. Then $\frac{d^r g_{n,a}}{dt^r} = 0$ at $t \in [a + \delta_n, +\infty) \cup (-\infty, a]$, and $\|\frac{d^r g_{n,a}}{dt^r}\|_\infty \leqslant \|\frac{d^r \omega_n}{dt^r}\|_\infty = (\frac{3}{\delta_n})^{r+1}\|\frac{d^r \omega}{dt^r}\|_\infty$ for any $r \in \mathbb{N}$, where $\|\cdot\|_\infty$ denotes the supremum norm. Define $h_{n,a}(x) = g_{n,a}(|x|_{-2})$, $x \in B$, for any $n \in \mathbb{N}$ and $a \in \mathbb{R}$.

Notations. For any Gâteaux differentiable function f on B and $x, y \in B$, let $\delta f(x; y)$ denote the Gâteaux derivative of f at x in the direction y; and by induction on $r \in \mathbb{N}$, for $y_1, \ldots, y_{r+1} \in B$, $\delta^{r+1} f(x; y_1, \ldots, y_{r+1})$ is defined by $\delta(\delta^r f(\cdot; y_1, \ldots, y_r))(x; y_{r+1})$.

Let $n \in \mathbb{N}$, $a \geqslant 0$, and $r \in \mathbb{N}$. By a direct computation and using the fact $\frac{d^r g_{n,a}}{dt^r}|_{t=0} = 0$, $\delta^r h_{n,a}(x, y_1, \ldots, y_r)$ exists for any $x \in B$ and $y_1, \ldots, y_r \in B$. Moreover, we have the following Formulas (1)–(3):

Formula (1).

$$\delta h_{n,a}\big(\phi_x(r,0,0); y\big) = \begin{cases} \frac{dg_{n,a}}{dt}|_{t=|\phi_x(r,0,0)|_{-2}} \frac{A_y^{r,0,0}}{|\phi_x(r,0,0)|_{-2}}, & \text{if } \phi_x(r,0,0) \neq 0, \\ 0, & \text{if } \phi_x(r,0,0) = 0. \end{cases}$$

Formula (2).

$$\delta^2 h_{n,a}\big(\phi_x(r,s,0); y, z\big) = \begin{cases} \frac{dg_{n,a}}{dt}|_{t=|\phi_x(r,s,0)|_{-2}} \frac{\langle y,z\rangle_{-2}|\phi_x(r,s,0)|_{-2}^2 - A_y^{r,s,0}A_z^{r,s,0}}{|\phi_x(r,s,0)|_{-2}^3} \\ \quad + \frac{d^2 g_{n,a}}{dt^2}|_{t=|\phi_x(r,s,0)|_{-2}} \frac{A_y^{r,s,0}A_z^{r,s,0}}{|\phi_x(r,s,0)|_{-2}^2}, & \text{if } \phi_x(r,s,0) \neq 0, \\ 0, & \text{if } \phi_x(r,s,0) = 0. \end{cases}$$

Formula (3).

$$\delta^3 h_{n,a}\big(\phi_x(r,s,u); y,z,w\big)$$
$$= \begin{cases} \left(\frac{d^2 g_{n,a}}{dt^2}\big|_{t=|\phi_x(r,s,u)|_{-2}} - \frac{\frac{dg_{n,a}}{dt}|_{t=|\phi_x(r,s,u)|_{-2}}}{|\phi_x(r,s,u)|_{-2}}\right) \\ \quad \times \frac{\{\langle y,z\rangle_{-2} A_w^{r,s,u} + \langle z,w\rangle_{-2} A_y^{r,s,u} + \langle w,y\rangle_{-2} A_z^{r,s,u}\}|\phi_x(r,s,u)|_{-2}^2 - 3A_y^{r,s,u}A_z^{r,s,u}A_w^{r,s,u}}{|\phi_x(r,s,u)|_{-2}^4} \\ \quad + \frac{d^3 g_{n,a}}{dt^3}\big|_{t=|\phi_x(r,s,u)|_{-2}} \frac{A_y^{r,s,u}A_z^{r,s,u}A_w^{r,s,u}}{|\phi_x(r,s,u)|_{-2}^3}, & \text{if } \phi_x(r,s,u) \neq 0, \\ 0, & \text{if } \phi_x(r,s,u) = 0, \end{cases}$$

where $\phi_x(r,s,u) = x + ry + sz + uw$, $A_y^{r,s,u} = \langle \phi_x(r,s,u), y\rangle_{-2}$, $A_z^{r,s,u} = \langle \phi_x(r,s,u), z\rangle_{-2}$, and $A_w^{r,s,u} = \langle \phi_x(r,s,u), w\rangle_{-2}$ for any $(r,s,u) \in \mathbb{R}^3$.

Proposition 5.7. *Fix $n \in \mathbb{N}$ and $a \geqslant 0$. Then $h_{n,a} \in \mathcal{U}Lip\text{-}1(B)$ and*

$$\|h_{n,a}\|_{\mathrm{ULip}} \leqslant \frac{9}{\delta_n^2} C_{B,H_{-2}} \left\| \frac{d\omega}{dt} \right\|_\infty, \tag{5.19}$$

where $C_{B,H_{-2}}$ is the operator norm of the canonical embedding of B into H_{-2}. In addition, let $f_{h_{n,a}}$ be defined as in (4.2). *Then $f_{h_{n,a}}$ is twice Fréchet differentiable on B, satisfying the following uniformly Lipschitz conditions: For any $x_1, x_2 \in B$,*

$$\big|f_{h_{n,a}}^{(1)}(x_1) - f_{h_{n,a}}^{(1)}(x_2)\big|_2 \leqslant \frac{81}{2\delta_n^3} \left\| \frac{d^2\omega}{dt^2} \right\|_\infty |x_1 - x_2|_{-2}; \tag{5.20}$$

$$\big\|f_{h_{n,a}}^{(2)}(x_1) - f_{h_{n,a}}^{(2)}(x_2)\big\|_{B,B^*} \leqslant \frac{270}{\delta_n^4} C_{B,H_{-2}}^3 \left\| \frac{d^3\omega}{dt^3} \right\|_\infty \|x_1 - x_2\|, \tag{5.21}$$

where $f_{h_{n,a}}^{(r)}(x)$ denotes the r-th order Fréchet derivative of $f_{h_{n,a}}$ at $x \in B$.

Proof. First, it is straightforward by Formulas (1)–(3) that $\delta^r h_{n,a}(x)$, $r = 1,2,3$ and $x \in B$, is r-linear on B with

$$\big\|\delta h_{n,a}(x;\cdot)\big\|_{B^*} \leqslant \frac{9}{\delta_n^2} C_{B,H_{-2}} \left\| \frac{d\omega}{dt} \right\|_\infty, \tag{5.22}$$

$$\big\|\delta^2 h_{n,a}(x;\cdot,\cdot)\big\|_{B,B^*} \leqslant \frac{81}{\delta_n^3} C_{B,H_{-2}}^2 \left\| \frac{d^2\omega}{dt^2} \right\|_\infty, \tag{5.23}$$

$$\big|\delta^2 h_{n,a}(x;\cdot,\cdot)\big|_{H_{-2},H_{-2}^*} \leqslant \frac{81}{\delta_n^3} \left\| \frac{d^2\omega}{dt^2} \right\|_\infty, \tag{5.24}$$

$$\big\|\delta^3 h_{n,a}(x;\cdot,\cdot,\cdot)\big\|_{B,L(B,B^*)} \leqslant \frac{810}{\delta_n^4} C_{B,H_{-2}}^3 \left\| \frac{d^3\omega}{dt^3} \right\|_\infty, \tag{5.25}$$

where $L(B, B^*)$ is the Banach space of all bounded linear operators from B into B^* with the operator norm, and the last three estimates are obtained by using the following two inequalities:

$$\sup_{u\in\mathbb{R}}\left|\frac{\frac{d^r g_{n,a}(u)}{du^r}}{u}\right| \leqslant \left\|\frac{d^{r+1} g_{n,a}(u)}{du^{r+1}}\right\|_\infty \quad (r = 1, 2);$$

$$\sup_{u\in\mathbb{R}}\left|\frac{\frac{dg_{n,a}(u)}{du}}{u^2}\right| = \sup_{u\in\mathbb{R}}\left|\frac{1}{u^2}\int_0^u (u-s)\frac{d^3 g_{n,a}(s)}{ds^3}\,ds\right| \leqslant \frac{\|\frac{d^3 g_{n,a}}{du^3}\|_\infty}{2}.$$

Let $x_1, x_2 \in B$ be arbitrarily given. Then, by (5.22) and the mean-value theorem for differentiation, there is $\vartheta \in (0, 1)$ such that

$$\begin{aligned}\big|h_{n,a}(x_1) - h_{n,a}(x_2)\big| &= \big|\delta h_{n,a}\big(x_2 + \vartheta(x_1 - x_2); x_1 - x_2\big)\big| \\ &\leqslant \frac{9}{\delta_n^2} C_{B,H_{-2}}\left\|\frac{d\omega}{dt}\right\|_\infty \|x_1 - x_2\|,\end{aligned}$$

which implies (5.19). Next, by (5.22), (5.23), and applying the mean-value theorem for differentiation, $f_{h_{n,a}}$ is a real-valued Fréchet differentiable function on B. Moreover, for any $x, z \in B$,

$$\big(z, f_{h_{n,a}}^{(1)}(x)\big)_{B,B^*} = \int_0^\infty \int_B \delta h_{n,a}\big(e^{-t}x + \sqrt{1-e^{-2t}}\,y; e^{-t}z\big)\,\mu_Z(dy)\,dt,$$

and then by (5.24) and applying the mean-value theorem for differentiation,

$$\big|\big(z, f_{h_{n,a}}^{(1)}(x_1) - f_{h_{n,a}}^{(1)}(x_2)\big)_{B,B^*}\big| \leqslant \frac{81}{\delta_n^3}\left\|\frac{d^2\omega}{dt^2}\right\|_\infty |x_1 - x_2|_{-2}|z|_{-2}\int_0^\infty e^{-2t}\,dt,$$

which implies (5.20). Similarly, by (5.23), (5.25), and applying the mean-value theorem for differentiation, $f_{h_{n,a}}^{(1)}$ is a B^*-valued Fréchet differentiable function on B, and, for any $x, z_1, z_2 \in B$,

$$f_{h_{n,a}}^{(2)}(x)(z_1, z_2) = \int_0^\infty \int_B \delta h_{n,a}^2\big(e^{-t}x + \sqrt{1-e^{-2t}}\,y; e^{-t}z_1, e^{-t}z_2\big)\,\mu_Z(dy)\,dt.$$

Then, by (5.25) and applying the mean-value theorem for differentiation,

$$\big|\big(f_{h_{n,a}}^{(2)}(x_1) - f_{h_{n,a}}^{(2)}(x_2)\big)(z_1, z_2)\big| \leqslant \frac{810}{\delta_n^4} C_{B,H_{-2}}^3\left\|\frac{d^3\omega}{dt^3}\right\|_\infty \|x_1 - x_2\|\|z_1\|\|z_2\|\int_0^\infty e^{-3t}\,dt,$$

which implies (5.21). The proof is complete. □

Let $a \geqslant 0$ and $n \in \mathbb{N}$. Observe that

$$\mathbb{E}\big[h_{n,a}(W_n)\big] = \Bigg(\int_{\{|W_n|_{-2} \leqslant a\}} + \int_{\{a<|W_n|_{-2}<a+\delta_n\}} + \int_{\{|W_n|_{-2} \geqslant a+\delta_n\}}\Bigg) h_{n,a}(W_n)\, d\mathcal{P}$$
$$= \mathcal{P}\big(\big\{|W_n|_{-2} \leqslant a\big\}\big) + \int_{\{a<|W_n|_{-2}<a+\delta_n\}} h_{n,a}(W_n)\, d\mathcal{P},$$

as well as

$$\int_B h_{n,a}(x)\, \mu_Z(dx) = \mu_Z\big(\big\{|x|_{-2} \leqslant a\big\}\big) + \int_{\{a<|x|_{-2}<a+\delta_n\}} h_{n,a}(x)\, \mu_Z(dx).$$

If $\mathcal{P}(\{|W_n|_{-2} \leqslant a\}) - \mu_Z(\{x \in B;\ |x|_{-2} \leqslant a\}) \geqslant 0$, then we have

$$\begin{aligned} &\big|\mathcal{P}\big(\big\{|W_n|_{-2} \leqslant a\big\}\big) - \mu_Z\big(\big\{x \in B;\ |x|_{-2} \leqslant a\big\}\big)\big| \\ &\quad \leqslant \bigg|\mathbb{E}\big[h_{n,a}(W_n)\big] - \int_B h_{n,a}(x)\, \mu_Z(dx)\bigg| \\ &\qquad + \int_{\{a-\delta_n<|x|_{-2}<a+\delta_n\}} h_{n,a}(x)\, \mu_Z(dx). \end{aligned} \tag{5.26}$$

If $\mathcal{P}(\{|W_n|_{-2} \leqslant a\}) - \mu_Z(\{x \in B;\ |x|_{-2} \leqslant a\}) < 0$, it follows from

$$\mathbb{E}\big[1 - h_{n,a-\delta_n}(W_n)\big] = \mathcal{P}\big(\big\{|W_n|_{-2} > a\big\}\big) + \int_{\{a-\delta_n \leqslant |W_n|_{-2} \leqslant a\}} \big(1 - h_{n,a-\delta_n}(W_n)\big)\, d\mathcal{P},$$

and

$$\int_B \big(1 - h_{n,a-\delta_n}(x)\big)\, \mu_Z(dx) = \mu_Z\big(\big\{|x|_{-2} > a\big\}\big) + \int_{\{a-\delta_n \leqslant |x|_{-2} \leqslant a\}} \big(1 - h_{n,a-\delta_n}(x)\big)\, \mu_Z(dx),$$

that

$$\begin{aligned} &\big|\mathcal{P}\big(\big\{|W_n|_{-2} \leqslant a\big\}\big) - \mu_Z\big(\big\{x \in B;\ |x|_{-2} \leqslant a\big\}\big)\big| \\ &\quad = \mathcal{P}\big(\big\{|W_n|_{-2} > a\big\}\big) - \mu_Z\big(\big\{x \in B;\ |x|_{-2} > a\big\}\big) \\ &\quad \leqslant \bigg|\mathbb{E}\big[h_{n,a-\delta_n}(W_n)\big] - \int_B h_{n,a-\delta_n}(x)\, \mu_Z(dx)\bigg| \\ &\qquad + \int_{\{a-\delta_n \leqslant |x|_{-2} \leqslant a+\delta_n\}} \big(1 - h_{n,a-\delta_n}(x)\big)\, \mu_Z(dx). \end{aligned} \tag{5.27}$$

Combining (5.26) with (5.27) yields that

$$\begin{aligned}
&\left|\mathcal{P}\left(\{|W_n|_{-2} \leqslant a\}\right) - \mu_Z\left(\{x \in B;\ |x|_{-2} \leqslant a\}\right)\right| \\
&\quad\leqslant \sup_{b\in\mathbb{R}}\left|\mathbb{E}\left[h_{n,b}(W_n)\right] - \int_B h_{n,b}(x)\,\mu_Z(dx)\right| + \mu_Z\left(\{x \in B;\ |x|_{-2} \in [a-\delta_n, a+\delta_n]\}\right) \\
&\quad\leqslant \sup_{b\geqslant 0}\left|\mathbb{E}\left[h_{n,b}(W_n)\right] - \int_B h_{n,b}(x)\,\mu_Z(dx)\right| \\
&\qquad + 2\mu_Z\left(\{x \in B;\ |x|_{-2} \in [0, \delta_n]\}\right) + \mu_Z\left(\{x \in B;\ |x|_{-2} \in [a-\delta_n, a+\delta_n]\}\right),
\end{aligned} \tag{5.28}$$

where the last inequality is obtained by the fact $0 \leqslant h_{n,b} \leqslant h_{n,0} \leqslant 1$ for any $b < 0$. For any $E \in \mathcal{B}(H_{-2})$, set $\mu_Z^{H_{-2}}(E) = \mu_Z(E \cap B)$. Then $\mu_Z^{H_{-2}}$ is a Gaussian measure on H_{-2} with mean zero, and, for any $n \in \mathbb{N}$ and $\alpha < \beta$ in $\mathbb{R}$,

$$\mu_Z^{H_{-2}}\left(\{x \in H_{-2};\ |x|_{-2} \in [\alpha, \beta]\}\right) = \mu_Z\left(\{x \in B;\ |x|_{-2} \in [\alpha, \beta]\}\right).$$

By [16, Lemma 2.1], there exists a constant $C_{\mu_Z} > 0$ such that

$$\mu_Z\left(\{x \in B;\ |x|_{-2} \in [\alpha, \beta]\}\right) \leqslant C_{\mu_Z}(\beta - \alpha), \quad \forall \alpha < \beta,$$

whence, by applying such an inequality to (5.28), we see that for any $a \geqslant 0$,

$$\begin{aligned}
&\left|\mathcal{P}\left(\{|W_n|_{-2} \leqslant a\}\right) - \mu_Z\left(\{x \in B;\ |x|_{-2} \leqslant a\}\right)\right| \\
&\quad\leqslant \sup_{b\geqslant 0}\left|\mathbb{E}\left[h_{n,b}(W_n)\right] - \int_B h_{n,b}(x)\,\mu_Z(dx)\right| + 4C_{\mu_Z}\delta_n.
\end{aligned} \tag{5.29}$$

Set $\delta_n = \frac{1}{\sqrt[10]{n}}$. By applying Theorem 5.5 and Proposition 5.7 to (5.29), we conclude the following

Theorem 5.8. *Let $\{X_1, X_2, \ldots\}$ be a sequence of independent, identically distributed B-valued random variables with finite first moment on the probability space $(\Omega, \mathcal{P})$. Suppose that X_1 has zero mean and the distribution μ_{X_1} has the same covariance function as μ_Z. Then, for any $n \in \mathbb{N}$ and $a \geqslant 0$,*

$$\left|\mathcal{P}\left(\left\{\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\right|_{-2} \leqslant a\right\}\right) - \mu_Z\left(\{x \in B;\ |x|_{-2} \leqslant a\}\right)\right| \leqslant M(Z, X_1)\frac{1}{\sqrt[10]{n}},$$

where $M(Z, X_1)$ is a constant depending only on Z and X_1. In fact,

$$\begin{aligned}
M(Z, X_1) = \max&\left\{270 C_{B,H_{-2}}^3\left\|\frac{d^3\omega}{dt^3}\right\|_\infty, \frac{81}{2}\left\|\frac{d^2\omega}{dt^2}\right\|_\infty\right\} \\
&\times\left\{\|S_{\mu_Z^K}\|_{K,K}\sqrt{\mathrm{Tr}_K(S_{\mu_Z^K})\left(4\mathbb{E}\left[|X_1|_{-1}^4\right] - \|S_{\mu_Z^K}\|_{\mathrm{HS}(K)}^2\right)} + \frac{4}{3}\mathbb{E}\left[\|X_1\|^3\right]\right\} + 4C_{\mu_Z}.
\end{aligned}$$

Remark 5.9. Note that the covariance operator of $\mu_Z^{H_{-2}}$ is $S^2_{\mu_Z^K}$. From the proof of [16, Lemma 2.1], such a constant C_{μ_Z} can be taken to be

$$\frac{1}{\pi}\left(2+\frac{1}{k_{j_1}k_{j_2}k_{j_3}}\right)\left(1+\|S_{\mu_Z^K}\|^2_{\mathrm{HS}(K)}\right),$$

where $k_{j_1}, k_{j_2}, k_{j_3}$ are any three eigenvalues of $S_{\mu_Z^K}$.

References

[1] A.D. Barbour, Stein's method and Poisson process convergence, J. Appl. Probab. 25A (1988) 175–184.
[2] A.D. Barbour, Stein's method for diffusion approximation, Probab. Theory Related Fields 84 (1990) 297–322.
[3] A.D. Barbour, L.H.Y. Chen, An Introduction to Stein's Method, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., vol. 4, Singapore Univ. Press, World Scientific, Singapore, 2005.
[4] V.I. Bogachev, Gaussian Measures, AMS Math. Surveys Monogr., vol. 62, Amer. Math. Soc., Providence, RI, 1998.
[5] S. Chatterjee, E. Meckes, Multivariate approximation using exchangeable pairs, ALEA Lat. Am. J. Probab. Math. Stat. 4 (2008) 257–283.
[6] L.H.Y. Chen, Poisson approximation for dependent trials, Ann. Probab. 3 (1975) 534–545.
[7] R.M. Dudley, J. Feldman, L. LeCam, On semi-norms and probabilities, and abstract Wiener spaces, Ann. Math. 93 (1971) 390–408.
[8] L. Goldstein, Y. Rinott, Multivariate normal approximation by Stein's method and size bias coupling, J. Appl. Probab. 33 (1996) 1–17.
[9] V. Goodman, Quasi-differentiable functions on Banach spaces, Proc. Amer. Math. Soc. 30 (1971) 367–370.
[10] F. Götze, On the rate of convergence in the multivariate CLT, Ann. Probab. 19 (1991) 724–739.
[11] L. Gross, Abstract Wiener spaces, in: Proc. 5th Berkeley Symp. Math. Statist. Probab., vol. 2, 1965, pp. 31–42.
[12] L. Gross, Potential theory on Hilbert space, J. Funct. Anal. 1 (1967) 123–181.
[13] K. Itô, The topological support of Gauss measure on Hilbert space, Nagoya Math. J. 38 (1970) 181–183.
[14] K. Itô, Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces, CBMS–NSF Regional Conf. Ser. in Appl. Math., vol. 47, SIAM, Philadelphia, 1984.
[15] J. Kuelbs, Gaussian measures on Banach spaces, J. Funct. Anal. 5 (1970) 354–367.
[16] J. Kuelbs, T. Kurtz, Berry–Esséen estimates in Hilbert space and an application to the law of the iterated logarithm, Ann. Probab. 2 (1974) 387–407.
[17] H.-H. Kuo, Gaussian Measures in Banach Spaces, Lecture Notes in Math., vol. 463, Springer-Verlag, Berlin/New York, 1975.
[18] H.-H. Kuo, Y.-J. Lee, Integration by parts formula and Stein lemma on abstract Wiener space, Commun. Stoch. Anal. (2011), in press.
[19] Y.-J. Lee, Sharp inequalities and regularity of heat semigroup on infinite dimensional spaces, J. Funct. Anal. 71 (1987) 69–87.
[20] E. Meckes, On Stein's method for multivariate normal approximation, in: High Dimensional Probability V: The Luminy Volume, in: Inst. Math. Stat. Collect., vol. 5, 2009, pp. 153–178.
[21] I. Nourdin, G. Peccati, Stein's method on Wiener chaos, Probab. Theory Related Fields 145 (1) (2009) 75–118.
[22] I. Nourdin, G. Peccati, A. Réveillac, Multivariate normal approximation using Stein's method and Malliavin calculus, Ann. Inst. Henri Poincaré Probab. Stat. 46 (1) (2010) 45–58.
[23] I. Nourdin, G. Peccati, G. Reinert, Second order Poincaré inequalities and CLTs on Wiener space, J. Funct. Anal. 257 (2009) 593–609.
[24] N. Obata, White Noise Calculus and Fock Space, Lecture Notes in Math., vol. 1577, Springer-Verlag, Berlin/Heidelberg, 1994.
[25] M.A. Piech, Parabolic equations associated with the number operator, Trans. Amer. Math. Soc. 194 (1974) 213–222.
[26] M.A. Piech, The Ornstein–Uhlenbeck semigroup in an infinite dimensional L^2 setting, J. Funct. Anal. 18 (1975) 271–285.
[27] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Encyclopedia Math. Appl., vol. 44, Cambridge University Press, Cambridge, 1992.
[28] R. Schatten, Norm Ideals of Completely Continuous Operators, Springer-Verlag, Berlin/Göttingen/Heidelberg, 1960.

[29] W. Schoutens, Stochastic Processes and Orthogonal Polynomials, Lecture Notes in Statist., vol. 146, Springer-Verlag, Berlin/Heidelberg, 2000.
[30] W. Schoutens, Orthogonal polynomials in Stein's method, J. Math. Anal. Appl. 253 (2001) 515–531.
[31] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in: Proc. Sixth Berkeley Symp. Math. Statist. Probab., vol. 2, Univ. California Press, Berkeley, CA, 1972, pp. 583–602.
[32] C. Stein, Approximation Computation of Expectations, IMS Lecture Notes Monogr. Ser., vol. 7, Institute of Mathematical Statistics, Hayward, CA, 1986.
[33] K.R. Stromberg, Probability for Analysts, Chapman & Hall, New York/London, 1994.
[34] Y. Yamasaki, Measures on Infinite Dimensional Spaces, Ser. Pure Math., vol. V, World Scientific, Singapore, 1985.