Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 125 (2015) 182–216
stochastic processes and their applications
www.elsevier.com/locate/spa

Nourdin–Peccati analysis on Wiener and Wiener–Poisson space for general distributions

Richard Eden[1], Juan Víquez*

Department of Mathematics, Purdue University, 150 N. University St., West Lafayette, IN 47907-2067, USA

Received 1 March 2012; received in revised form 1 September 2014; accepted 1 September 2014
Available online 10 September 2014

Abstract

Given a reference random variable, we study the solution of its Stein equation and obtain universal bounds on its first and second derivatives. We then extend the analysis of Nourdin and Peccati by bounding the Fortet–Mourier and Wasserstein distances from more general random variables such as members of the Exponential and Pearson families. Using these results, we obtain non-central limit theorems, generalizing the ideas applied to their analysis of convergence to Normal random variables. We do these in both Wiener space and the more general Wiener–Poisson space. In the former space, we study conditions for convergence under several particular cases and characterize when two random variables have the same distribution. In the latter space we give sufficient conditions for a sequence of multiple (Wiener–Poisson) integrals to converge to a Normal random variable.

Keywords: Malliavin calculus; Stein's method; Pearson distribution; Convergence in distribution

1. Introduction

Recent years have seen exciting research on combining Stein's method with Malliavin calculus in proving central and non-central limit theorems. The delicate combination of these tools can be attributed to Nourdin and Peccati who intertwined an integration by parts formula from

* Correspondence to: Mathematics Department, Universidad de Costa Rica, Costa Rica. Tel.: +506 2260 6424.
E-mail addresses: reden@ateneo.edu (R. Eden), viquezejin@gmail.com (J. Víquez).
[1] Richard Eden is now working at the Mathematics Department, Ateneo de Manila University, Philippines.

http://dx.doi.org/10.1016/j.spa.2014.09.001
0304-4149/

Malliavin calculus with an ordinary differential equation called a Stein equation. Much work has been done to compare Normal or Gamma random variables (r.v.'s) with another r.v. (having unknown distribution). See [12,13,19,20] for results on the convergence of multiple (Wiener) integrals to a standard Normal or Gamma law. [3,26] discuss Cramer's theorem for Normal and Gamma distributions applied to multiple integrals. [28] gives probability tail bounds in terms of the Normal probability tail, with [8] applying the same techniques to give tail bounds in terms of the probability tail of other r.v.'s (e.g. Pearson distributions).

In [15], Nourdin and Peccati found a clever link between Stein's method and Malliavin calculus. This was used to derive the Nourdin–Peccati upper bound (NP bound) on the Wasserstein, Total Variation, Fortet–Mourier and Kolmogorov distances of a generic r.v. from a Normal r.v., and lay the groundwork for comparisons to a more general r.v. (with such results leading to non-central limit theorems). These authors and Reinert (see [16]) applied this NP bound to obtain a second order Poincaré-type inequality useful in proving central limit theorems (CLTs) in Wiener space. Specifically, they proved CLTs for linear functionals of Gaussian subordinated fields. Particular instances are when the subordinated process is fractional Brownian motion (fBm) or the solution to the Ornstein–Uhlenbeck (O–U) stochastic differential equations (SDE) driven by fBm. They also characterized convergence in distribution to a Normal r.v. for multiple stochastic integrals.

Later in [21] these ideas were applied to prove the NP bound in Poisson space (pure jump processes), which was used to obtain Berry–Esséen bounds for arbitrary tensor powers of O–U kernels. Keeping in line with attempts to extend these results as far as possible, [29] proved an NP bound in Wiener–Poisson space. The author applied similar ideas found in [16] to derive a second order Poincaré-type inequality and use it to prove CLTs for a continuous average of a product of two O–U processes (one in Wiener space and the other in Poisson space) which lives in the second chaos of Wiener–Poisson space. Also, it was proved that under mild conditions, the small jump part of a functional in the first Poisson chaos is approximately equal in law to a functional in the first Wiener chaos with the same kernel (useful when simulating a fractional Lévy process as a process with finitely many jumps plus a fBm). All these results show the importance of this NP bound and the potential it has as an effective tool in proving non-central limit theorems, CLTs and characterizations.

Let Z be absolutely continuous with respect to Lebesgue measure with known density. Typical instances are when Z is Normal, Gamma, or another member of the Pearson family of distributions. X is another r.v. whose properties are not as easy to determine as with Z, our "target" r.v. We may have a hunch that X has the same distribution as Z, or in the case of sequences, a belief that $\{X_n\}$ converges in law to the distribution of Z. We thus want to compare X with Z. How different are the laws of X and Z for instance (and we need to make precise the sense in which they are different)? What conditions will ensure that X has the same law as Z? For a sequence $\{X_n\}$, what sufficient conditions ensure convergence to Z in distribution? In this regard, we wish to measure the distance between (the laws of) X and Z by a metric $d_{\mathcal{H}}$ which induces a topology that is equal to or stronger than the topology of convergence in distribution: if $d_{\mathcal{H}}(X_n, Z) \to 0$, then $X_n \to Z$ in distribution.

The motivation for this paper is to find the widest generalization of the NP bound by applying it to a target r.v. which is neither Normal nor Gamma, and in both Wiener space and Wiener–Poisson space. This is worked out in [10] but the conditions needed to apply the NP bound are quite restrictive (it was also carried out only in Wiener space). The conditions we are introducing here are more general, and are still wide enough in scope to cover a Z belonging to the Exponential family or the Pearson family. We point out that Wiener–Poisson space is more inclusive than Wiener space (which can be identified with a subspace of the former). In fact, it includes processes with jumps, and therefore considers Poisson space too as a subspace (also by identification). Nevertheless, even if Wiener space is less general, we can apply our techniques to a wider class of target r.v. Z than in Wiener–Poisson space (which requires boundedness of the second derivative of the solution of Stein's equation, something not needed in Wiener space).

Our main results are the NP bounds on $d_{\mathcal{H}}(X, Z)$ in Wiener space and in Wiener–Poisson space. The main result in Wiener space (Theorem 13) is

$$\begin{aligned} d_{\mathcal{H}}(X, Z) &\le k\mathbb{E}\left|g_*(X) - g_X\right| \\ &\le k\sqrt{\left|\mathbb{E}\left[g_*(X)^2\right] - \mathbb{E}\left[g_*(Z)^2\right]\right| + \left|\mathbb{E}\left[XG_*(X)\right] - \mathbb{E}\left[ZG_*(Z)\right]\right| + \left|\mathbb{E}\left[g_X^2\right] - \mathbb{E}\left[g_Z^2\right]\right|}. \end{aligned}$$

The main result in Wiener–Poisson space (Theorem 25) is

$$d_{\mathcal{H}}(X, Z) \le k\left(\mathbb{E}\left|g_*(X) - g_X\right| + \mathbb{E}\left[\left\langle \left|x\,(DX)^2\right|, \left|-DL^{-1}X\right|\right\rangle_{\mathfrak{H}}\right]\right).$$

Here, $g_X := \mathbb{E}[\langle DX, -DL^{-1}X\rangle_{\mathfrak{H}}|X]$ is a random variable defined using Malliavin calculus operators, specifically, the Malliavin derivative D and the inverse of the infinitesimal generator L of the O–U semigroup. It would be helpful to think of g_X as an object belonging exclusively to X. On the other hand, $g_*(\cdot)$ is a function whose support is the support of Z, taking on nonnegative numbers as values and $g_Z := g_*(Z)$. It will depend only on the density of Z, and is independent of the structure of X. As such, it is an object belonging solely to Z. In the second term of the first bound above, G_* is an antiderivative of g_*, provided it exists. If Z has the necessary (Malliavin) differentiability properties, $g_*(\cdot)$ actually coincides with $\mathbb{E}[\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z = \cdot]$ (P_Z-a.s.), thus explaining the choice of notation g_Z for $g_*(Z)$ which is similar to g_X. This also allows us to make sense of the NP bounds in the following way: if we want to know how different the laws of X and Z are, then we need to know how different (in the L^1 sense) $g_X = \mathbb{E}[\langle DX, -DL^{-1}X\rangle_{\mathfrak{H}}|X]$ and $g_*(X) = \mathbb{E}[\langle DZ, -DL^{-1}Z\rangle_{\mathfrak{H}}|Z = X]$ are. In Wiener–Poisson space, we consider in addition how close the jump part $\mathbb{E}[\langle|x\,(DX)^2\,|, |-DL^{-1}X|\rangle_{\mathfrak{H}}]$ is to 0, which makes sense since Z belongs to Wiener space (subspace of the Wiener–Poisson space without jumps).

In our bounds above, k is a constant that does not depend on X but on Z and the metric we are using. For convergence problems, we do not need its specific value since the convergence will follow from the convergence of $\mathbb{E}\left|g_*(X_n) - g_{X_n}\right|$ to 0. This presupposes we have such a constant k. This constant appears as a bound ($\|\phi\|_\infty \le k$) for some function ϕ, which is related to the solution of the underlying Stein equation. In particular, since we have a Stein equation for each Z (the target r.v.), k depends on Z. Finding such a bound k is easy when Z is Normal: g_* is constant, and consequently, the Stein equation is simpler. If g_* vanishes at a finite endpoint of the support of Z, the challenge now is to find a bound for $\|\phi/g_*\|_\infty$. To the best of our knowledge, [10] (Kusuoka and Tudor) presents the first attempt to find such a sup norm bound when Z is not Normal. Their result is presented below as Lemma 7. In Theorem 9 we improve their result, and this paves the way for the needed bound we stated for the general non-Normal case.

The paper is organized as follows. In Section 2, we review the operators we need from Malliavin calculus. We also define the functions g_* and G_* as well as the random variables g_X and g_Z, studying carefully their properties (needed in the subsequent sections). Section 3 contains preliminaries on Stein's method. Here we find universal bounds on the first and second derivatives of the solution of the general Stein equation. Our main result in Wiener space is in Section 4, where we give a tractable upper inequality which is easier to compute. We also characterize when the law of X is the same as that of Z. Said result is applied to specific cases when g_* is a polynomial and when $\{X_n\}$ is a sequence of multiple integrals. As an example, we prove the convergence of a bilinear functional of a Gaussian subordinated field to a χ^2 r.v. by computing some moments and showing their convergence to desired values. In Section 5, we extend the main result to the more general Wiener–Poisson space. Here, we work out some sufficient conditions for convergence to a Normal law and convergence of the fourth moment.

2. Elements of Malliavin calculus and tools

For the sake of completeness, we include here a brief survey of the needed Malliavin calculus objects. The r.v. $\langle DX, -DL^{-1}X\rangle_{\mathfrak{H}}$ is a key element that bridges Stein's method and Malliavin calculus. D is the Malliavin derivative operator and L is the generator of the Ornstein–Uhlenbeck semigroup.

2.1. Wiener space

Nualart presents in Chapter 1 of [18] a very good exposition on Malliavin calculus in Wiener space. We mention here the elements that we need. Let $\mathfrak{H}$ be a real separable Hilbert space. Assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ over which $W = \{W(h) : h \in \mathfrak{H}\}$ is an isonormal Gaussian process. By definition, this means W is a centered Gaussian family such that $\mathbb{E}[W(h_1)W(h_2)] = \langle h_1, h_2\rangle_{\mathfrak{H}}$. We may also assume that $\mathcal{F}$ is the σ-field generated by W. The white noise case is when $\mathfrak{H} = L^2(T, \mathcal{B}, \mu)$ where $(T, \mathcal{B})$ is a measurable space and μ is a σ-finite atomless measure. The Gaussian process W is then characterized by the family of r.v.'s $\{W(A) : A \in \mathcal{B}, \mu(A) < \infty\}$ where $W(A) = W(\mathbf{1}_A)$. We can then think of W as an $L^2(\Omega, \mathcal{F}, \mathbb{P})$ random measure on $(T, \mathcal{B})$. This is called the white noise measure based on μ. An important example is when $T = [0, \infty)$ and μ is Lebesgue measure. In this case, if we write $W_t = W\left(\mathbf{1}_{[0,t]}\right)$ for $t \geq 0$, then $\{W_t\}_{t\geq 0}$ is a standard Brownian motion embedded in our isonormal Gaussian process.

The qth Hermite polynomial H_q is given by $H_q(x) = (-1)^q e^{x^2/2} \frac{d^q}{dx^q}\left(e^{-x^2/2}\right)$ for $q \geq 1$ and $H_0(x) = 1$. The qth Wiener chaos $\mathcal{H}_q$ is defined as the subspace of $L^2(\Omega) = L^2(\Omega, \mathcal{F}, \mathbb{P})$ generated by the r.v.'s $\left\{H_q(W(h)) : h \in \mathfrak{H}, \|h\|_{\mathfrak{H}} = 1\right\}$. In the white noise case $\mathfrak{H} = L^2_\mu([0, 1])$, each Wiener chaos consists of iterated multiple (Wiener) integrals

$$I_q(f) := q! \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{q-1}} f\left(t_1, t_2, \ldots, t_q\right) dW_{t_q} \cdots dW_{t_2} dW_{t_1}$$

with respect to W, where $f \in \mathfrak{H}^{\odot q}$ is a symmetric nonrandom kernel. When f is nonsymmetric, we let $\widetilde{f}$ denote its symmetrization, and $I_q(f) = I_q(\widetilde{f})$.

All elements of $\mathcal{H}_1$ are Gaussian and all elements of $\mathcal{H}_0$ are deterministic. It is well-known that $L^2(\Omega)$ can be decomposed into an infinite orthogonal sum of the Wiener chaoses, i.e. $L^2(\Omega) = \oplus_{q=0}^{\infty} \mathcal{H}_q$. In the white noise case, any $F \in L^2(\Omega)$ admits a Wiener chaos decomposition of multiple integrals

$$F = \sum_{q=0}^{\infty} I_q\left(f_q\right) \tag{1}$$

where each symmetric $f_q \in \mathfrak{H}^{\odot q} = L^2_\mu(T^q)$ is uniquely determined by F. Note that $I_0(f_0) = f_0 = \mathbb{E}[F]$ and $\mathbb{E}\left[I_q\left(f_q\right)\right] = 0$ for $q \geq 1$.

Consider an orthonormal system $\{e_k : k \geq 1\}$ in $\mathfrak{H}$. For $f \in \mathfrak{H}^{\otimes p}$ and $g \in \mathfrak{H}^{\otimes q}$, the contraction of order $r \leq \min\{p, q\}$ is the element $f \otimes_r g \in \mathfrak{H}^{\otimes(p+q-2r)}$ defined by

$$f \otimes_r g = \sum_{i_1,\ldots,i_r}^{\infty} \left\langle f, e_{i_1} \otimes \cdots \otimes e_{i_r} \right\rangle_{\mathfrak{H}^{\otimes r}} \left\langle g, e_{i_1} \otimes \cdots \otimes e_{i_r} \right\rangle_{\mathfrak{H}^{\otimes r}}.$$

Even if f and g are symmetric, $f \otimes_r g$ may be nonsymmetric so we denote its symmetrization by $f \widetilde{\otimes}_r g$. In the white noise case $\mathfrak{H} = L^2_\mu(T)$, the contraction is given by integrating out r variables. Thus, if $f \in L^2_\mu(T^p)$ and $q \in L^2_\mu(T^q)$, we have $f \otimes_r g \in L^2_\mu\left(T^{p+q-2r}\right)$ and

$$\begin{aligned}(f \otimes_r g)\left(t_1, \ldots, t_{p+q-2r}\right) = \int_{T^r} & f\left(t_1, \ldots, t_{p-r}, s_1, \ldots, s_r\right) \\ & \times g\left(t_{p+1}, \ldots, t_{p+q-r}, s_1, \ldots, s_r\right) d\mu(s_1) \cdots d\mu(s_r).\end{aligned}$$

The product of two multiple integrals is

$$I_q(f)\, I_p(g) = \sum_{r=0}^{p\wedge q} r! \binom{p}{r}\binom{q}{r} I_{q+p-2r}(f \otimes_r g). \tag{2}$$

The Malliavin derivative of a random variable $F \in L^2(\Omega)$ is an $\mathfrak{H}$-valued random variable denoted by DF. In the white noise case $\mathfrak{H} = L^2_\mu(T)$, if $F = I_1(f) = \int_T f(t)\, dW_t$, then D maps F to an $L^2_\mu(T)$-valued element: $D_r F = f(r)$ for $r \in T$. In general, if $F \in L^2(\Omega)$ admits the decomposition (1), then

$$D_r F = \sum_{q=1}^{\infty} q I_{q-1}\left(f_q(r, \cdot)\right). \tag{3}$$

It is possible to iterate this definition to obtain a well defined form for D^k. We denote by $\mathbb{D}^{k,p}$ the domain of D^k in $L^p(\Omega)$, that is F is in $\mathbb{D}^{k,p}$ if and only if $\sum_{j=0}^{k} \mathbb{E}\left[\left\|D^j F\right\|^p_{L^p_\mu(T)}\right] < \infty$. In the white noise case, F with the above decomposition is in $\mathbb{D}^{1,2}$ if and only if $\mathbb{E}\left[\|DF\|^2_{L^2_\mu(T)}\right] = \sum_{q=1}^{\infty} q \cdot q! \left\|f_q\right\|^2_{L^2_\mu(T^q)} < \infty$. We use the following notation: $\mathbb{D}^\infty = \cap_{k\geq 1} \cap_{p\geq 1} \mathbb{D}^{k,p}$. D satisfies the chain rule formula: $D(f(F)) = f'(F)\, DF$ when $F \in \mathbb{D}^{1,2}$ and f is continuously differentiable with bounded derivative. One may relax this to f Lipschitz as long as F has an absolutely continuous law. By approximation, it is possible to prove that this chain rule holds also when f' is not bounded, but we require that $F \in \mathbb{D}^\infty$ and f' is continuous with at most polynomial growth.

D has an adjoint, the divergence operator δ, so that if $F \in \operatorname{Dom}\delta \subset L^2(\Omega; \mathfrak{H})$, then $\delta(F) \in L^2(\Omega)$ and $\mathbb{E}[\delta(F)\, G] = \mathbb{E}\left[\langle F, DG\rangle_{\mathfrak{H}}\right]$ for any $G \in \mathbb{D}^{1,2}$. In the white noise case, δ is called the Skorohod integral: for $F \in \operatorname{Dom}\delta \subset L^2_{\mu\times\mathbb{P}}(T \times \Omega)$ with chaos representation $F(t) = \sum_{q=0}^{\infty} I_q\left(f_q(t, \cdot)\right)$ where each $f_q \in L^2_{\mu^{\otimes(q+1)}}$ is symmetric in the last q variables, $\delta(F) = \sum_{q=0}^{\infty} I_{q+1}\left(\widetilde{f_q}\right)$ if $\sum_{q=0}^{\infty} (q+1)! \left\|\widetilde{f_q}\right\|^2_{L^2_{\mu^{\otimes(q+1)}}} < \infty$, i.e. $F \in \operatorname{Dom}\delta$.

One other operator we need, L, acts on F as in (1) in this way: $LF = -\sum_{q=1}^{\infty} q I_q\left(f_q\right)$. Its domain consists of F for which $\sum_{q=1}^{\infty} q^2 \cdot q! \left\|f_q\right\|^2_{L^2_\mu(T^q)} < \infty$. L also happens to be the infinitesimal generator of the Ornstein–Uhlenbeck semigroup T_t, defined by $T_t F = \sum_{q=0}^{\infty} e^{-qt} I_q\left(f_q\right)$.

One important relation is $\delta DF = -LF$. More than L, we need its pseudo-inverse L^{-1} defined by $L^{-1}F = -\sum_{q=1}^{\infty} \frac{1}{q} I_q\left(f_q\right)$. It easily follows that $LL^{-1}F = F - \mathbb{E}[F]$.

2.2. Wiener–Poisson space

Assume a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ over which $\mathcal{L} = \{\mathcal{L}_t\}_{t\geq 0}$ is a Lévy process. By definition, this means $\mathcal{L}$ has stationary and independent increments, is continuous in probability, and $\mathcal{L}_0 = 0$. Suppose $\mathcal{L}$ is càdlàg, centered, and $\mathbb{E}\left[\mathcal{L}_1^2\right] < \infty$. We may also assume $\mathcal{F}$ is generated by $\mathcal{L}$. Let $\mathcal{L}$ have Lévy triplet $\left(0, \sigma^2, \nu\right)$ and thus, Lévy–Itô decomposition $\mathcal{L}_t = \sigma W_t + \iint_{[0,t)\times\mathbb{R}_0} x d\widetilde{N}(s,x)$ where $W = \{W_t\}_{t\geq 0}$ is a standard Brownian motion, $\widetilde{N}$ is the compensated jump measure (defined in terms of ν) and $\mathbb{R}_0 = \mathbb{R} - \{0\}$. See [1,22] for more about Lévy processes.

Consider now the measure μ on $\mathcal{B}\left(\mathbb{R}^+ \times \mathbb{R}\right)$ where $\mathbb{R}^+ = \{t : t \geq 0\}$ and

$$d\mu(t,x) = \sigma^2 dt\delta_0(x) + x^2 dt d\nu(x)\left(1 - \delta_0(x)\right).$$

Analogous to a Gaussian process W being extended to a random measure (which we also denoted by W) in Wiener space, $\mathcal{L}$ can be extended to a random measure M (see [9]) on $\left(\mathbb{R}^+ \times \mathbb{R}, \mathcal{B}\left(\mathbb{R}^+ \times \mathbb{R}\right)\right)$. This is used to construct (in an analogous way to the Itô integral construction) an integral on step functions, and then by linearity and continuity, extended to $L^2_{\mu^{\otimes q}} = L^2\left(\left(\mathbb{R}^+ \times \mathbb{R}\right)^q, \mathcal{B}\left(\mathbb{R}^+ \times \mathbb{R}\right)^q, \mu^{\otimes q}\right)$. We also denote it by I_q. As in Wiener space,

1. $I_q(f) = I_q\left(\widetilde{f}\right)$;
2. I_q is linear;
3. $\mathbb{E}\left[I_q(f) I_p(g)\right] = \mathbf{1}_{\{q=p\}} q! \int_{\left(\mathbb{R}^+\times\mathbb{R}\right)^q} \widetilde{f}\widetilde{g} d\mu^{\otimes q}$.

Thus, when $F = I_q(f)$, $\mathbb{E}[F^2] = \mathbb{E}[I_q(f)^2] = q!\|\widetilde{f}\|^2_{L^{\otimes q}_\mu}$.

Contractions are defined slightly differently. Suppose $f \in L^2_{\mu^{\otimes q}}$ and $g \in L^2_{\mu^{\otimes p}}$. Let $r \leq \min\{q,p\}$ and $s \leq \min\{q,p\} - r$. The contraction $f \otimes_r^s g \in L^2_{\mu^{\otimes(q+p-2r-s)}}$ is defined by integrating out r variables and sharing s of the remaining variables:

$$\left(f \otimes_r^s g\right)(z,u,v) = \left(\prod_{i=1}^{s} x_i\right) \langle f(\cdot,z,u), g(\cdot,z,v)\rangle_{L^2_{\mu^{\otimes r}}}$$

where $z \in \left(\mathbb{R}^+ \times \mathbb{R}_0\right)^s$, $z_i = (t_i, x_i)$, $u \in \left(\mathbb{R}^+ \times \mathbb{R}_0\right)^{q-r-s}$ and $v \in \left(\mathbb{R}^+ \times \mathbb{R}_0\right)^{p-r-s}$. Its symmetrization is $f \widetilde{\otimes}_r^s g$. We need the following product formula later (see [11] for the proof):

$$I_q(f) I_p(g) = \sum_{r=0}^{p\wedge q} \sum_{s=0}^{p\wedge q - r} r!s! \binom{p}{r}\binom{q}{r}\binom{p-r}{s}\binom{q-r}{s} I_{q+p-2r-s}\left(f \otimes_r^s g\right). \quad (4)$$

We may think of this as a more general version of the product formula (2) where we only consider $s = 0$ since there are no jump components to be shared (which appear in the definition of $f \otimes_r^s g$).

We have briefly narrated a setup parallel to what was done in Wiener space. See [24] for a more detailed exposition. This time though, we have only considered $\mathfrak{H} = L^2\big(\mathbb{R}^+ \times \mathbb{R}, \mathcal{B}\left(\mathbb{R}^+ \times \mathbb{R}\right), \mu\big)$ as underlying Hilbert space, with inner product $\langle f, g\rangle_{\mathfrak{H}} = \int_{\mathbb{R}^+\times\mathbb{R}} f(z) g(z) d\mu(z)$. There is as yet no Malliavin calculus theory developed for a more general abstract Hilbert space. While we do not have a chaos decomposition via orthogonal polynomials (like Hermite polynomials in Wiener space; see [7]), we still have a comparable decomposition proved by Itô (Theorem 2, [9]): for $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$,

$$F = \sum_{q=0}^{\infty} I_q\left(f_q\right) \quad \text{where } f_q \in L^2_{\mu^{\otimes q}}. \tag{5}$$

With this decomposition, we can define the Malliavin derivative operator and Skorohod integral operator. Define $\operatorname{Dom} D$ as the set of $F \in L^2(\Omega)$ for which $\sum_{q=1}^{\infty} q q! \left\| f_q \right\|^2_{L^2_{\mu^{\otimes q}}} < \infty$ and

$$D_z F = \sum_{q=1}^{\infty} q I_{q-1}\left(f_q(z, \cdot)\right).$$

It is instructive to consider the derivatives $D_{t,0}$ and D_z where $z = (t, x)$ has $x \neq 0$. This will enable us to better understand the similarities, and where they end, between the Malliavin calculus of Wiener space and that of Wiener–Poisson space. See [24,23] for more details on the following discussion. We consider two spaces on which we can embed $\operatorname{Dom} D$. For $F \in L^2(\Omega)$, we say $F \in \operatorname{Dom} D^0$ iff $\sum_{q=1}^{\infty} q q! \int_{\mathbb{R}^+} \left\| f_q((t, 0), \cdot)\right\|^2_{L^2_{\mu^{\otimes(q-1)}}} dt < \infty$ and $F \in \operatorname{Dom} D^J$ iff $\sum_{q=1}^{\infty} q q! \int_{\mathbb{R}^+ \times \mathbb{R}_0} \left\| f_q(z, \cdot)\right\|^2_{L^2_{\mu^{\otimes(q-1)}}} d\mu(z) < \infty$. In fact, $\operatorname{Dom} D = \operatorname{Dom} D^0 \cap \operatorname{Dom} D^J$. Since W and $\widetilde{N}$ are independent, we can think of Ω as a cross product of the form $\Omega_W \times \Omega_J$ where $\Omega_W = \mathcal{C}\left(\mathbb{R}^+\right)$ and Ω_J consists of the sequences $((t_1, x_1), (t_2, x_2), \ldots) \in \left(\mathbb{R}^+ \times \mathbb{R}_0\right)^{\mathbb{N}}$ (with a few other technical conditions).

- The derivative $D_{t,0}$ can be interpreted as the derivative with respect to the Brownian motion part. In fact, if $\nu = 0$, then $D_{t,0} F = \frac{1}{\sigma} D^W_t F$ where D^W is the classical Malliavin derivative (defined in Wiener space); the $\frac{1}{\sigma}$ comes from the fact that we are differentiating with respect to σW_t and not just W_t. From the isometry $L^2(\Omega) \simeq L^2\left(\Omega_W; L^2(\Omega_J)\right)$, consider $F \in L^2(\Omega)$ as an element of $L^2\left(\Omega_W; L^2(\Omega_J)\right)$. A smooth F then has the form $F = \sum_{i=1}^{n} G_i H_i$ where each G_i is a smooth Brownian random variable and $H_i \in L^2(\Omega_J)$. We can then define D^W by $D^W F = \sum_{i=1}^{n} \left(D^W G_i\right) H_i$, where $D^W G_i$ is the classical Malliavin derivative. It can be shown that this definition can be extended to a subspace $\operatorname{Dom} D^W \subset \operatorname{Dom} D^0$, so that for $F \in \operatorname{Dom} D^W$, as expected,

$$D_{t,0} F = \frac{1}{\sigma} D^W_t F. \tag{6}$$

For functionals of the form $F = f(G, H) \in L^2(\Omega)$ having $G \in \operatorname{Dom} D^W,\ H \in L^2(\Omega_J)$, and such that f is continuously differentiable with bounded partial derivatives in the first variable, we have a chain rule result: $F \in \operatorname{Dom} D^0$ and $D_{t,0} F = \frac{1}{\sigma} \frac{\partial f}{\partial x}(G, H) D^W_t G$. We may loosen the restriction on f to a.e. differentiability if G is absolutely continuous.
- The derivative $D_z,\ z = (t, x)$ with $x \neq 0$, is a difference operator: for $F \in \operatorname{Dom} D^J$

$$D_z F = \frac{F\left(\omega_{t,x}\right) - F(\omega)}{x}$$

where, if $\Psi_z F$ is the right-hand expression, then $\mathbb{E}\left[\int_{\mathbb{R}^+ \times \mathbb{R}_0} (\Psi_z F)^2 d\mu(z)\right] < \infty$. The idea is to introduce a jump of size x at time t which is captured by the realization $\omega_{t,x}$. For $\omega = \left(\omega^W, \omega^J\right)$, we define $\omega_{t,x}$ by simply adding the time–jump pair (t, x) to ω^J. For

$F = f(G, H) \in L^2(\Omega)$ with $G \in L^2(\Omega_J)$, $H \in \operatorname{Dom} D^J$ and f continuous, we have this chain rule result:

$$D_z F = \frac{f\left(G, H\left(\omega_{t,x}\right)\right) - f(G, H(\omega))}{x} = \frac{f(G, xD_z H + H(\omega)) - f(G, H(\omega))}{x}.$$

If f is differentiable, then by the mean value theorem, for some random $\theta_z \in (0, 1)$,

$$D_z F = \frac{\partial f}{\partial y}(G, \theta_z x D_z H + H(\omega)) D_z H.$$

The following unified chain rule will be very useful (see Proposition 2 in [29]): If $F \in \operatorname{Dom} D^W \cap \operatorname{Dom} D^J$, $DF \in L^2_\mu$, $f \in \mathcal{C}^{k-1}$ has a bounded first derivative and $f^{(k-1)}$ is a.e. differentiable, then for $z \in (t, x) \in \mathbb{R}^+ \times \mathbb{R}$,

$$D_z f(F) = \sum_{n=1}^{k-1} \frac{f^{(n)}(F)}{n!} x^{n-1} (D_z F)^n + \int_0^{D_z F} \frac{f^{(k)}(F + xu)}{(k-1)!} x^{k-1} (D_z F - u)^{k-1} du. \quad (7)$$

In the case where $f^{(k-1)}$ is differentiable everywhere, the chain rule is

$$D_z f(F) = \sum_{n=1}^{k-1} \frac{f^{(n)}(F)}{n!} x^{n-1} (D_z F)^n + \frac{f^{(k)}(F + \theta_z x D_z F)}{k!} x^{k-1} (D_z F)^k \quad (8)$$

for some function $\theta_z \in (0, 1)$ for all $z = (t, x) \in \mathbb{R}^+ \times \mathbb{R}$.

We now define the adjoint of D (see [24] again). Suppose $F \in L^2\left(\mathbb{R}^+ \times \mathbb{R} \times \Omega, \mathcal{B}\left(\mathbb{R}^+ \times \mathbb{R}\right) \times \mathcal{F}, \mu \times \mathbb{P}\right)$ with $F(z) = \sum_{q=0}^\infty I_q\left(f_q(z, \cdot)\right)$ where each $f_q \in L^2_{\mu^{\otimes(q+1)}}$ is symmetric in the last q variables. In this case, the Skorohod integral of F is $\delta(F) = \sum_{q=0}^\infty I_{q+1}\left(\widetilde{f}_q\right)$ where $\sum_{q=0}^\infty (q+1)! \left\| \widetilde{f}_q \right\|^2_{L^2_{\mu^{\otimes(q+1)}}} < \infty$, i.e. $F \in \operatorname{Dom} \delta$ (by definition). Furthermore, $\mathbb{E}[\delta(F) G] = \mathbb{E}\left[\langle F, DG \rangle_{L^2_\mu}\right]$ for any $G \in \operatorname{Dom} D$.

Finally, we define as before $L = -\delta D$: for F as in (5), $LF = -\sum_{q=1}^\infty q I_q\left(f_q\right)$. The pseudo-inverse is defined by $L^{-1} F = -\sum_{q=1}^\infty \frac{1}{q} I_q\left(f_q\right)$. We have again $LL^{-1} F = F - \mathbb{E}[F]$.

Remark 1. Write $\vec{z}_q = (z_1, \ldots, z_q)$, with $z_i = (t_i, x_i)$ for all i. Define

$$\widetilde{W} = \left\{ F = \sum_{q=0}^\infty I_q(f_q) \in \operatorname{Dom} D^0 \ : \ f_q \in L^2_{\mu^{\otimes q}}, \text{ and for every } q, \right.$$
$$\left. f_q(\vec{z}_q) = 0 \text{ if } x_i \neq 0 \text{ for some } i \right\}.$$

Notice from the previous discussion that if $f_q(\vec{z}_q) = 0$ because $x_i \neq 0$, then $I_q(f_q)$ coincides with an iterated multiple (Wiener) integral. Therefore, Wiener space can be seen as a subspace of Wiener–Poisson space (similarly for Poisson space as a subspace). Moreover, $\widetilde{W}$ coincides with the subspace $\mathbb{D}^{1,2}$ (through embedding). The relevance of these facts is that if we have a r.v. $F \in \widetilde{W}$, then the chain rule formula and the Malliavin calculus operators are exactly (up to a constant) the same as those in Wiener space (as explained earlier in this subsection). Furthermore, the results (from other papers) in Wiener space can be replicated in $\widetilde{W}$ and so the conclusions will hold in Wiener–Poisson space, but within $\widetilde{W}$. From now on, $\mathbb{D}^{1,2}$ will mean the subspace $\mathbb{D}^{1,2}$ in Wiener space or the respective embedding $\widetilde{W}$ in Wiener–Poisson space.

2.3. The random variables g_X, g_Z and the functions g_*, G_*

From this point on, $\mathfrak{H}$ will be taken as $L^2\left(\mathbb{R}^+ \times \mathbb{R}, \mathcal{B}\left(\mathbb{R}^+ \times \mathbb{R}\right), \mu\right)$ if we are in Wiener–Poisson space. Now suppose F has mean 0. We have the following integration by parts formulas.

- If $F \in \operatorname{Dom} D^W \cap \operatorname{Dom} D^J$ and $f \in \mathcal{C}^1$ with bounded first derivative a.e. differentiable,

$$\mathbb{E}\left[Ff(F)\right] = \mathbb{E}\left[\left\langle -DL^{-1}F, DF\right\rangle_{\mathfrak{H}} f'(F)\right] + \mathbb{E}\left[\left\langle -DL^{-1}F, \int_0^{DF} f''(F+xu)x(DF-u)du\right\rangle_{\mathfrak{H}}\right]. \tag{9}$$

- If $F \in \operatorname{Dom} D^W \cap \operatorname{Dom} D^J$ and f is twice differentiable with bounded first derivative,

$$\mathbb{E}\left[Ff(F)\right] = \mathbb{E}\left[\left\langle -DL^{-1}F, DF\right\rangle_{\mathfrak{H}} f'(F)\right] + \mathbb{E}\left[\left\langle -DL^{-1}F, \frac{f''(F+\theta.xDF)}{2}x(DF)^2\right\rangle_{\mathfrak{H}}\right]. \tag{10}$$

- If $F \in \mathbb{D}^{1,2}$ (see Remark 1) and f is continuously differentiable with bounded derivative (or f is Lipschitz if F has a density),

$$\mathbb{E}\left[Ff(F)\right] = \mathbb{E}\left[\left\langle -DL^{-1}F, DF\right\rangle_{\mathfrak{H}} f'(F)\right]. \tag{11}$$

Remark 2. Notice that the ideas employed to prove that the chain rule ($Df(F) = f'(F)DF$) holds if f is a polynomial and $F \in \mathbb{D}^\infty$, can be reproduced through the relation (6) obtaining the applicability of the chain rule for functionals in a fixed Wiener–Poisson chaos. Therefore, using formula (10), it follows that for $X = I_q(g)$ (in a fixed Wiener–Poisson chaos),

$$\mathbb{E}\left[X^{r+1}\right] = \frac{r}{q}\mathbb{E}\left[X^{r-1}\|DX\|^2_{\mathfrak{H}}\right] + \frac{r(r-1)}{2q}\mathbb{E}\left[\left\langle x(DX)^3, (X+\theta.xDX)^{r-2}\right\rangle_{\mathfrak{H}}\right].$$

These formulas provide the link to the use of Malliavin calculus techniques in solving problems related to Stein's method. Since $F = LL^{-1}F = -\delta DL^{-1}F$, we have

$$\mathbb{E}\left[Ff(F)\right] = \mathbb{E}\left[-\delta DL^{-1}F \cdot f(F)\right] = \mathbb{E}\left[\left\langle -DL^{-1}F, Df(F)\right\rangle_{\mathfrak{H}}\right].$$

A direct application of the chain rule for Wiener–Poisson space, choosing $k = 2$ in (7) and (8), yields (9) and (10) respectively, and an application of the respective chain rule in Wiener space yields (11).

Assumption A. Z has mean 0 and support (l, u) with $-\infty \le l < 0 < u \le \infty$. The density ρ_* of Z is known, and it is continuous in its support. X is either in $\mathbb{D}^{1,2}$ (Wiener space case) or in $\operatorname{Dom} D^W \cap \operatorname{Dom} D^J$ (Wiener–Poisson space case), and it also has mean 0.

Caution: Notice that in the previous subsection we used $x \in \mathbb{R}$ to denote the jump component of $z \in \mathbb{R}^+ \times \mathbb{R}$ in our state space. On the other hand, we are using Z to denote the target r.v. and X the r.v. with unknown distribution. A confusion may arise in the usage of x and X, or z and Z. However, we will stick with current notation for consistency with existing literature. In this regard, we urge the reader to keep in mind that x represents the size of the jump while X is a random variable *not* (directly) related to x. On the other hand, z is a jump (time of the jump, size of the jump) while Z is the target r.v. which has no jumps.

Remark 3. In some results, we will consider instead of X a sequence $\{X_n\}$ of random variables. In this case, we have the same assumptions (and corresponding functionals, defined below) for each X_n. The continuity assumption of the density ρ_* is not strong at all, since general processes like solutions of stochastic differential equations driven by Brownian motion or (under mild conditions) fractional Brownian motion (for example see [2]) have continuous densities.

Define the random variable $g_X = \mathbb{E}\left[\left\langle DX, -DL^{-1}X\right\rangle_{\mathfrak{H}} \middle| X\right]$ for any Malliavin differentiable r.v. X in the domain of L^{-1}. Nourdin and Peccati proved that $g_X \geq 0$ almost surely (Proposition 3.9, [15]). Closely related is the function

$$g_*(z) = \begin{cases} \dfrac{\int_z^u y\rho_*(y)\,dy}{\rho_*(z)} = -\dfrac{\int_l^z y\rho_*(y)\,dy}{\rho_*(z)} & \text{if } z \in (l,u) \\ 0 & \text{if } z \notin (l,u). \end{cases} \tag{12}$$

Let $g_Z = g_*(Z)$. It must be pointed out that $\varphi(z) := \int_z^u y\rho_*(y)\,dy = -\int_l^z y\rho_*(y)\,dy > 0$ for all $z \in (l,u)$. Since ρ_* is (necessarily) bounded (Assumption A), $\varphi(z)/\rho_*(z)$ is strictly positive (inside the support). Furthermore, $g_*(z) > 0$ for every $z \in (l,u)$. Notice that using this definition of g_* we can conclude that

$$\left(g_*(z)\rho_*(z)\right)' = \varphi'(z) = -z\rho_*(z).$$

One can retrieve the density ρ_* given g_* using the following noteworthy density formula Stein [25] proved[2]:

$$\rho_*(z) = \frac{\mathbb{E}|Z|}{2g_*(z)} \exp\left(-\int_0^z \frac{y}{g_*(y)}dy\right). \tag{13}$$

Proposition 4. *g_* necessarily satisfies the following:*

$$\int_l^0 \frac{y}{g_*(y)}dy = -\infty \qquad \int_0^u \frac{y}{g_*(y)}dy = \infty. \tag{14}$$

Proof. With $\varphi(z)$ defined as before, Nourdin and Viens (Theorem 3.1 [17]) showed that

$$\int_0^z \frac{y}{g_*(y)}dy = \ln\frac{\varphi(0)}{\varphi(z)}.$$

Since $\varphi(z) \to 0$ as $z \to u$ and as $z \to l$, the result follows. ■

Conversely, given the density ρ_* of Z, we can compute g_* using (12). Some examples of known distributions with their g_* are given in Table 1. Recall that $g_*(z) = 0$ outside the support.

Remark 5.
- The necessary conditions in Proposition 4 are actually not new. Stein (Lemma VI.3 [25]) has pointed out that these are necessary for a continuous function g_*, strictly positive on an interval (l,u), to correspond to a unique probability density function ρ_* having mean 0, with g_* and ρ_* related by (12) and (13).
- Suppose $g_*(x) = \alpha(x-l)^p$ for some constant $\alpha > 0$ and the support of Z is (l,∞). Then $\int_0^\infty \frac{x}{g_*(x)}dx = \infty$ if and only if $p \leq 2$, and $\int_l^0 \frac{x}{g_*(x)}dx = -\infty$ if and only if $1 \leq p$. Similarly, if $g_*(x) = \alpha(u-x)^q$ over the support $(-\infty,u)$, $1 \leq q \leq 2$ necessarily. Also, if

[2] Nourdin and Viens proved it in the case where $Z \in \mathbb{D}^{1,2}$ in [17].

Table 1
Common distributions with their ρ_* and g_*.

Normal $(-\infty,\infty)$; $\sigma>0$, $C^{-1}=\sqrt{2\pi}\sigma$	
$\rho_*(z)=Ce^{-z^2/(2\sigma^2)}$	$g_*(z)=\sigma^2$
Gamma (l,∞); $l=-rs$, $r>0$, $s>0$, $C^{-1}=s^r\Gamma(r)$	
$\rho_*(z)=C\,(z-l)^{r-1}\,e^{-(z-l)/s}$	$g_*(z)=s\,(z-l)$
$\chi^2(l,\infty)$; $l=-v$, d.f. $v>0$, $C^{-1}=2^{v/2}\Gamma(v/2)$	
$\rho_*(z)=C\,(z-l)^{\frac{v}{2}-1}\,e^{-(z-l)/2}$	$g_*(z)=2\,(z-l)$
Exponential (l,∞); $l=-\frac{1}{\lambda}$, $C=\lambda>0$	
$\rho_*(z)=Ce^{-\lambda(z-l)}$	$g_*(z)=\frac{1}{\lambda}\,(z-l)$
Beta (l,u); $l=-\frac{r}{r+s}$, $r>0$, $s>0$, $u=1+l$, $C^{-1}=\beta(r,s)$	
$\rho_*(z)=C(z-l)^{r-1}(1+l-z)^{s-1}$	$g_*(z)=\frac{1}{r+s}(z-l)(1+l-z)$
Pearson Type IV $(-\infty,\infty)$; $t=-\frac{s}{2(r-1)}$, $r>\frac{3}{2}$	
$\rho_*(z)=C\left(1+(z-t)^2\right)^{-r}e^{s\tan^{-1}(z-t)}$	$g_*(z)=\frac{1}{2(r-1)}\left(1+(z-t)^2\right)$
Student's $T(-\infty,\infty)$; d.f. $v>2$, $C=\frac{\Gamma((v+1)/2)}{\sqrt{v\pi}\,\Gamma(v/2)}$	
$\rho_*(z)=C\left(1+z^2/v\right)^{-(v+1)/2}$	$g_*(z)=\frac{v}{v-1}\left(1+z^2/v\right)$
Inverse Gamma (l,∞); $l=-\frac{s}{r-1}$, $r>3$, $s>0$, $C=\frac{s^{r-1}}{\Gamma(r-1)}$	
$\rho_*(z)=C(z-l)^{-r}e^{-s/(z-l)}$	$g_*(z)=\frac{1}{r-2}(z-l)^2$
Uniform (l,u); $u=-l>0$, $C^{-1}=2u$	
$\rho_*(z)=C$	$g_*(z)=\frac{1}{2}\left(u^2-z^2\right)$
Pareto (l,∞); $r>2$, $l<0$, $C=r(-l)^r(r-1)^r$	
$\rho_*(z)=C(z-rl)^{-(r+1)}$	$g_*(z)=\frac{1}{r-1}(z-l)(z-rl)$
Laplace $(-\infty,\infty)$; $r>0$, $C=r/2$	
$\rho_*(z)=Ce^{-r\lvert z\rvert}$	$g_*(z)=\frac{1}{r^2}(1+r\lvert z\rvert)$
Lognormal (l,∞); $l=-e^{\mu+\sigma^2/2}$, $\sigma>0$, $p(z)=\frac{\ln(z-l)-\mu}{\sigma}$, $C^{-1}=-\sqrt{2\pi}\sigma e^{2\mu}/l$	
$\rho_*(z)=C\exp\left(-\frac{[p(z)+\sigma]^2}{2}\right)$	$g_*(z)=\sigma e^{2\mu}\exp\left(\frac{[p(z)+\sigma]^2}{2}\right)\int_{p(z)-\sigma}^{p(z)}e^{-s^2/2}ds$

$g_*(x)=O(x^p)$ and the support is $(-\infty,\infty)$, then $p\le 2$. If $g_*(x)=\alpha\,(u-x)^q\,(x-l)^p$ over the support (l,u), then $p\ge 1$ and $q\ge 1$ necessarily.

Let $G_*(z)=\int_0^z g_*(y)\,dy$ be the indefinite integral of g_* (assuming $g_*\in L^1(l,u)$). Consider the situations where the chain rule formula on $G_*(z)$ is applicable. If we take $f=G_*$ in (11), then

$$\mathbb{E}\left[g_X g_*(X)\right]=\mathbb{E}\left[G_*(X)\,X\right]. \tag{15}$$

Assumption A′. Along with Assumption A, one of the following conditions is satisfied:

- g_* has at most polynomial growth and $X\in\mathbb{D}^\infty$.
- g_* is a polynomial of degree m and $X\in\mathbb{D}^{1,m+2}$.
- $\|g_*\|_\infty<\infty$ and the support of the law of X is contained in (l,u).

3. Stein's method and the Stein equation

Stein's method is a set of procedures that is often used to measure distances between random variables such as X and Z. More precisely, we are measuring the distance between the laws of X and Z. These distances take the form

$$d_{\mathcal{H}}(X,Z) = \sup_{h\in\mathcal{H}} |\mathbb{E}[h(X)] - \mathbb{E}[h(Z)]| \tag{16}$$

where $\mathcal{H}$ is a suitable family of functions. If we take $\mathcal{H}_W = \{h : \|h\|_L \leq 1\}$ where $\|\cdot\|_L$ is the Lipschitz seminorm, then $d_W = d_{\mathcal{H}_W}$ is called Wasserstein distance. The bounded Wasserstein (Fortet–Mourier) distance corresponds to $\mathcal{H}_{FM} = \{h : \|h\|_L + \|h\|_\infty \leq 1\}$. Clearly, $d_{FM} \leq d_W$. d_{FM} is important because it metrizes convergence in distribution: $d_{FM}(X_n, Z) \to 0$ if and only if $X_n \xrightarrow{\text{Law}} Z$. d_W on the other hand induces a topology stronger than that of convergence in distribution.

Nourdin and Peccati [15] mentioned other useful metrics. We have the Total Variation distance when $\mathcal{H}_{TV} = \{\mathbf{1}_B : B \text{ is Borel}\}$ and the Kolmogorov distance when $\mathcal{H}_K = \left\{\mathbf{1}_{(-\infty,z]} : z \in \mathbb{R}\right\}$. The latter for example is suited for the analysis of probability tails. However, in this paper, we will only consider d_W and d_{FM} as we try to find bounds for $d_{\mathcal{H}}(X,Z)$ by exploiting properties of Lipschitz functions $h \in \mathcal{H}$.

A Stein equation is at the root of Stein's method. Given Z and a test function h, the Stein equation is the differential equation

$$g_*(x) f'(x) - xf(x) = h(x) - \mathbb{E}[h(Z)]. \tag{17}$$

Observe that $f = f_h$ in (23) is a solution. If the law of X is "close" to the law of Z, then we expect $\mathbb{E}[h(X)] - \mathbb{E}[h(Z)]$ to be close to 0, for h belonging to a large class of functions. Consequently, $\mathbb{E}\left[g_*(X) f'(X) - Xf(X)\right]$ would have to be close to 0. In fact, subject to certain technical conditions, the left-hand side of Eq. (17) provides a characterization of the law of Z: $\mathbb{E}\left[g_*(X) f'(X) - Xf(X)\right] = 0$ if and only if $X \overset{\text{Law}}{=} Z$ (in the equation, information about the law of Z is coded in g_*). The following proposition states this result in its precise form. For a quick proof, see Proposition 6.4 in [15]. The first statement is Lemma 1 in [25] by Stein.

Lemma 6 (*Stein's Lemma*)**.**

1. *If f is continuous, piecewise continuously differentiable, and $\mathbb{E}\left[g_*(Z)\left|f'(Z)\right|\right] < \infty$, then*

$$\mathbb{E}\left[g_*(Z) f'(Z) - Zf(Z)\right] = 0. \tag{18}$$

2. *If for every differentiable f, $x \mapsto \left|g_*(x) f'(x)\right| + |xf(x)|$ is bounded and*

$$\mathbb{E}\left[g_*(X) f'(X) - Xf(X)\right] = 0, \tag{19}$$

then $X \overset{\text{Law}}{=} Z$.

Let $\mathcal{H} = \mathcal{H}_{FM}$ or $\mathcal{H} = \mathcal{H}_W$. Using (17) on (16), we have

$$d_{\mathcal{H}}(X,Z) \leq \sup_{h\in\mathcal{H}} \left|\mathbb{E}\left[g_*(X) f_h'(X) - Xf_h(X)\right]\right| \tag{20}$$

where each f_h is the solution given by (23) for the Stein equation, for a corresponding $h \in \mathcal{H}$. Here the integration by parts formulas (9) and (11) allow us to rewrite the term $\mathbb{E}[Xf_h(X)]$ in terms of the derivatives of f_h and the r.v. g_X, as we pointed out before. For instance, in Wiener space,

$$\begin{aligned} d_{\mathcal{H}}(X,Z) &\le \sup_{h\in\mathcal{H}} \left|\mathbb{E}\left[g_*(X) f_h'(X) - g_X f_h'(X)\right]\right| \\ &= \sup_{h\in\mathcal{H}} \left|\mathbb{E}\left[f_h'(X)(g_*(X)-g_X)\right]\right|. \end{aligned} \tag{21}$$

Thus, to ensure that the distance between X and Z is small, $g_*(X)$ should be close to g_X. We also need to have a good control of $f_h'(X)$. One way of addressing this, taking note of Corollary 6.5 in [15], is by assuming a universal bound for $\mathbb{E}\left[f_h'(X)^2\right]$ for all $h \in \mathcal{H}$ since

$$d_{\mathcal{H}}(X,Z) \le \sqrt{\sup_{h\in\mathcal{H}} \mathbb{E}\left[f_h'(X)^2\right]} \times \sqrt{\mathbb{E}\left[(g_*(X)-g_X)^2\right]}. \tag{22}$$

The first factor is intractable since it requires us to consider conditions on X in relation to all solutions f_h. If however we have a uniform bound for f_h', then we can avoid imposing an additional restriction on X. In this case, we only need to worry about how close $g_*(X)$ is to g_X in $L^2(\Omega)$. In fact, such a bound allows us to just consider how close $g_*(X)$ is to g_X in $L^1(\Omega)$. It is then interesting to see how information about the law of Z is contained in its Malliavin derivative. Notice though that this discussion needs to be modified slightly in Wiener–Poisson space, since the integration by parts formula (9) involves also the second derivative. Thus, we need to control (in a uniform way) both the first and second derivatives of the solution of the Stein equation. Due to this extra requirement, as will be seen later, we will not be able to apply our tools to as wide a scope of target r.v. Z, as we would be able to do in Wiener space.

3.1. Bound for f'

The solution of the Stein equation (17) that we are interested in is the function f_h given by (23). We emphasize here that each such solution is determined by a particular function h (of course, the r.v. Z is also used as input). In the sequel, for the sake of brevity, we drop the subscript h from the solution f and its derivatives.

The Normal case in Wiener space:

If Z is standard Normal ($g_*(z) = 1$), the Stein equation is $f'(x) - xf(x) = h(x) - \mathbb{E}[h(Z)]$ and it has solution $f(x) = e^{x^2/2}\int_{-\infty}^{x}[h(y)-\mathbb{E}[h(Z)]]e^{-y^2/2}dy$. Stein proved (Lemma II.3 *in* [25]) *that $\|f'\|_\infty \le 2\|h-\mathbb{E}[h(Z)]\|_\infty$. In fact, $\|f'\|_\infty \le \min\{2\|h-\mathbb{E}[h(Z)]\|_\infty, 4\|h'\|_\infty\}$ (see Lemma* 2.3 [5]). *For $h \in \mathcal{H}_{FM}$, $\|f'\|_\infty \le 4$. It follows from* (21) *that $d_{FM}(X,Z) \le k\mathbb{E}[|1-g_X|] \le k\sqrt{\mathbb{E}\left[(1-g_X)^2\right]}$ with $k = 4$. Similar estimates for $h \in \mathcal{H}_W$ lead to a bound for d_W of the same form but with $k = 1$ (Lemma* 4.2 [4], *Lemma* 1.2 [15]). *How close the law of X is to the standard Normal law depends on how close g_X is to $g_Z = 1$ (in the L^1 sense).*

In the general case, the Stein equation (17) has solution

$$\begin{aligned} f_h(x) &= \frac{1}{g_*(x)\rho_*(x)}\int_l^x [h(y)-m_h]\rho_*(y)\,dy \\ &= \frac{-1}{g_*(x)\rho_*(x)}\int_x^u [h(y)-m_h]\rho_*(y)\,dy \end{aligned} \tag{23}$$

for $x \in (l, u)$, where $m_h := \mathbb{E}[h(Z)]$. If $x \notin (l, u)$ (in case the support is not $\mathbb{R}$), it follows easily from (17) and since $g_*(x) = 0$, that $f_h(x) = -\frac{h(x)-m_h}{x}$. We then see that if $l > -\infty$, by L'Hôpital's rule,

$$\lim_{x\to l^+} f_h(x) = \lim_{x\to l^+} \frac{[h(x) - m_h]\,\rho_*(x)}{-x\rho_*(x)} = \lim_{x\to l^-} f_h(x)$$

so that f_h is continuous at $x = l$ (and similarly, also at $x = u$ if $u < \infty$).

The proof of the bound for f' when Z is Normal can be adapted to find a constant bound for $g_* f'$ in the non-Normal case. If g_* is uniformly bounded below by a positive number, we easily get a uniform bound for f'. Unfortunately, this is not always the case. In Table 1 we can see several examples of target r.v.'s for which g_* can get arbitrarily close to 0 in its support (for example, when Z is Gamma and $g_*(z) = s(z-l)_+$). Kusuoka and Tudor in [10, Proposition 3] proved the following proposition to address this issue. We state it in the following form using notation and assumptions we have set.

Lemma 7. *Suppose we have the following conditions on* g_*.

1. *If* $u < \infty$, *then* $\underline{\lim}_{x\to u}\, g_*(x)/(u-x) > 0$.
2. *If* $l > -\infty$, *then* $\underline{\lim}_{x\to l}\, g_*(x)/(x-l) > 0$.
3. *If* $u = \infty$, *then* $\underline{\lim}_{x\to u}\, g_*(x) > 0$.
4. *If* $l = -\infty$, *then* $\underline{\lim}_{x\to l}\, g_*(x) > 0$.

Then the solution f *of the Stein equation* (17), *for a given test function* h *with* $\|h\|_\infty < \infty$ *and* $\|h'\|_\infty < \infty$, *has derivative bounded as follows:*

$$\|f'\|_\infty \le k\left(\|h\|_\infty + \|h'\|_\infty\right) \tag{24}$$

where the constant k *depends on* Z *alone, and not on* h.

Unfortunately, conditions **1** and **2** are too restrictive. Consider for instance a r.v. Z with support (l, ∞) and $g_*(x) = \alpha(x)(x-l)^p$, where $\alpha(x)$ is uniformly bounded below by some $\alpha_0 > 0$. From Remark 5, $1 \le p \le 2$ necessarily. Among all g_* of this form, Lemma 7 is thus only able to assure the needed boundedness of f' when $p = 1$. For instance, when Z is Inverse Gamma or Lognormal, condition **2** fails (see the corresponding g_* in Table 1). This stresses the need for less restrictive conditions on g_* that would allow us to include these cases and much more. The first requirement in order to achieve this is a good representation of the derivative f'.

Proposition 8. *For* $x \in (l, u)$, *the derivative* f' *of the solution* $f = f_h$ *given in* (23), *of the Stein equation* (17), *is*

$$f'(x) = \frac{1}{g_*^2(x)\,\rho_*(x)} \int_x^u \int_l^x [1 - \Phi(s)]\,\Phi(t)\left[h'(t) - h'(s)\right] dt\,ds$$

where $\Phi(x) = \int_l^x \rho_*(t)dt$ *is the cumulative distribution function of* Z.

Proof. First,

$$\begin{aligned} h(x) - m_h &= \int_l^x [h(x) - h(s)]\,\rho_*(s)\,ds + \int_x^u [h(x) - h(s)]\,\rho_*(s)\,ds \\ &= \int_l^x \left[\int_s^x h'(t)\,dt\right] \rho_*(s)\,ds - \int_x^u \left[\int_x^s h'(t)\,dt\right] \rho_*(s)\,ds \end{aligned}$$

$$= \int_l^x \left[\int_l^t \rho_*(s)\,ds\right] h'(t)\,dt - \int_x^u \left[\int_t^u \rho_*(s)\,ds\right] h'(t)\,dt$$
$$= \int_l^x \Phi(t)\,h'(t)\,dt - \int_x^u [1-\Phi(t)]\,h'(t)\,dt$$

and so, from (23),

$$\begin{aligned}
g_*(x)\,\rho_*(x)\,f(x) &= \int_l^x [h(y)-m_h]\,\rho_*(y)\,dy \\
&= \int_l^x \left[\int_l^y \Phi(t)\,h'(t)\,dt\right]\rho_*(y)\,dy \\
&\quad - \int_l^x \left[\int_y^u [1-\Phi(t)]\,h'(t)\,dt\right]\rho_*(y)\,dy \\
&= \int_l^x \left[\int_t^x \rho_*(y)\,dy\right]\Phi(t)\,h'(t)\,dt \\
&\quad - \int_l^x \left[\int_l^t \rho_*(y)\,dy\right][1-\Phi(t)]\,h'(t)\,dt \\
&\quad - \int_x^u \left[\int_l^x \rho_*(y)\,dy\right][1-\Phi(t)]\,h'(t)\,dt \\
&= \int_l^x [\Phi(x)-\Phi(t)]\,\Phi(t)\,h'(t)\,dt - \int_l^x \Phi(t)\,[1-\Phi(t)]\,h'(t)\,dt \\
&\quad - \int_x^u \Phi(x)\,[1-\Phi(t)]\,h'(t)\,dt.
\end{aligned}$$

Canceling some terms and solving for f,

$$f(x) = -\frac{1-\Phi(x)}{g_*(x)\,\rho_*(x)} \int_l^x \Phi(t)\,h'(t)\,dt - \frac{\Phi(x)}{g_*(x)\,\rho_*(x)} \int_x^u [1-\Phi(t)]\,h'(t)\,dt. \tag{25}$$

Observe that if $x < 0$,

$$0 = \mathbb{E}[Z] = \int_l^x t\rho_*(t)dt + \int_x^u t\rho_*(t)dt \le x\Phi(x) + g_*(x)\rho_*(x)$$

while if $x > 0$,

$$0 = \mathbb{E}[Z] = \int_l^x t\rho_*(t)dt + \int_x^u t\rho_*(t)dt \ge -g_*(x)\rho_*(x) + x[1-\Phi(x)].$$

Therefore, $0 \le -x\Phi(x) \le g_*(x)\rho_*(x) \to 0$ as $x \to l$ and $0 \le x[1-\Phi(x)] \le g_*(x)\rho_*(x) \to 0$ as $x \to u$. When we then integrate by parts,

$$\int_l^x \Phi(t)\,dt = t\,\Phi(t)\Big|_l^x - \int_l^x t\rho_*(t)\,dt = x\,\Phi(x) + g_*(x)\,\rho_*(x) \tag{26}$$

$$\begin{aligned}
\int_x^u [1-\Phi(t)]\,dt &= t\,[1-\Phi(t)]\Big|_x^u + \int_x^u t\rho_*(t)\,dt \\
&= -x\,[1-\Phi(x)] + g_*(x)\,\rho_*(x). \qquad (27)
\end{aligned}$$

Finally, from (17),

$$\begin{aligned}
g_*(x)\, f'(x) &= xf(x) + h(x) - m_h \\
&= \left(-\frac{x\,[1-\Phi(x)]}{g_*(x)\,\rho_*(x)} + 1\right)\int_l^x \Phi(t)\, h'(t)\, dt \\
&\quad - \left(\frac{x\,\Phi(x)}{g_*(x)\,\rho_*(x)} + 1\right)\int_x^u [1-\Phi(t)]\, h'(t)\, dt \\
&= \frac{1}{g_*(x)\,\rho_*(x)}\int_x^u [1-\Phi(s)]\, ds \int_l^x \Phi(t)\, h'(t)\, dt \\
&\quad - \frac{1}{g_*(x)\,\rho_*(x)}\int_l^x \Phi(t)\, dt \int_x^u [1-\Phi(s)]\, h'(s)\, ds
\end{aligned}$$

which leads to the given form of f'. ■

The bound (24) is not directly suited for d_W where we do not have a prescribed bound on $\|h\|_\infty$. A workaround, as pointed out in [10], is that for each $h \in \mathcal{H}_W$, we pass on the analysis to a sequence $\{h_n\}$ converging to h uniformly in every compact set, where $\{h_n\} \subset \left\{h \in C_0^1 : \left\|h'\right\|_\infty \le 1\right\}$. However, with the help of the previous lemma, we can overcome this complication by giving a bound for f' in terms of only $\|h'\|_\infty$. Recall that if h is Lipschitz, it is a.e. differentiable and $\left\|h'\right\|_\infty = \|h\|_L$.[3] Thus, the upper bound obtained here is immediately well suited for all $f \in \mathcal{F}_{FM}$ and for all $f \in \mathcal{F}_W$.

Theorem 9. *If applicable, assume conditions* **3** *and* **4** *from Lemma 7. Suppose there exists a positive function* $\tilde{g} \in \mathcal{C}^1(l,u)$ *such that*

1. $0 < \underline{\lim}_{x\to u}\, g_*(x)/\tilde{g}(x) \le \overline{\lim}_{x\to u}\, g_*(x)/\tilde{g}(x) < \infty$ *and* $\tilde{g}'(u^-) := \lim_{x\to u^-} \tilde{g}'(x) \in \boldsymbol{R}$ *exists.*[4]
2. $0 < \underline{\lim}_{x\to l}\, g_*(x)/\tilde{g}(x) \le \overline{\lim}_{x\to l}\, g_*(x)/\tilde{g}(x) < \infty$ *and* $\tilde{g}'(l^+) := \lim_{x\to l^+} \tilde{g}'(x) \in \boldsymbol{R}$ *exists.*

Then the solution f of the Stein equation (17), for a given test function h with $\left\|h'\right\|_\infty < \infty$, has derivative bounded as follows:

$$\left\|f'\right\|_\infty \le k\left\|h'\right\|_\infty \tag{28}$$

where the constant k depends on Z alone, and not on h.

Proof. If the support is not $\mathbb{R}$, suppose $x \notin (l,u)$ so that $f(x) = -\frac{h(x)-m_h}{x}$, and so $f'(x) = -\frac{h'(x)}{x} + \frac{h(x)-m_h}{x^2}$. The first term of f' is bounded as $\left|-h'(x)/x\right| \le \left\|h'\right\|_\infty/|l|$ if $x < l$ when $l > -\infty$ ($\left\|h'\right\|_\infty/u$ if $x > u$ when $u < \infty$). For the second term,

$$\begin{aligned}
\left|\frac{h(x)-m_h}{x^2}\right| &= \frac{1}{x^2}\left|\int_l^u h(x)\rho_*(y)dy - \int_l^u h(y)\rho_*(y)dy\right| \\
&\le \frac{1}{x^2}\int_l^u |h(x)-h(y)|\,\rho_*(y)dy
\end{aligned}$$

[3] This is shown as follows: if $x \le y$, $|h(y)-h(x)| = \left|\int_x^y h'(z)dz\right| \le \int_x^y \left|h'(z)\right| dz \le \|h'\|_\infty(y-x) = \|h'\|_\infty|y-x|$, implying that $\left\|h'\right\|_\infty \ge \|h\|_L$. And trivially, $\left\|h'\right\|_\infty \le \|h\|_L$.

[4] $\boldsymbol{R}$ stands for the extended real numbers, i.e. $\boldsymbol{R} = [-\infty, \infty]$.

$$\leq \frac{\|h\|_L}{x^2}\int_l^u |x-y|\rho_*(y)dy$$
$$\leq \frac{\|h\|_L}{x^2}(|x|+\mathbb{E}[|Z|]) = \|h\|_L\left(\frac{1}{|x|}+\frac{\mathbb{E}[|Z|]}{x^2}\right).$$

If $x < l$ when $l > -\infty$, the second factor is bounded by $1/|l| + \mathbb{E}[|Z|]/l^2$ (we have a similar bound when $u < \infty$).

Assume now that x is in the support of Z. Note that from Proposition 8,

$$\left|f'(x)\right| \leq \frac{2\left\|h'\right\|_\infty}{g_*^2(x)\,\rho_*(x)}\int_x^u [1-\Phi(s)]\,ds\int_l^x \Phi(t)\,dt. \tag{29}$$

Fix l' and u' s.t. $l < l' < 0 < u' < u$. Since $g_*(x)\rho_*(x)$ is continuous and strictly positive on $[l', u']$, it attains its minimum $m := \inf_{[l',u']} g_*(x)\rho_*(x) > 0$ on this compact set. Also by continuity of the density $M := \sup_{[l',u']}\rho_*(x) < \infty$, and $g_*(x) = \frac{g_*(x)\rho_*(x)}{\rho_*(x)} \geq \frac{m}{M} > 0$ on $\left[l', u'\right]$, so $g_*^2(x)\rho_*(x) \geq \frac{m^2}{M}$. By the continuity and positivity of $I_1(x) := \int_x^u [1-\Phi(s)]\,ds$ and $I_2(x) := \int_l^x \Phi(t)\,dt$ we conclude that $K := \sup_{[l',u']}(I_1(x)\vee I_2(x)) < \infty$. By (29), $|f'(x)| \leq \frac{2MK^2}{m^2}\left\|h'\right\|_\infty$ on $\left[l', u'\right]$.

Since l' and u' were arbitrarily chosen, we only need to prove now that $\overline{\lim}_{x\to l}\left|f'(x)\right| \leq k_1\|h'\|_\infty$ and $\overline{\lim}_{x\to u}\left|f'(x)\right| \leq k_2\|h'\|_\infty$ for some finite constants k_1 and k_2. Due to the symmetry of the arguments it suffices to prove just one of these limits. Suppose l' was chosen small enough so that $\tilde{g} \in \mathcal{C}^1\left(l, l'\right)$, and for some constants $0 < c \leq C < \infty$, $cg_*(x) \leq \tilde{g}(x) \leq Cg_*(x)$ on $\left(l, l'\right)$.

- **Case 1:** $l > -\infty$.

  We show that the limit of the right-hand side of (29) is finite as $x \to l$. Note that in this case, $\int_x^u [1-\Phi(s)]\,ds = g_*(x)\,\rho_*(x) - x\,[1-\Phi(x)] \to |l|$. By L'Hôpital's rule,

$$\begin{aligned}
\overline{\lim_{x\to l}}\left|f'(x)\right| &\leq 2\left\|h'\right\|_\infty |l|\,\overline{\lim_{x\to l}}\,\frac{C\int_l^x \Phi(t)\,dt}{\tilde{g}(x)\,g_*(x)\,\rho_*(x)}\\
&\leq 2\left\|h'\right\|_\infty |l|\,C\,\overline{\lim_{x\to l}}\,\frac{\Phi(x)}{-x\tilde{g}(x)\,\rho_*(x)+\tilde{g}'(x)\,g_*(x)\,\rho_*(x)}\\
&\leq 2\left\|h'\right\|_\infty |l|\,C\,\overline{\lim_{x\to l}}\,\frac{\Phi(x)}{\left[-cx+\tilde{g}'(x)\right]g_*(x)\,\rho_*(x)}\\
&\leq \frac{2\left\|h'\right\|_\infty |l|\,C}{\tilde{g}'\left(l^+\right)-cl}\,\overline{\lim_{x\to l}}\,\frac{\rho_*(x)}{-x\rho_*(x)} = \frac{2\left\|h'\right\|_\infty C}{\tilde{g}'\left(l^+\right)-cl}.
\end{aligned}$$

  Since $\tilde{g}\left(l^+\right) := \lim_{z\to l^+}\tilde{g}(z) = 0$ and $\tilde{g} \geq 0$, we may assume l' is small enough so $\tilde{g}' \geq 0$ on $\left(l, l'\right)$. Consequently, $\tilde{g}'\left(l^+\right) \neq cl < 0$.

- **Case 2:** $l = -\infty$.

  Since $\underline{\lim}_{x\to-\infty}\, g_*(x) > 0$, we may suppose l' is small enough so that for some constant $m_0 > 0$, $g_*(x) \geq m_0$ over $\left(-\infty, l'\right)$.

$$\begin{aligned}\overline{\lim_{x\to-\infty}}\left|f'(x)\right| &\le 2\|h'\|_\infty \overline{\lim_{x\to-\infty}} \frac{\left(g_*(x)\rho_*(x) - x\left[1-\Phi(x)\right]\right)\int_{-\infty}^{x}\Phi(t)\,dt}{g_*^2(x)\rho_*(x)} \\ &\le 2\left\|h'\right\|_\infty \left(\overline{\lim_{x\to-\infty}} \frac{\int_{-\infty}^{x}\Phi(t)\,dt}{m_0} + \overline{\lim_{x\to-\infty}} \frac{-x\int_{-\infty}^{x}\Phi(t)\,dt}{g_*^2(x)\rho_*(x)}\right) \\ &= 2\left\|h'\right\|_\infty \overline{\lim_{x\to-\infty}} \frac{|x|\int_{-\infty}^{x}\Phi(t)dt}{g_*^2(x)\rho_*(x)}.\end{aligned}$$

There are two subcases to consider depending on the behavior of $\tilde{g}(x)$ as $x \to -\infty$. From the continuity of $\tilde{g}$ and the existence of $\tilde{g}'(l^+)$, $L := \lim_{x\to-\infty}\tilde{g}(x)$ necessarily exists. If $L < \infty$, then $\lim_{x\to-\infty}\frac{\tilde{g}(x)}{|x|} = 0$. If $L = \infty$, then by L'Hôpital's rule, $\lim_{x\to-\infty}\frac{\tilde{g}(x)}{|x|} = -\lim_{x\to-\infty}\tilde{g}'(x) = -\tilde{g}'(l^+)$. In either case, $\lim_{x\to-\infty}\frac{\tilde{g}(x)}{|x|}$ exists.

– **Subcase 1:** $\lim_{x\to-\infty}\frac{\tilde{g}(x)}{|x|} = \infty$
  Note that by (26), $\int_{-\infty}^{x}\Phi(t)dt = x\Phi(x) + g_*(x)\rho_*(x) \le g_*(x)\rho_*(x)$ so

$$\frac{|x|\int_{-\infty}^{x}\Phi(t)dt}{g_*^2(x)\rho_*(x)} \le C\frac{|x|g_*(x)\rho_*(x)}{\tilde{g}(x)g_*(x)\rho_*(x)} = C\frac{|x|}{\tilde{g}(x)}.$$

  Therefore

$$\overline{\lim_{x\to-\infty}}\left|f'(x)\right| \le 2\left\|h'\right\|_\infty C\,\overline{\lim_{x\to-\infty}}\frac{|x|}{\tilde{g}(x)} = 0 < \infty.$$

– **Subcase 2:** $\lim_{x\to-\infty}\frac{\tilde{g}(x)}{|x|} < \infty$
  Similarly from (26),

$$\frac{|x|\int_{-\infty}^{x}\Phi(t)dt}{g_*^2(x)\rho_*(x)} \le \frac{\int_{-\infty}^{x}\frac{|x|}{|t|}g_*(t)\rho_*(t)dt}{m_0 g_*(x)\rho_*(x)} \le \frac{\int_{-\infty}^{x}g_*(t)\rho_*(t)dt}{m_0 g_*(x)\rho_*(x)}.$$

  Therefore,

$$\begin{aligned}\overline{\lim_{x\to-\infty}}\left|f'(x)\right| &\le \frac{2\|h'\|_\infty}{m_0}\overline{\lim_{x\to-\infty}}\frac{\int_{-\infty}^{x}g_*(t)\rho_*(t)dt}{g_*(x)\rho_*(x)} \le \frac{2\|h'\|_\infty}{m_0}\overline{\lim_{x\to-\infty}}\frac{g_*(x)\rho_*(x)}{-x\rho_*(x)} \\ &\le \frac{2\|h'\|_\infty}{m_0}\overline{\lim_{x\to-\infty}}\frac{\tilde{g}(x)}{c|x|} < \infty.\end{aligned}$$

The proof that $\overline{\lim}_{x\to u}\left|f'(x)\right| \le k_2\|h'\|_\infty$ for some $k_2 < \infty$ is similar. ■

Note that if g_* is uniformly bounded below in a neighborhood of $l > -\infty$ (or for $u < \infty$) then condition **2** (**1** in the case of u) from Theorem 9 is not required (see discussion before Lemma 7). In the statement of the previous theorem, we can take $\tilde{g} = g_*$ if g_* is continuously differentiable (at least locally $\mathcal{C}^1$ close to the endpoints of the support), and in this case the conditions are trivially met. In other words, if we can check that $g_* \in \mathcal{C}^1(l,u)$ then bound (28) is automatically true (given the existence of $\tilde{g}'(u^-)$ and $\tilde{g}'(l^+)$). These new conditions are met by all r.v.'s in the Exponential family, Pearson family, and practically any other r.v. whose density is $\mathcal{C}^1$ and is strictly positive in its support. If g_* is not continuously differentiable, we can still get the bound but we are required to approximate g_* by a continuously differentiable function $\tilde{g}$ near the endpoints of the support. For example, consider the Laplace distribution where $g_*(x) = \frac{1}{c^2}(1 + c|x|)$ (see Table 1). In this case g_* is differentiable everywhere except at 0. Therefore we can choose $\tilde{g}(x) = g_*(x)$ for all $x \in \left(-\infty, l'\right) \cup \left(u', \infty\right)$ (with $-\infty < l' < 0 < u' < \infty$) and $\tilde{g}(x) = \phi(x)$ on (l', u') where ϕ is a smooth function such that $\tilde{g}$ is differentiable at l' and u'.

Assumption B. We have the following conditions on g_*.

1. For some positive $\tilde{g} \in \mathcal{C}^1(l, u)$,
   (a) $0 < \underline{\lim}_{x\to u}\, g_*(x)/\tilde{g}(x) \le \overline{\lim}_{x\to u}\, g_*(x)/\tilde{g}(x) < \infty$.
   (b) $0 < \underline{\lim}_{x\to l}\, g_*(x)/\tilde{g}(x) \le \overline{\lim}_{x\to l}\, g_*(x)/\tilde{g}(x) < \infty$.
   (c) $\tilde{g}'(l^+)$ and $\tilde{g}'(u^-)$ exist.
2. If $u = \infty$, then $\underline{\lim}_{x\to u}\, g_*(x) > 0$.
3. If $l = -\infty$, then $\underline{\lim}_{x\to l}\, g_*(x) > 0$.

3.2. Bound for f''

We reiterate that when we refer to a solution f of the Stein equation, we mean the solution f_h given by (23). It is in fact determined by a test function h, but we will drop here the subscript h for brevity.

For our convergence in distribution results in Wiener–Poisson space, we need a boundedness result for f''. The existence of f'' demands more conditions on g_* such as differentiability, which is understandable since we are requiring greater regularity in the solution of the Stein equation. In this setting, the existence of f'' will also immediately force most conditions of Theorem 9 to be satisfied. If we want to work with d_W or d_{FM}, we need to consider Lipschitz functions h, and for any such test function, we can only hope for it to be differentiable almost everywhere. Consequently, f'' must be understood in the almost everywhere sense, i.e., f'' is a version of the second derivative of f such that wherever the second derivative does not exist, f'' will have a value of 0.

Before setting out to find a bound, we point out the unfortunate fact that our results here will not apply to as wide a range of target r.v. Z as what happened for the first derivative. More specifically, we will not be able to give a finite bound for $\left|f''(x)\right|$ when $l > -\infty$ or $u < \infty$, as we were able to do for $\left|f'(x)\right|$ in Theorem 9. See the paragraph before Remark 10 for a counterexample: a bounded Lipschitz function h such that if the support of Z is $(l, \infty) \subsetneq \mathbb{R}$, then $f''(x)$ does not tend to a finite limit as $x \to l$. A similar counterexample can be constructed for a r.v. Z with support $(-\infty, u) \subsetneq \mathbb{R}$, or with support $(l, u) \subsetneq \mathbb{R}$.

First, we make preliminary computations on f''. Differentiating (17) gives us the second derivative

$$f''(x) = \frac{x - g_*'(x)}{g_*(x)} f'(x) + \frac{1}{g_*(x)} f(x) + \frac{1}{g_*(x)} h'(x)$$

which, after considering the form of f in Eq. (25) and of f' given in Proposition 8, reduces to

$$f''(x) = \frac{A(x)\int_l^x \Phi(t)\,h'(t)\,dt + B(x)\int_x^u [1-\Phi(s)]\,h'(s)\,ds + g_*^2(x)\,\rho_*(x)\,h'(x)}{g_*^3(x)\,\rho_*(x)} \tag{30}$$

where, with the help of (26) and (27),

$$\begin{aligned} A(x) &= \left(x - g_*'(x)\right)\int_x^u [1-\Phi(s)]\,ds - g_*(x)\,(1-\Phi(x)) \\ &= g_*(x)\,\rho_*(x)\left(x - g_*'(x)\right) - Q(x)\,(1-\Phi(x)) \end{aligned} \tag{31}$$

$$B(x) = -\left(x - g_*'(x)\right)\int_l^x \Phi(t)\,dt - g_*(x)\,\Phi(x)$$
$$= g_*(x)\,\rho_*(x)\left(g_*'(x) - x\right) - Q(x)\Phi(x). \tag{32}$$

Here, we defined for our convenience the function Q as

$$Q(x) = x^2 - xg_*'(x) + g_*(x). \tag{33}$$

Let $d(x) = g_*^3(x)\,\rho_*(x)$ and $n(x) = f''(x)\,d(x)$, the indicated denominator and numerator, respectively, of $f''(x)$. As $x \to l$, both $d(x)$ and $n(x)$ tend to 0. If h' happens to be differentiable, then by L'Hôpital's rule, $\lim_{x\to l} f''(x) = \lim_{x\to l} n'(x)/d'(x)$. It can be shown that

$$A'(x) = \left(2 - g_*''(x)\right)\int_x^u [1 - \Phi(s)]\,ds \tag{34}$$

and $B'(x) = -\left(2 - g_*''(x)\right)\int_l^x \Phi(t)\,dt$. Therefore

$$\begin{aligned}
n'(x) = {} & A'(x)\int_l^x \Phi(t)\,h'(t)\,dt + A(x)\,\Phi(x)\,h'(x) + B'(x)\\
& \times \int_x^u [1 - \Phi(s)]\,h'(s)\,ds - B(x)\,[1 - \Phi(x)]\,h'(x)\\
& + \left[-xg_*(x)\,\rho_*(x) + g_*'(x)\,g_*(x)\,\rho_*(x)\right]h'(x) + g_*^2(x)\,\rho_*(x)\,h''(x)\\
= {} & \left(2 - g_*''(x)\right)\int_x^u [1 - \Phi(s)]\,ds\int_l^x \Phi(t)\,h'(t)\,dt - \left(2 - g_*''(x)\right)\\
& \times \int_l^x \Phi(t)\,dt\int_x^u [1 - \Phi(s)]\,h'(s)\,ds + [A(x)\,\Phi(x) - B(x)\,(1 - \Phi(x))\\
& - \left(x - g_*'(x)\right)g_*(x)\,\rho_*(x)]\,h'(x) + g_*^2(x)\,\rho_*(x)\,h''(x)\\
= {} & \left(2 - g_*''(x)\right)g_*^2(x)\,\rho_*(x)\,f'(x) + 0\cdot h'(x) + g_*^2(x)\,\rho_*(x)\,h''(x)
\end{aligned}$$

and so

$$\begin{aligned}
\lim_{x\to l} f''(x) &= \lim_{x\to l}\frac{\left(2 - g_*''(x)\right)g_*^2(x)\,\rho_*(x)\,f'(x) + g_*^2(x)\,\rho_*(x)\,h''(x)}{\left(2g_*'(x) - x\right)g_*^2(x)\,\rho_*(x)}\\
&= \lim_{x\to l}\frac{2 - g_*''(x)}{2g_*'(x) - x}f'(x) + \lim_{x\to l}\frac{h''(x)}{2g_*'(x) - x}.
\end{aligned}$$

Define the function $h(x) = \frac{4}{3}(x - l)^{3/2}$ on $(l, 0)$, $h(x) = \frac{4}{3}|l|^{3/2}$ on $[0, \infty)$ and $h(x) = 0$ on $(-\infty, l]$. This function is clearly Lipschitz. Note that $h''(x) = \frac{1}{\sqrt{x-l}}$ on $(l, 0)$. We now consider the same assumptions from Theorem 9 and see that $\overline{\lim}_{x\to l}\left|f'(x)\right| \le k\left\|h'\right\|_\infty$ and $\lim_{x\to l}\frac{h''(x)}{2g_*'(x)-x} = \infty$. We have thus found a Lipschitz function h for which $\lim_{x\to l}\left|f''(x)\right| = \infty$.

Remark 10. From the above discussion we cannot expect to have a universal bound on the second derivative of f unless the support of the target r.v. is $(-\infty, \infty)$. This is consistent with the known NP bound in Wiener–Poisson space developed in [29], where Z was Normal and hence had $(-\infty, \infty)$ for support. For the rest of this subsection, we will then assume that $l = -\infty$ and $u = \infty$.

Lemma 11. *Suppose that on some* $(-\infty, -R] \cup [R, \infty)$ *for some constant* $R \geq 0$, g_* *is twice differentiable with* $g_*''(x) < 2$, *and* $\left|\frac{x-g_*'(x)}{Q(x)}\right|$ *is bounded as* $|x| \to \infty$. *Then for some constant* $R' \geq R$, $A(x) \leq 0$ *and* $B(x) \leq 0$ *for all* $x \in (-\infty, -R'] \cup [R', \infty)$.

Proof. Recall the functions A, B and Q in (31)–(33). On $(-\infty, -R] \cup [R, \infty)$, we define the functions

$$r(x) = -\frac{A(x)}{Q(x)} = 1 - \Phi(x) - \frac{x - g_*'(x)}{Q(x)} g_*(x)\rho_*(x)$$
$$s(x) = -\frac{B(x)}{Q(x)} = 1 - r(x) = \Phi(x) + \frac{x - g_*'(x)}{Q(x)} g_*(x)\rho_*(x).$$

Then, using (34) and (27),

$$\begin{aligned}
[Q(x)]^2 r'(x) &= -A'(x)Q(x) + A(x)Q'(x) \\
&= -(2 - g_*''(x)) \int_x^\infty [1 - \Phi(s)]ds\, Q(x) + A(x)(2x - xg_*''(x)) \\
\frac{[Q(x)]^2 r'(x)}{2 - g_*''(x)} &= -Q(x) \int_x^\infty [1 - \Phi(s)]ds + xA(x) \\
&= -Q(x)\left[-x[1 - \Phi(x)] + g_*(x)\rho_*(x)\right] + x\left[g_*(x)\,\rho_*(x)\left(x - g_*'(x)\right)\right. \\
&\quad \left. - Q(x)\,(1 - \Phi(x))\right] \\
&= -Q(x)g_*(x)\rho_*(x) + xg_*(x)\rho_*(x)(x - g_*'(x)) = -g_*^2(x)\rho_*(x).
\end{aligned}$$

Since $g_*''(x) < 2$, then $r'(x) < 0$ for all $x \in [R, \infty)$. As $x \to \infty$, $1 - \Phi(x) \to 0$, $g_*(x)\rho_*(x) \to 0$ and $\left|\frac{x-g_*'(x)}{Q(x)}\right|$ is bounded. Therefore, $r(x) \geq \lim_{x\to\infty} r(x) = 0$ on $[R, \infty)$. Consequently, $\lim_{x\to\infty} s(x) = 1$ so that by the continuity of s, there is some $R_u' \geq R$ such that $s(x) \geq 0$ on $[R_u', \infty)$.

Similar statements can be proved for r and s on $(-\infty, -R]$. The computations above show that $s'(x) = -r'(x) > 0$ on $(-\infty, -R]$ and so $s(x) \geq \lim_{x\to-\infty} s(x) = 0$ for all $x \in (-\infty, -R]$. Because it follows that $\lim_{x\to-\infty} r(x) = 1$, then for some $R_d' \geq R$, $r(x) \geq 0$ on $(-\infty, -R_d']$. If we take $R' = \max\{R_u', R_d'\}$, $r(x) \geq 0$ and $s(x) \geq 0$ on $I := (-\infty, R'] \cup [R', \infty)$. Therefore, for any $x \in I$, $A(x)$ and $B(x)$ have the same sign. Now we show they are both non-positive. Observe that

$$\begin{aligned}
D(x) &:= A(x) \int_{-\infty}^x \Phi(t)dt + B(x) \int_x^\infty [1 - \Phi(s)]ds \\
&= \left[\left(x - g_*'(x)\right) \int_x^\infty [1 - \Phi(s)]\,ds - g_*(x)\,(1 - \Phi(x))\right] \int_{-\infty}^x \Phi(t)\,dt \\
&\quad + \left[-\left(x - g_*'(x)\right) \int_{-\infty}^x \Phi(t)\,dt - g_*(x)\,\Phi(x)\right] \int_x^\infty [1 - \Phi(s)]\,ds \\
&= -g_*(x)\,(1 - \Phi(x)) \int_{-\infty}^x \Phi(t)\,dt - g_*(x)\,\Phi(x) \int_x^\infty [1 - \Phi(s)]\,ds \\
&= -g_*(x)\,(1 - \Phi(x))\,(g_*(x)\,\rho_*(x) + x\,\Phi(x)) \\
&\quad - g_*(x)\,\Phi(x)\,(g_*(x)\,\rho_*(x) - x\,[1 - \Phi(x)]) \\
&= -g_*^2(x)\rho_*(x) \leq 0.
\end{aligned}$$

Therefore, $A(x) \leq 0$ and $B(x) \leq 0$ for all $x \in I$. ■

Theorem 12. *Suppose Assumption B holds, and that on some* $(-\infty, -R] \cup [R, \infty)$ *for some constant* $R \geq 0$, g_* *is twice differentiable with* $g_*''(x) < 2$, *and* $\left|\frac{x-g_*'(x)}{Q(x)}\right|$ *is bounded as* $|x| \to \infty$. *Then the solution* $f = f_h$ *given by (23) of the Stein equation (17), for a given test function* h *with* $\|h'\|_\infty < \infty$, *has second derivative bounded as follows:*

$$\|f''\|_\infty \leq k \|h'\|_\infty \tag{35}$$

where the constant k *depends on* Z *alone, and not on* h.

Proof. Recall the functions A and B in (31) and (32). From the preceding lemma, $A(x) \leq 0$ and $B(x) \leq 0$ for all $x \in I = (-\infty, -R'] \cup [R', \infty)$ for some constant $R' \geq 0$. Therefore, from (30), and using D defined in the proof of the preceding lemma, for all $x \in I$,

$$\begin{aligned}
\left|f''(x)\right| &\leq \frac{-A(x)}{g_*^3(x)\rho_*(x)} \int_{-\infty}^{x} \Phi(t)\,dt \cdot \|h'\|_\infty \\
&\quad + \frac{-B(x)}{g_*^3(x)\rho_*(x)} \int_{x}^{\infty} [1-\Phi(s)]\,ds \cdot \|h'\|_\infty + \frac{\left|h'(x)\right|}{g_*(x)} \\
\frac{g_*^3(x)\rho_*(x)\left|f''(x)\right|}{\|h'\|_\infty} &\leq -A(x)\int_{-\infty}^{x} \Phi(t)\,dt - B(x)\int_{x}^{u}[1-\Phi(s)]\,ds + g_*^2(x)\rho_*(x) \\
&= -D(x) + g_*^2(x)\rho_*(x) = 2g_*^2(x)\rho_*(x).
\end{aligned}$$

Due to the continuity of g_* and conditions of Assumption B when $l = -\infty$ and $u = \infty$, there is some $m_0 > 0$ such that $g_*(x) > m_0$ for all $x \in \mathbb{R}$. Then, $|f''(x)| \leq \frac{2\|h'\|_\infty}{g_*(x)} \leq \frac{2}{m_0}\|h'\|_\infty = k\|h'\|_\infty$ for all $x \in I$. Lastly, for $x \in [-R', R']$, we can glean from (30) the bound

$$|f''(x)| \leq \|h'\|_\infty \frac{|A(x)| \int_{-\infty}^{x} \Phi(t)dt + |B(x)| \int_{x}^{\infty}[1-\Phi(s)]ds + g_*^2(x)\rho_*(x)}{g_*^3(x)\rho_*(x)}.$$

The factor following $\|h'\|_\infty$ is continuous, and thus bounded on $[-R', R']$. This finishes the proof. ■

We point out that Lemma 11 is a more general version of Lemma 7 in [8] (note that Φ there is defined as the upper probability tail). The lemma there prescribed our conditions to hold on $(-\infty, \infty)$, while in Lemma 11, we showed that we could weaken the conditions involving g_* (double differentiability, boundedness) so that these need only hold over some union $(-\infty, -R] \cup [R, \infty)$ (where R may be strictly positive). Lest one think the conditions of Theorem 12 are too restrictive, a closer look will show that they are all satisfied by the g_* of members of the Pearson family having $(-\infty, \infty)$ as its support. Examples are the Pearson Type IV, Normal, and Student's T distributions (see Table 1 to check the conditions). The conditions are also satisfied by the g_* of Laplace distributions, which we note is twice differentiable everywhere except at the origin. We collect these conditions in the following assumption, which we will have need of in Section 5.

Assumption B′. Along with Assumption B, and $l = -\infty$ and $u = \infty$, the following hold.

1. For some constant $R \geq 0$, g_* is twice differentiable and $g_*''(x) < 2$ for all $x \in (-\infty, -R] \cup [R, \infty)$.
2. $\overline{\lim}_{x \to \pm\infty} \left|\frac{x-g_*'(x)}{Q(x)}\right| < \infty$.

4. NP bound in Wiener space

From the results in Section 3.1 all solutions $f = f_h$ given by (23), of the Stein equation, belong to a set $\mathcal{F}_{\mathcal{H}} = \{f \in \mathcal{C}^1(l,u) : \|f'\|_\infty \le k\}$. While the constant k will not depend on the specific test function h used, it may still be driven by general characteristics of the members of $\mathcal{H}$, the family of test functions used. See the beginning of Section 3.1 for different choices of k under various families $\mathcal{H}$, when Z is standard Normal.

Theorem 13 (*NP Bound*). *Let $d_{\mathcal{H}}$ be d_W or d_{FM}. Under Assumptions A and B,*

$$d_{\mathcal{H}}(X,Z) \le k\mathbb{E}\left|g_*(X) - g_X\right| \tag{36}$$

$$\le k\sqrt{\left|\mathbb{E}\left[g_*(X)^2\right] - \mathbb{E}\left[g_*(Z)^2\right]\right| + \left|\mathbb{E}\left[g_*(X)g_X\right] - \mathbb{E}\left[g_*(Z)g_Z\right]\right| + \left|\mathbb{E}\left[g_X^2\right] - \mathbb{E}\left[g_Z^2\right]\right|}. \tag{37}$$

Let $G_*(x)$ be an antiderivative of $g_*(x)$. Under Assumptions A′ and B, with $ZG_*(Z) \in L^1(\Omega)$,

$$d_{\mathcal{H}}(X,Z) \le k\sqrt{\left|\mathbb{E}\left[g_*(X)^2\right] - \mathbb{E}\left[g_*(Z)^2\right]\right| + \left|\mathbb{E}\left[XG_*(X)\right] - \mathbb{E}\left[ZG_*(Z)\right]\right| + \left|\mathbb{E}\left[g_X^2\right] - \mathbb{E}\left[g_Z^2\right]\right|}. \tag{38}$$

In both statements, k is a finite constant depending only on Z and on $d_{\mathcal{H}}$.

Proof. The first bound in (36) follows from (21) and Theorem 9. The second bound follows from Hölder's Inequality. Let $\Delta = \mathbb{E}\left[(g_*(X) - g_X)^2\right]^{1/2}$. Since $(g_*(Z) - g_Z)^2 = 0$ a.s.,

$$\Delta^2 = \mathbb{E}\left[g_*(X)^2\right] - 2\mathbb{E}\left[g_*(X)g_X\right] + \mathbb{E}\left[g_X^2\right] - \left(\mathbb{E}\left[g_*(Z)^2\right] - 2\mathbb{E}\left[g_*(Z)g_Z\right] + \mathbb{E}\left[g_Z^2\right]\right)$$

and (37) follows. From (15) and Assumption A′ we have $\mathbb{E}\left[g_*(F)g_F\right] = \mathbb{E}\left[FG_*(F)\right]$, which proves (38). ■

The first inequality also follows from Theorem 1 and Eq. (19) in Kusuoka and Tudor [10]. The setup in their paper involves functions b and a. The function b is any function for which $\int_l^u b(x)\rho_*(x)\,dx = 0$ along with a few other mild conditions: $b > 0$ near l, $b < 0$ near u, $b\rho_*$ is continuous and bounded on (l,u). They then defined $a(x) = 2\int_l^x b(y)\rho_*(y)\,dy/\rho_*(x)$. Then for W a standard Brownian motion, the SDE

$$dY_t = b(Y_t)\,dt + \sqrt{a(Y_t)}dW_t \tag{39}$$

has a unique Markovian weak solution with invariant density ρ_*. With a and b as given above, from Theorem 1 in [10],

$$d_{\mathcal{H}}(X,Z) \le k\mathbb{E}\left|\frac{a(X)}{2} - \left\langle DX, DL^{-1}\{b(X) - \mathbb{E}b(X)\}\right\rangle_{\mathfrak{H}}\right| + k\left|\mathbb{E}b(X)\right|. \tag{40}$$

If we take $b(x) = -x$, it follows that $a(x) = 2g_*(x)$. If X is centered, the right-hand side of (40) quickly reduces to $k\mathbb{E}\left|g_*(X) - g_X\right|$.

While the results in [10] appear more general, taking $b(x) = -x$ suffices. A careful analysis will reveal that the proofs of their main results depend only on the density ρ_* and the choice of b. While each choice of b arguably yields a different diffusion process Y, the invariant density is still ρ_*. Their analytical proofs are in fact independent of the stochastic differential equation

(39) and the diffusion process arising from it. For the present paper, we only need comparisons with the law of the reference variable Z. To this end, knowing the density ρ_* will suffice. The computations using $b(x) = -x$ and $a(x) = 2g_*(x)$ are much easier and this is reflected in the simplicity of (36) compared to (40).

Furthermore, as shown in the next theorem, the bounds we get from taking $b(x) = -x$ (see Theorem 13) are tight. Indeed, nothing is lost by choosing b this way.

Theorem 14 (*Law Characterization*)**.** *Under Assumptions A′ and B, $X \overset{\text{Law}}{=} Z$ if and only if all of the following are satisfied.*

1. $\mathbb{E}\left[g_*(X)^2\right] = \mathbb{E}\left[g_*(Z)^2\right]$.
2. $\mathbb{E}\left[XG_*(X)\right] = \mathbb{E}\left[ZG_*(Z)\right]$.
3. $\mathbb{E}\left[g_X^2\right] = \mathbb{E}\left[g_Z^2\right]$.

Proof. If the three conditions are satisfied, Theorem 13 implies $d_{\mathcal{H}}(X, Z) = 0$.

Now suppose $X \overset{\text{Law}}{=} Z$. They then have the same density ρ_* so **1** and **2** immediately follow. We next prove that $g_X \overset{\text{Law}}{=} g_Z$, imitating the technique Nourdin and Viens used to prove (12) (see Theorem 3.1 [17]). Let f be a continuous function with compact support, and F any antiderivative of f.

$$\begin{aligned}
\mathbb{E}\left[f(X)\, g_X\right] = \mathbb{E}\left[XF(X)\right] &= \int_l^u \left[x\rho_*(x)\right] F(x)\, dx \\
&= -F(x) \int_x^u y\rho_*(y)\, dy \Big|_{x \to l}^{x \to u} + \int_l^u f(x) \left[\int_x^u y\rho_*(y)\, dy\right] dx \\
&= \int_l^u f(x) \frac{\int_x^u y\rho_*(y)\, dy}{\rho_*(x)} \rho_*(x)\, dx = \mathbb{E}\left[f(X) \frac{\int_X^u y\rho_*(y)\, dy}{\rho_*(X)}\right]
\end{aligned}$$

so $g_X = \int_X^u y\rho_*(y)\, dy / \rho_*(X)$ a.s. This has the same distribution as $\int_Z^u y\rho_*(y)\, dy / \rho_*(Z)$, equal to g_Z a.s., so **3** then follows. ■

Remark 15. We see that $\mathbb{E}\left[g_*(Z)^2\right] = \mathbb{E}\left[ZG_*(Z)\right] = \mathbb{E}\left[g_Z^2\right]$ (by integration by parts formula). Thus, for X to have the same law as Z, it is necessary and sufficient that $\mathbb{E}\left[g_*(X)^2\right]$, $\mathbb{E}\left[XG_*(X)\right]$ and $\mathbb{E}\left[g_X^2\right]$ (which a priori need not be all the same) are all equal to $\mathbb{E}\left[g_Z^2\right]$. The three conditions in Theorem 14 are stated in their current form due to the symmetry involved.

That $\mathbb{E}\left[g_*(Z)^2\right] = \mathbb{E}\left[ZG_*(Z)\right] = \mathbb{E}\left[g_Z^2\right]$ are all equal depends on the specific structure of Z itself, and it is rooted in how g_* (and thus G_* as well) is defined in terms of the law of Z. Specifically, it is because $g_*(Z) = g_Z$ that we are able to use the integration by parts formula (11) on $g_*(Z)$. If we evaluate the function g_* at the random variable X, we cannot expect $g_*(X)$ to be equal to g_X because g_* is an object that "belongs" to Z. However, if X and Z are to be "almost" the same in law, we would expect X to "almost" satisfy the same relations/equations for Z, e.g. $\mathbb{E}\left[g_*(X)^2\right]$ "$=$" $\mathbb{E}\left[XG_*(X)\right]$. If g_* is a polynomial, then this amounts to checking that the moments of X satisfy the same conditions met by the moments of Z. Granted, this method of moments is not sufficient. Hence, the need for condition **3**, $\mathbb{E}\left[g_X^2\right] = \mathbb{E}\left[g_Z^2\right]$, in Theorem 14.

The following versions of Theorems 14 and 13 for sequences are useful.

Corollary 16. *Under Assumptions A (or A′) and B, $X_n \to Z$ in distribution if all of the following are satisfied.*

1. $\mathbb{E}\left[g_*(X_n)^2\right] \to \mathbb{E}\left[g_*(Z)^2\right]$.
2. $\mathbb{E}\left[g_*(X_n)\,g_{X_n}\right] \to \mathbb{E}\left[g_*(Z)\,g_Z\right]$ *(under Assumption A).*
   $\mathbb{E}\left[X_n G_*(X_n)\right] \to \mathbb{E}\left[Z G_*(Z)\right]$ *(under Assumption A′).*
3. $\mathbb{E}\left[g_{X_n}^2\right] \to \mathbb{E}\left[g_Z^2\right]$.

Corollary 17. *Under Assumptions A and B, $X_n \to Z$ in distribution if $g_*(X_n) - g_{X_n} \to 0$ in $L^1(\Omega)$.*

Remark 18. If we normalize so that $\operatorname{Var} X = \operatorname{Var} Z$, condition **3** in Theorem 14 can be replaced by $\operatorname{Var} g_X = \operatorname{Var} g_Z$ since $\mathbb{E}[g_X] = \operatorname{Var} X$. This also allows us to replace the term $\left|\mathbb{E}\left[g_X^2\right] - \mathbb{E}\left[g_Z^2\right]\right|$ in Theorem 13 by $|\operatorname{Var} g_X - \operatorname{Var} g_Z|$. In Corollary 16, we can replace condition **3** by $\operatorname{Var} g_{X_n} \to \operatorname{Var} g_Z$ if $\mathbb{E}\left[X_n^2\right] \to \mathbb{E}\left[Z^2\right]$.

If Z is Normal with variance σ^2 so $g_*(y) = \sigma^2$, $G_*(y) = \sigma^2 y$ and $g_Z = \sigma^2$. If $\operatorname{Var} X = \sigma^2$, then

$$\begin{aligned} d_{\mathcal{H}}(X, Z) &\le k\sqrt{\left|\sigma^4 - \sigma^4\right| + \sigma^2\left|\mathbb{E}\left[X^2\right] - \mathbb{E}\left[Z^2\right]\right| + |\operatorname{Var} g_X - \operatorname{Var} g_Z|} \\ &= k\sqrt{\operatorname{Var} g_X} \end{aligned} \tag{41}$$

where $k = 4$ if $d_{\mathcal{H}} = d_{FM}$ and $k = 1$ if $d_{\mathcal{H}} = d_W$. This retrieves Theorem 3.3 in [14]. If we have a bound on $\operatorname{Var} g_X$, this may be used to bound the distance. A Poincaré-type inequality may be used in this regard. See [16] (also for an explanation of the notation used below) where they use such a bound on $\operatorname{Var} g_X$ to get the following result:

$$d_{\mathcal{H}}(X, Z) \le \frac{k\sqrt{10}}{2\sigma}\left(\mathbb{E}\left[\left\|D^2 X \otimes_1 D^2 X\right\|_{\mathfrak{H}^{\otimes 2}}^2\right]\right)^{1/2}\left(\mathbb{E}\left[\|DX\|_{\mathfrak{H}}^4\right]\right)^{1/2}. \tag{42}$$

This was used in [16,29] to prove CLTs for functionals of Gaussian subordinated fields (applied to fBm and the solution of the O–U SDE driven by fBm, for all $H \in (0, 1)$).

4.1. Convergence when g_* is a polynomial

Many of the common random variables belong to the Pearson family of distributions, all of whose members are characterized by their g_* being polynomials of degree at most 2, i.e. $g_*(y) = \alpha y^2 + \beta y + \gamma$ in the support of Z. Some member distributions in this family are Normal (g_* is constant), Gamma (g_* has degree 1), Beta (g_* is quadratic with positive discriminant), Student's T-distribution (g_* is quadratic with negative discriminant) and Inverse Gamma (g_* is quadratic with zero discriminant).

Refer to [6,25] for more information about Pearson distributions, and [8] for Stein's method applied to comparisons of probability tails with a Pearson Z. From Remark 5, if the support of Z is unbounded and g_* is a polynomial, then Z is necessarily Pearson. If Z has bounded support and g_* is a polynomial, g_* may have degree exceeding 2 and in this case, Z is not Pearson.

Corollary 19. *If g_* is a polynomial $g_*(x) = \sum_{k=0}^{m} a_k x^k$, for the convergence $X_n \to Z$ in distribution, conditions* **1** *and* **2** *in Corollary 16 can be replaced by these conditions (respectively):*

$\mathbb{E}\left[X_n^k\right] \to \mathbb{E}\left[Z^k\right]$ for $k = 1, \ldots, 2m$, and $\mathbb{E}\left[X_n^k g_{X_n}\right] \to \mathbb{E}\left[Z^k g_Z\right]$ for $k = 1, \ldots, m$. Under Assumption A′ *(i.e., $X_n \in \mathbb{D}^{1,m+2}$), the two conditions can be replaced by $\mathbb{E}\left[X_n^k\right] \to \mathbb{E}\left[Z^k\right]$ for $k = 1, \ldots, \max\{2m, m+2\}$.*

Proof. $g_*^2(x)$ has order $2m$ while $xG_*(x)$ has order $m+2$. The matching moments ensure condition **1** in Corollary 16 is satisfied, and under Assumption A′ also condition **2** is fulfilled. ∎

Suppose $g_*(x) = \sum_{k=0}^m a_k x^k$. Note that

$$\mathbb{E}\left[g_*(Z)^2\right] = \mathbb{E}\left[\left(\sum_{k=0}^m a_k Z^k\right)^2\right] = \sum_{k=0}^{2m}\left(\sum_{i=0}^k a_i a_{k-i}\right)\mathbb{E}\left[Z^k\right]$$

while

$$\mathbb{E}\left[ZG_*(Z)\right] = \sum_{k=0}^m \frac{a_k}{k+1}\mathbb{E}\left[Z^{k+2}\right].$$

We noted earlier that $\mathbb{E}\left[g_*(Z)^2\right]$ and $\mathbb{E}\left[ZG_*(Z)\right]$ are equal. While the polynomial coefficients of the different moments of Z are different, and more moments may be involved in one expression compared to the other, the coefficients and the moments themselves should take care of this apparent difference to ensure equality under the expectation.

Suppose Z is Pearson with $g_Z = g_*(Z) = \alpha Z^2 + \beta Z + \gamma$. We can prove the following recursive formula for the moments of Z (see [8], end of Section 5.1): $\mathbb{E}\left[Z^{r+1}\right] = \frac{r\beta}{1-r\alpha}\mathbb{E}\left[Z^r\right] + \frac{r\gamma}{1-r\alpha}\mathbb{E}\left[Z^{r-1}\right]$. Therefore,

$$\mathbb{E}\left[g_Z\right] = \mathbb{E}\left[Z^2\right] = \frac{\gamma}{1-\alpha}$$

$$2\mathbb{E}\left[Zg_Z\right] = \mathbb{E}\left[Z^3\right] = \frac{2\beta\gamma}{(1-\alpha)(1-2\alpha)}$$

$$3\mathbb{E}\left[Z^2 g_Z\right] = \mathbb{E}\left[Z^4\right] = \frac{6\beta^2\gamma + (1-2\alpha)3\gamma^2}{(1-\alpha)(1-2\alpha)(1-3\alpha)}$$

and

$$\mathbb{E}\left[g_Z^2\right] = \frac{\beta^2\gamma(1-\alpha) + \gamma^2(1-2\alpha)^2}{(1-\alpha)(1-2\alpha)(1-3\alpha)} \tag{43}$$

$$\operatorname{Var} g_Z = \mathbb{E}\left[g_*^2(Z)\right] - \left(\mathbb{E}\left[g_*(Z)\right]\right)^2 = \frac{\beta^2\gamma(1-\alpha)^2 + 2\alpha^2\gamma^2(1-2\alpha)}{(1-2\alpha)(1-3\alpha)(1-\alpha)^2}. \tag{44}$$

Corollary 20. *Suppose Z is a Pearson random variable and for the sequence $\{X_n\}$,* $\operatorname{Var} X_n = \mathbb{E}[X_n^2] = \mathbb{E}[g_{X_n}] \to \frac{\gamma}{1-\alpha}$. *The following are sufficient conditions so that $X_n \to Z$ in distribution.*

1. *When Z is Normal ($\alpha = \beta = 0$),* $\operatorname{Var} g_{X_n} \to 0$.
2. *When Z is Gamma ($\alpha = 0$),* $\operatorname{Var} g_{X_n} \to \beta^2\gamma$ *and*
   - *under Assumption* A, $\mathbb{E}\left[X_n g_{X_n}\right] \to \beta\gamma$.
   - *under* Assumption A′, $2\mathbb{E}\left[X_n g_{X_n}\right] = \mathbb{E}\left[X_n^3\right] \to 2\beta\gamma$.

3. *In the general case where* $\alpha \neq 0$, $\operatorname{Var} g_{X_n} \to \frac{\beta^2\gamma(1-\alpha)^2+2\alpha^2\gamma^2(1-2\alpha)}{(1-2\alpha)(1-3\alpha)(1-\alpha)^2}$ *and*
   - *under* Assumption A,
     $2\mathbb{E}\left[X_n g_{X_n}\right], \mathbb{E}\left[X_n^3\right] \to \frac{2\beta\gamma}{(1-\alpha)(1-2\alpha)}$, *and* $3\mathbb{E}\left[X_n^2 g_{X_n}\right], \mathbb{E}\left[X_n^4\right] \to \frac{6\beta^2\gamma+(1-2\alpha)3\gamma^2}{(1-\alpha)(1-2\alpha)(1-3\alpha)}$.
   - *under* Assumption A′,
     $2\mathbb{E}\left[X_n g_{X_n}\right] = \mathbb{E}\left[X_n^3\right] \to \frac{2\beta\gamma}{(1-\alpha)(1-2\alpha)}$, *and* $3\mathbb{E}\left[X_n^2 g_{X_n}\right] = \mathbb{E}\left[X_n^4\right] \to \frac{6\beta^2\gamma+(1-2\alpha)3\gamma^2}{(1-\alpha)(1-2\alpha)(1-3\alpha)}$.

Proof. Apply Corollary 19 directly. ■

The first statement is the version for sequences of Corollary 3.4 in [17]. Alternatively, we could replace $\operatorname{Var} g_{X_n} \to 0$ by $\mathbb{E}\left[g_{X_n}^2\right] \to \gamma^2$. For the Gamma convergence, we can replace $\operatorname{Var} g_{X_n} \to \beta^2\gamma$ by $\mathbb{E}\left[g_{X_n}^2\right] \to \beta^2\gamma + \gamma^2$. When $\alpha \neq 0$, we can work with (43) instead of (44) so the statement will be in terms of $\mathbb{E}\left[g_{X_n}^2\right] \to \frac{\beta^2\gamma(1-\alpha)+\gamma^2(1-2\alpha)^2}{(1-\alpha)(1-2\alpha)(1-3\alpha)}$.

The next result follows from Corollary 17.

Corollary 21. *Suppose* Z *is a Pearson random variable.* $X_n \to Z$ *in distribution if* $g_{X_n} - \alpha X_n^2 - \beta X_n \to \gamma$ *in* $L^1(\Omega)$.

4.2. Convergence in a fixed Wiener chaos

When X is inside a fixed Wiener chaos so $X = I_q(f)$, we have more structure available. For example, $\langle DX, -DL^{-1}X\rangle_{\mathfrak{H}} = \frac{1}{q}\|DX\|_{\mathfrak{H}}^2$. Therefore, if $Z \stackrel{\text{Law}}{=} \mathcal{N}(0,\sigma^2)$ and $\mathbb{E}\left[\left(I_q(f)\right)^2\right] = \sigma^2$, (41) gives us the bound

$$d_{\mathcal{H}}(X, Z) \le k\sqrt{\operatorname{Var} g_X} \le k\sqrt{\operatorname{Var}\left(\frac{1}{q}\|DX\|_{\mathfrak{H}}^2\right)}.$$

One may then use bounds like

$$\operatorname{Var}\left(\frac{1}{q}\|DX\|_{\mathfrak{H}}^2\right) \stackrel{(a)}{=} \frac{1}{q^2}\mathbb{E}\left[\left(\|DX\|_{\mathfrak{H}}^2 - q\sigma^2\right)^2\right] \stackrel{(b)}{\le} \frac{q-1}{3q}\left(\mathbb{E}\left[X^4\right] - 3\sigma^4\right) \tag{45}$$

to further cap the distance. Equality (a) follows from $\mathbb{E}\left[\frac{1}{q}\|DX\|_{\mathfrak{H}}\right] = \mathbb{E}[g_X] = \sigma^2$ and inequality (b) from Lemma 3.5 in [14]. These are quite important and known results which yield CLTs for functionals on a fixed Wiener chaos. For instance, if we have a sequence $\{X_n\} = \{I_q(f_n)\}$ where $\mathbb{E}\left[\left(I_q(f_n)\right)^2\right] \to \sigma^2$, then the following conditions are equivalent:

1. $X_n \to Z$ in distribution;
2. $\mathbb{E}\left[X_n^4\right] \to 3\sigma^4$;
3. $\|f_n \otimes_r f_n\|_{\mathfrak{H}^{\otimes(2q-2r)}} \to 0$ for all $r = 1, \ldots, q-1$;
4. $\|DX_n\|_{\mathfrak{H}}^2 \to q\sigma^2$ in $L^2(\Omega)$;
5. $\left\|D^2X_n \otimes_1 D^2X_n\right\|_{\mathfrak{H}^{\otimes 2}}^2 \to 0$ in $L^2(\Omega)$.

See [20] for (1) $\iff$ (2) $\iff$ (3), [19] for (1) $\iff$ (4), and [16] for (1) $\iff$ (5). These in some sense highlight the tightness of inequality (38) with the help of bounds like (42) and (45).

Corollary 22. *If $X_n = I_q(f_n)$ with $q \geq 1$, then condition* **3** *in Corollary 16 can be replaced by* $\mathbb{E}\left[\|DX_n\|_{\mathfrak{H}}^4\right] \to q^2 \mathbb{E}\left[g_*^2(Z)\right]$.

Proof. This is a direct consequence of $\left\langle DX_n, -DL^{-1}X_n \right\rangle_{\mathfrak{H}} = \frac{1}{q}\|DX_n\|_{\mathfrak{H}}^2$ and $\mathbb{E}\left[g_Z^2\right] = \mathbb{E}\left[g_*^2(Z)\right]$. ■

From this and Corollary 21, we have the following result for the convergence in a fixed Wiener chaos to a Pearson random variable.

Corollary 23. *Let Z be Pearson with $g_*(z) = \alpha z^2 + \beta z + \gamma$ in its support. Fix $q \geq 2$. Suppose $X_n = I_q(f_n)$ and $\mathbb{E}\left[X_n^2\right] \to \frac{\gamma}{1-\alpha}$. If $\|DX_n\|_{\mathfrak{H}}^2 - q\alpha X_n^2 - q\beta X_n \to q\gamma$ in $L^1(\Omega)$, then $X_n \to Z$ in distribution.*

Remark 24. Special cases of the above corollary are known results.

- Let Z be Normal with variance 1, i.e. $g_*(z) = 1$. Suppose $\mathbb{E}\left[X_n^2\right] \to 1$. Then $X_n \to Z$ in distribution if $\|DX_n\|_{\mathfrak{H}}^2 \to q$ in $L^2(\Omega)$. See [19].
- Let Z be Gamma with $g_*(z) = (2z + 2v)_+$, i.e. $\beta = 2$ and $\gamma = 2v$, where the parameters are chosen for consistency with the discussion in [13]. Suppose $\mathbb{E}\left[X_n^2\right] \to 2v$. Then $X_n \to Z$ in distribution if $\|DX_n\|_{\mathfrak{H}}^2 - 2qX_n \to 2qv$ in $L^2(\Omega)$.

The result in the first item of this remark is known as the Nualart–Ortiz-Latorre criterion. In [27], the authors used it to prove that

$$C\sqrt{N}\ln(N)\left(\widehat{H}_N - H\right) \xrightarrow[N\to\infty]{} \mathcal{N}(0,1)$$

where $\widehat{H}_N$ is an estimator of the Hurst parameter H for fBm when $H \in \left(\frac{1}{3}, \frac{1}{2}\right)$ (see [27] for details).

5. NP bound in Wiener–Poisson space

In Wiener–Poisson space, if we repeat the process before Eq. (21) and use (9), the correct integration by parts formula, we get

$$d_{\mathcal{H}}(X,Z) \leq \sup_{f\in\mathfrak{F}_{\mathcal{H}}} \Bigg| \mathbb{E}\left[f'(X)\left(g_*(X) - g_X\right)\right] + \mathbb{E}\left[\left\langle \int_0^{D_zX} f''(X+xu)x(D_zX-u)du, -DL^{-1}X \right\rangle_{\mathfrak{H}}\right] \Bigg|. \tag{46}$$

Here, we remind the reader that $z = (t,x) \in \mathbb{R}^+ \times \mathbb{R}$. It becomes evident that we need to find universal bounds on the first and second derivatives of f. Recall from Section 3.2 that we only have such bounds when $l = -\infty$ and $u = \infty$. With this in mind, we have $\mathfrak{F}_{\mathcal{H}} = \{f \in \mathcal{C}^1 : f' \text{ is Lipschitz}, \ \|f'\|_\infty < k_1, \|f''\|_\infty < k_2\}$, where k_1 and k_2 depend only on the distance $d_{\mathcal{H}}$. The following is a generalization of Theorem 2 in [29] (where Z was standard Normal) and an extension of Theorem 13 to Wiener–Poisson space.

Theorem 25 (*NP Bound*)**.** *Let $d_{\mathcal{H}}$ be d_W or d_{FM}. Under Assumptions A and B′,*

$$d_{\mathcal{H}}(X,Z) \leq k\left(\mathbb{E}\left|g_*(X) - g_X\right| + \mathbb{E}\left[\left\langle \left|x(DX)^2\right|, \left|-DL^{-1}X\right| \right\rangle_{\mathfrak{H}}\right]\right)$$

where k is a finite constant depending only on Z and on $d_{\mathcal{H}}$.

Proof. This follows immediately from (46) since $|\langle a,b\rangle|_{\mathfrak{H}} \leq \langle |a|,|b|\rangle_{\mathfrak{H}}$ and $\left|\int f d\mu\right| \leq \int |f| d\mu$. ■

This upper bound (with Z Normal) was first developed for Poisson space in [21], where it was used to prove several CLTs for Poisson functionals. In [29] it was used to prove CLTs for Wiener–Poisson functionals.

Corollary 26. *Under Assumptions A and B′, $X_n \to Z$ in distribution if both statements are true.*

1. $g_*(X_n) - g_{X_n} \to 0$ *in* $L^1(\Omega)$.
2. $\left\langle \left|x\,(DX_n)^2\right|, \left|-DL^{-1}X_n\right|\right\rangle_{\mathfrak{H}} \to 0$ *in* $L^1(\Omega)$.

The following preliminary computations are needed for Theorem 28.

Proposition 27. *Let $X_n = I_q(f_n)$, with $\mathbb{E}\left[X_n^2\right] = q!\,\|f_n\|^2_{\mathfrak{H}^{\otimes q}} \to 1$. Assume that for $r = 0,\dots,q-1$ and $s = 0,\dots,q-r$, $\left\| f_n \otimes_r^s f_n \right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}} \mathbf{1}_{\{s=0,r\neq 0\}\cup\{s\neq 0,r=0\}} \to 0$. Then as $n\to\infty$,*

1. $\mathbb{E}\left[\|DX_n\|^4_{\mathfrak{H}}\right] \to q^2$;
2. $\|DX_n\|^2_{\mathfrak{H}} \to q$ *in* $L^2(\Omega)$;
3. $\mathbb{E}\left[\int_{\mathbb{R}^+\times\mathbb{R}} x^2\,(D_zX_n)^4\,d\mu(z)\right] \to 0$;
4. $\mathbb{E}\left[X_n^4\right] \to 3$.

Proof. Since $D_zX_n = qI_{q-1}(f_n(z,\cdot))$, we can apply the product formula (4) to get

$$\begin{aligned}
\|DX_n\|^2_{\mathfrak{H}} &= \langle DX_n, DX_n\rangle_{\mathfrak{H}} = q^2\int I_{q-1}(f_n(z,\cdot))\,I_{q-1}(f_n(z,\cdot))\,d\mu(z)\\
&= q^2\int\sum_{r=0}^{q-1}\sum_{s=0}^{q-1-r} r!s!\binom{q-1}{r}^2\binom{q-1-r}{s}^2\\
&\quad\times I_{2q-2-2r-s}\big(f_n(z,\cdot)\otimes_r^s f_n(z,\cdot)\big)d\mu(z)\\
&= q^2\sum_{p=1}^{q}\sum_{s=0}^{q-p}(p-1)!s!\binom{q-1}{p-1}^2\binom{q-p}{s}^2\\
&\quad\times\int I_{2q-2p-s}\left(f_n(z,\cdot)\otimes_{p-1}^s f_n(z,\cdot)\right)d\mu(z)\\
&= \sum_{p=1}^{q}\sum_{s=0}^{q-p} pp!s!\binom{q}{p}^2\binom{q-p}{s}^2 I_{2q-2p-s}\\
&\quad\times\left(\int f_n(z,\cdot)\otimes_{p-1}^s f_n(z,\cdot)\,d\mu(z)\right)\\
&= \sum_{r=1}^{q}\sum_{s=0}^{q-r} rr!s!\binom{q}{r}^2\binom{q-r}{s}^2 I_{2q-2r-s}\left(f_n\otimes_r^s f_n\right).
\end{aligned}$$

Also by orthogonality of chaoses,

$$\begin{aligned}
\mathbb{E}\left[\|DX_n\|^4_{\mathfrak{H}}\right] &= \sum_{r,R=1}^{q}\sum_{s=0}^{q-r}\sum_{S=0}^{q-R} rRr!R!s!S!\binom{q}{r}^2\binom{q}{R}^2\binom{q-r}{s}^2\binom{q-R}{S}^2\\
&\quad\times\mathbb{E}\left[I_{2q-2r-s}\left(f_n\otimes_r^s f_n\right)I_{2q-2R-S}\left(f_n\otimes_R^S f_n\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\substack{r,R=1\\ r\neq R}}^{q} \sum_{s=0}^{q-r}\sum_{S=0}^{q-R} rRr!R!s!S! \binom{q}{r}^2 \binom{q}{R}^2 \binom{q-r}{s}^2 \binom{q-R}{S}^2 \\
&\quad \times \mathbf{1}_{\{2r+s=2R+S\}} (2q-2r-s)! \left\langle f_n \widetilde{\otimes}_r^s f_n, f_n \widetilde{\otimes}_R^S f_n \right\rangle_{\mathfrak{H}^{\otimes(2q-2r-s)}} \\
&\quad + \sum_{r=1}^{q-1}\sum_{s=0}^{q-r} r^2 (r!)^2 (s!)^2 \binom{q}{r}^4 \binom{q-r}{s}^4 (2q-2r-s)! \\
&\quad \times \left\| f_n \widetilde{\otimes}_r^s f_n \right\|^2_{\mathfrak{H}^{\otimes(2q-2r-s)}} + q^2 \left(q! \left\| f_n \right\|^2_{\mathfrak{H}^{\otimes q}} \right)^2 .
\end{aligned}
$$

Since $\|\widetilde{g}\|_{\mathfrak{H}} \leq \|g\|_{\mathfrak{H}}$ for nonsymmetric g (this follows by a simple application of the triangle inequality), then $\left\| f_n \widetilde{\otimes}_r^s f_n \right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}} \leq \left\| f_n \otimes_r^s f_n \right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}}$. Use this fact along with Hölder's inequality in the following:

$$
\begin{aligned}
\left\langle f_n \widetilde{\otimes}_r^s f_n, f_n \widetilde{\otimes}_R^S f_n \right\rangle_{\mathfrak{H}^{\otimes(2q-2r-s)}} &\leq \left\| f_n \widetilde{\otimes}_r^s f_n \right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}} \left\| f_n \widetilde{\otimes}_R^S f_n \right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}} \\
&\leq \left\| f_n \otimes_r^s f_n \right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}} \left\| f_n \otimes_R^S f_n \right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}} .
\end{aligned}
$$

$\mathbb{E}\left[\|DX_n\|^4_{\mathfrak{H}}\right] \to q^2$ then follows from the assumptions on the kernels' contractions, proving the first point.

On the other hand,

$$
\begin{aligned}
\mathbb{E}\left[\left(\|DX_n\|^2_{\mathfrak{H}} - q\right)^2\right] &= \mathbb{E}\left[\|DX_n\|^4_{\mathfrak{H}} - 2q\,\|DX_n\|^2_{\mathfrak{H}} + q^2\right] \\
&= \mathbb{E}\left[\|DX_n\|^4_{\mathfrak{H}}\right] - 2q \cdot q\mathbb{E}\left[X_n^2\right] + q^2 \to 0
\end{aligned}
$$

so $\|DX_n\|^2_{\mathfrak{H}} \to q$ in $L^2(\Omega)$ proving the second point.

For the third point we have,

$$
\begin{aligned}
(D_z X_n)^2 &= q^2 \sum_{r=0}^{q-1}\sum_{s=0}^{q-1-r} r!s! \binom{q-1}{r}^2 \binom{q-1-r}{s}^2 \\
&\quad \times I_{2q-2-2r-s}\big(f_n(z,\cdot) \otimes_r^s f_n(z,\cdot)\big) \\
(D_z X_n)^4 &= q^4 \sum_{r=0}^{q-1}\sum_{R=0}^{q-1}\sum_{s=0}^{q-1-r}\sum_{S=0}^{q-1-R} r!R!s!S! \binom{q-1}{r}^2 \binom{q-1}{R}^2 \\
&\quad \times \binom{q-1-r}{s}^2 \binom{q-1-R}{S}^2 \\
&\quad \times I_{2q-2-2r-s}\big(f_n(z,\cdot) \otimes_r^s f_n(z,\cdot)\big) I_{2q-2-2R-S}\left(f_n(z,\cdot) \otimes_R^S f_n(z,\cdot)\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\left[\int x^2 (D_z X_n)^4 \, d\mu(z)\right] &= q^4 \sum_{r=0}^{q-1}\sum_{R=0}^{q-1}\sum_{s=0}^{q-1-r}\sum_{S=0}^{q-1-R} r!R!s!S! \binom{q-1}{r}^2 \\
&\quad \times \binom{q-1}{R}^2 \binom{q-1-r}{s}^2 \\
&\quad \times \binom{q-1-R}{S}^2
\end{aligned}
$$

$$\times \int \mathbb{E}\left[I_{2q-2-2r-s}\left(xf_n(z,\cdot)\otimes_r^s f_n(z,\cdot)\right)\right.$$
$$\left.\times I_{2q-2-2R-S}\left(xf_n(z,\cdot)\otimes_R^S f_n(z,\cdot)\right)\right] d\mu(z).$$

The expectation, when $2r+s=2R+S$, is bounded by

$$\begin{aligned}
&(2q-2r-s-2)!\left|\left\langle xf_n(z,\cdot)\widetilde{\otimes}_r^s f_n(z,\cdot), xf_n(z,\cdot)\widetilde{\otimes}_R^S f_n(z,\cdot)\right\rangle_{\mathfrak{H}^{\otimes(2q-2R-S-2)}}\right| \\
&\quad\le (2q-2r-s-2)!\left\|xf_n(z,\cdot)\widetilde{\otimes}_r^s f_n(z,\cdot)\right\|_{\mathfrak{H}^{\otimes(2q-2r-s-2)}} \\
&\qquad\times \left\|xf_n(z,\cdot)\widetilde{\otimes}_R^S f_n(z,\cdot)\right\|_{\mathfrak{H}^{\otimes(2q-2R-S-2)}}.
\end{aligned}$$

Modulo the constant factor $(2q-2r-s-2)!$, the integral of the expectation is bounded by

$$\begin{aligned}
&\int \left\|xf_n(z,\cdot)\widetilde{\otimes}_r^s f_n(z,\cdot)\right\|_{\mathfrak{H}^{\otimes(2q-2r-s-2)}}\left\|xf_n(z,\cdot)\widetilde{\otimes}_R^S f_n(z,\cdot)\right\|_{\mathfrak{H}^{\otimes(2q-2R-S-2)}} d\mu(z) \\
&\quad= \left\langle \left\|xf_n(z,\cdot)\widetilde{\otimes}_r^s f_n(z,\cdot)\right\|_{\mathfrak{H}^{\otimes(2q-2r-s-2)}}, \left\|xf_n(z,\cdot)\widetilde{\otimes}_R^S f_n(z,\cdot)\right\|_{\mathfrak{H}^{\otimes(2q-2R-S-2)}}\right\rangle_{\mathfrak{H}} \\
&\quad\le \left\|\left\|xf_n(z,\cdot)\otimes_r^s f_n\right\|_{H^{\otimes(2q-2r-s-2)}}\right\|_{\mathfrak{H}} \times \left\|\,\|xf_n(z,\cdot)\otimes_R^S f_n(z,\cdot)\|_{\mathfrak{H}^{\otimes(2q-2R-S-2)}}\right\|_{\mathfrak{H}}.
\end{aligned}$$

We will work out the first factor:

$$\begin{aligned}
\left\|\,\left\|xf_n(z,\cdot)\otimes_r^s f_n(z,\cdot)\right\|_{\mathfrak{H}^{\otimes(2q-2r-s-2)}}\right\|_{\mathfrak{H}}^2 &= \int \left\|xf_n(z,\cdot)\otimes_r^s f_n(z,\cdot)\right\|^2_{\mathfrak{H}^{\otimes(2q-2r-s-2)}} d\mu(z) \\
&= \int \left\|\left(f_n\otimes_r^{s+1} f_n\right)(z,\cdot)\right\|^2_{\mathfrak{H}^{\otimes(2q-2r-s-2)}} d\mu(z) \\
&= \left\|f_n\otimes_r^{s+1} f_n\right\|^2_{\mathfrak{H}^{\otimes(2q-2r-s-1)}}.
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathbb{E}\left[\int x^2 (D_z X_n)^4 d\mu(z)\right] &\le q^4 \sum_{r,R=0}^{q-1}\sum_{s=0}^{q-1-r}\sum_{S=0}^{q-1-R} r!R!s!S!\binom{q-1}{r}^2 \\
&\quad\times \binom{q-1}{R}^2\binom{q-1-r}{s}^2 \\
&\quad\times \binom{q-1-R}{S}^2 \mathbf{1}_{\{2r+s=2R+S\}}(2q-2r-s-2)! \\
&\quad\times \left\|f_n\otimes_r^{s+1} f_n\right\|^2_{\mathfrak{H}^{\otimes(2q-2r-s-1)}}\left\|f_n\otimes_R^{S+1} f_n\right\|^2_{\mathfrak{H}^{\otimes(2q-2R-S-1)}} \\
&= q^4 \sum_{r,R=0}^{q-1}\sum_{t=1}^{q-r}\sum_{T=1}^{q-R} r!R!(t-1)!(T-1)! \\
&\quad\times \binom{q-1}{r}^2\binom{q-1}{R}^2 \\
&\quad\times \binom{q-1-r}{t-1}^2\binom{q-1-R}{T-1}^2
\end{aligned}$$

$$\times \mathbf{1}_{\{2r+t=2R+T\}} (2q-2r-t-1)! \\ \times \left\| f_n \otimes_r^t f_n \right\|^2_{\mathfrak{H}^{\otimes(2q-2r-t)}} \times \left\| f_n \otimes_R^T f_n \right\|^2_{\mathfrak{H}^{\otimes(2q-2R-T)}}.$$

The third point then follows.

Finally, for the fourth point, we use the integration by parts formula for functionals in a fixed Wiener–Poisson chaos explained in Remark 2,

$$\begin{aligned} \mathbb{E}\left[X_n^4\right] &= \frac{3}{q}\mathbb{E}\left[X_n^2 \|DX_n\|_H^2\right] + \frac{3}{q}\mathbb{E}\left[\left\langle x\,(DX_n)^3\,, X_n + \theta_z x DX_n\right\rangle_{\mathfrak{H}}\right] \\ &= \frac{3}{q}U_n + \frac{3}{q}(V_n + W_n) \end{aligned}$$

where

$$\begin{aligned} V_n &= \mathbb{E}\left[\left\langle x\,(DX_n)^3\,, X_n\right\rangle_{\mathfrak{H}}\right] = \mathbb{E}\left[\left\langle x\,(DX_n)^2\,, X_n\,(DX_n)\right\rangle_{\mathfrak{H}}\right] \\ W_n &= \mathbb{E}\left[\left\langle x\,(DX_n)^3\,, \theta_z x DX_n\right\rangle_{\mathfrak{H}}\right] = \mathbb{E}\left[\int \theta_z x^2 (D_z X_n)^4\, d\mu(z)\right] \end{aligned}$$

and $U_n = \mathbb{E}\left[X_n^2 \|DX_n\|_{\mathfrak{H}}^2\right]$. It is sufficient to prove that $U_n \to q$, $V_n \to 0$ and $W_n \to 0$ as $n \to \infty$.

To compute U_n, note that $X_n^2 = \sum_{r=0}^{q} \sum_{s=0}^{q-r} r!s! \binom{q}{r}^2 \binom{q-r}{s}^2 I_{2q-2r-s}\left(f_n \otimes_r^s f_n\right)$. Using our expression for $\|DX_n\|_{\mathfrak{H}}^2$ above,

$$\begin{aligned} U_n &= \sum_{r=0}^{q}\sum_{R=1}^{q}\sum_{s=0}^{q-r}\sum_{S=0}^{q-R} Rr!R!s!S! \binom{q}{r}^2 \binom{q}{R}^2 \binom{q-r}{s}^2 \binom{q-R}{S}^2 \\ &\quad \times \mathbb{E}\left[I_{2q-2r-s}\left(f_n \otimes_r^s f_n\right) I_{2q-2R-S}\left(f_n \otimes_R^S f_n\right)\right] \\ &= \sum_{r=0}^{q}\sum_{R=1}^{q}\sum_{s=0}^{q-r}\sum_{S=0}^{q-R} Rr!R!s!S! \binom{q}{r}^2 \binom{q}{R}^2 \binom{q-r}{s}^2 \binom{q-R}{S}^2 \\ &\quad \times \mathbf{1}_{\{2r+s=2R+S\}} (2q-2r-s)! \left\langle f_n \widetilde{\otimes}_r^s f_n, f_n \widetilde{\otimes}_R^S f_n\right\rangle_{\mathfrak{H}^{\otimes(2q-2r-s)}} \\ &= \sum_{\substack{r=0,R=1\\ r\neq R}}^{q} \sum_{s=0}^{q-r}\sum_{S=0}^{q-R} Rr!R!s!S! \binom{q}{r}^2 \binom{q}{R}^2 \binom{q-r}{s}^2 \binom{q-R}{S}^2 \\ &\quad \times \mathbf{1}_{\{2r+s=2R+S\}} (2q-2r-s)! \left\langle f_n \widetilde{\otimes}_r^s f_n, f_n \widetilde{\otimes}_R^S f_n\right\rangle_{\mathfrak{H}^{\otimes(2q-2r-s)}} \\ &\quad + \sum_{r=1}^{q-1}\sum_{s=0}^{q-r} r\,(r!)^2 (s!)^2 \binom{q}{r}^4 \binom{q-r}{s}^4 (2q-2r-s)! \left\| f_n \widetilde{\otimes}_r^s f_n \right\|^2_{\mathfrak{H}^{\otimes(2q-2r-s)}} \\ &\quad + q\left(q! \|f_n\|^2_{\mathfrak{H}^{\otimes q}}\right)^2. \end{aligned}$$

We can again apply Hölder's inequality on the inner product, and conclude that all the terms go to 0 except the last term which goes to q. Therefore, $U_n \to q$ as $n \to \infty$.

Observe that

$$|V_n| \leq \mathbb{E}\left[\left\|x\,(DX_n)^2\right\|_{\mathfrak{H}} \|X_n\,(DX_n)\|_{\mathfrak{H}}\right] \leq \sqrt{\mathbb{E}\left[\left\|x\,(DX_n)^2\right\|_{\mathfrak{H}}^2\right]}\sqrt{\mathbb{E}\left[\|X_n\,(DX_n)\|_{\mathfrak{H}}^2\right]}$$
$$= \sqrt{\mathbb{E}\left[\int x^2\,(D_zX_n)^4\,d\mu\,(z)\right]}\sqrt{\mathbb{E}\left[\|X_n\,(DX_n)\|_{\mathfrak{H}}^2\right]}.$$

Note that

$$\mathbb{E}\left[\|X_n\,(DX_n)\|_{\mathfrak{H}}^2\right] = \mathbb{E}\left[X_n^2\,\|DX_n\|_{\mathfrak{H}}^2\right] = U_n \to q.$$

From the third point, $\mathbb{E}\left[\int_{\mathbb{R}^+\times\mathbb{R}} x^2\,(D_zX_n)^4\,d\mu\,(z)\right] \to 0$ so $V_n \to 0$ as $n \to \infty$.

Finally for W_n,

$$|W_n| = \mathbb{E}\left[\int \theta_z x^2\,(D_zX_n)^4\,d\mu\,(z)\right] \leq \mathbb{E}\left[\int x^2\,(D_zX_n)^4\,d\mu\,(z)\right] \longrightarrow 0.$$

Putting them together we get the fourth point: $\mathbb{E}[X_n^4] \to 3$ as $n \to \infty$. ■

In Wiener space, convergence in a fixed Wiener chaos to a standard normal distribution is characterized by the convergence of the fourth moments to 3 or of the convergence of the norm of certain contractions to 0 (see the list preceding Corollary 22). We would then like to see if the same situation holds in Wiener–Poisson space. At this point, this appears to be an open question. We then finish with the following theorem which shows convergence in distribution and of the fourth moments to 3 if certain contractions converge to 0.

Theorem 28. *Suppose Assumptions A and B′ hold. Let $X_n = I_q\,(f_n)$, with $\mathbb{E}\left[X_n^2\right] = q!\,\|f_n\|_{\mathfrak{H}^{\otimes q}}^2 \to 1$. Assume that for $r = 0, \ldots, q-1$ and $s = 0, \ldots, q-r$, $\left\|f_n \otimes_r^s f_n\right\|_{\mathfrak{H}^{\otimes(2q-2r-s)}} \mathbf{1}_{\{s=0,r\neq0\}\cup\{s\neq0,r=0\}} \to 0$. Then as $n \to \infty$,*

- $\mathbb{E}[X_n^4] \to 3$.
- $X_n \to \mathcal{N}\,(0, 1)$ *in distribution.*

Proof. The first assertion is the fourth point in Proposition 27. For the second point, we refer to Corollary 26 to see that it suffices to prove $g_*\,(X_n) - g_{X_n} \to 0$ in $L^2\,(\Omega)$ and $\left\langle |x|\,(DX_n)^2\,, |DX_n|\right\rangle_{\mathfrak{H}} \to 0$ in $L^1(\Omega)$ as $n \to \infty$. These are immediate when we note that

$$g_*(X_n) - g_{X_n} = 1 - \frac{1}{q}\|DX_n\|_{\mathfrak{H}}^2 \to 0 \quad \text{in } L^2(\Omega) \text{ (by point 2 of Proposition 27)}$$

and

$$\mathbb{E}\left[\left\langle |x|\,(DX_n)^2\,, |DX_n|\right\rangle_{\mathfrak{H}}\right] \leq \mathbb{E}\left[\left\|x\,(DX_n)^2\right\|_{\mathfrak{H}} \|DX_n\|_{\mathfrak{H}}\right]$$
$$\leq \sqrt{\mathbb{E}\left[\left\|x\,(DX_n)^2\right\|_{\mathfrak{H}}^2\right]}\sqrt{\mathbb{E}\left[\|DX_n\|_{\mathfrak{H}}^2\right]}$$
$$= \sqrt{\mathbb{E}\left[\int x^2\,(DX_n)^4\,d\mu\right]}\sqrt{q\mathbb{E}\left[X_n^2\right]} \to 0$$

(by point 3 of Proposition 27). ■

Acknowledgments

We are deeply grateful to our respective Ph.D. advisors, Frederi Viens and Fabrice Baudoin, for their very helpful comments in making this paper, and primarily for introducing us to Malliavin calculus. We would also like to thank David Nualart for his valuable insights, and the anonymous referees for their careful reading and astute observations.

References

[1] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge University Press, Cambridge, 2004.
[2] F. Baudoin, C. Ouyang, S. Tindel, Upper bounds for the density of solutions of stochastic differential equations driven by fractional Brownian motions, Ann. Inst. H. Poincaré 50 (1) (2014) 111–135.
[3] S. Bourguin, C. Tudor, Cramér theorem for Gamma random variables, Electron. Comm. Probab. 16 (2011) 365–378.
[4] S. Chatterjee, A new method of Normal approximation, Ann. Probab. 36 (4) (2008) 1584–1610.
[5] L. Chen, Q.-M. Shao, Stein's method for normal approximation, in: An Introduction to Stein's Method, in: Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., Singapore Univ. Press, Singapore, 2005, pp. 1–59.
[6] P. Diaconis, S. Zabell, Closed form summation for classical distributions: variations on a theme of De Moivre, Statist. Sci. 14 (3) (1991) 284–302.
[7] G. Di Nunno, On orthogonal polynomials and the Malliavin derivative for Lévy stochastic measures, Anal. Probab. Sém. Congr. 16 (2008) 55–69.
[8] R. Eden, F. Viens, General upper and lower tail estimates using Malliavin calculus and Stein's equations, in: Seminar on Stochastic Analysis, Random Fields and Applications VII—Progress in Probability. Vol. 67, 2013, pp. 55–84.
[9] K. Itô, Spectral type of the shift transformation of differential processes with stationary increments, Trans. Amer. Math. Soc. 81 (2) (1956) 252–263.
[10] S. Kusuoka, C. Tudor, Stein method for invariant measures of diffusions via Malliavin calculus, Stochastic Process. Appl. 122 (4) (2012) 1627–1651.
[11] Y. Lee, H. Shih, The product formula of multiple Lévy–Itô integrals, Bull. Inst. Math. Acad. Sin. 32 (2) (2004) 71–95.
[12] S. Noreddine, I. Nourdin, On the Gaussian approximation of vector-valued multiple integrals, J. Multivariate Anal. 102 (6) (2011) 1008–1017.
[13] I. Nourdin, G. Peccati, Noncentral convergence of multiple integrals, Ann. Probab. 37 (4) (2009) 1412–1426.
[14] I. Nourdin, G. Peccati, Stein's method meets Malliavin calculus: a short survey with new estimates, in: J. Duan, S. Luo, C. Wang (Eds.), Recent Development in Stochastic Dynamics and Stochastic Analysis, World Scientific, 2009.
[15] I. Nourdin, G. Peccati, Stein's method on Wiener chaos, Probab. Theory Related Fields 145 (1) (2009) 75–118.
[16] I. Nourdin, G. Peccati, G. Reinert, Second order Poincaré inequalities and CLTs on Wiener space, J. Funct. Anal. 257 (2) (2009) 593–609.
[17] I. Nourdin, F. Viens, Density formula and concentration inequalities with Malliavin calculus, Electron. J. Probab. 14 (2009) 2287–2309.
[18] D. Nualart, The Malliavin Calculus and Related Topics, second ed., Springer Verlag, Berlin, 2006.
[19] D. Nualart, S. Ortiz-Latorre, Central limit theorems for multiple stochastic integrals and Malliavin calculus, Stochastic Process. Appl. 118 (2008) 614–628.
[20] D. Nualart, G. Peccati, Central limit theorems for sequences of multiple stochastic integrals, Ann. Probab. 33 (1) (2005) 177–193.
[21] G. Peccati, J.L. Solé, M.S. Taqqu, F. Utzet, Stein's method and normal approximation of Poisson functionals, Ann. Probab. 38 (2) (2010) 443–478.
[22] K. Sato, Lévy Processes and Infinitely Divisible Distributions, Cambridge University Press, Cambridge, 1999.
[23] J.L. Solé, F. Utzet, J. Vives, Canonical Lévy process and Malliavin calculus, Stochastic Process. Appl. 117 (2) (2007) 165–187.
[24] J.L. Solé, F. Utzet, J. Vives, Chaos expansions and Malliavin calculus for Lévy processes, in: Stochastic Analysis and Applications—Abel Symposia, vol. 2, Springer, 2007, pp. 595–612.
[25] C. Stein, Approximate Computation of Expectations, in: Institute of Mathematical Statistics Lecture Notes—Monograph Series, vol. 7, 1986.
[26] C. Tudor, Asymptotic Cramér's theorem and analysis on Wiener space, in: Séminaire de Probabilité XLIII, in: Lecture Notes in Mathematics, 2011, pp. 309–325.

[27] C. Tudor, F. Viens, Variations and estimators for the selfsimilarity order through Malliavin calculus, Ann. Probab. 6 (2009) 2093–2134.
[28] F. Viens, Stein's lemma, Malliavin Calculus, and tail bounds, with application to polymer fluctuation exponent, Stochastic Process. Appl. 119 (2009) 3671–3698.
[29] J.J. Viquez, On the second order Poincaré inequality and CLT on Wiener–Poisson space. Preprint http://arxiv.org/abs/1104.1837.