Multivariate second order Poincaré inequalities for Poisson functionals

Matthias Schulte[*], J. E. Yukich[†]

March 29, 2018

Abstract

Given a vector $F = (F_1, \dots, F_m)$ of Poisson functionals $F_1, \dots, F_m$, we establish quantitative bounds for the proximity between F and an m-dimensional centered Gaussian random vector N_Σ with covariance matrix $\Sigma \in \mathbb{R}^{m\times m}$. We derive results for the d_2- and d_3-distances based on smooth test functions as well as for the d_{convex}-distance and the $d_{\mathbb{H}_\ell}$-distance given by

$$d_{\mathbb{H}_\ell}(F, N_\Sigma) := \sup_{h\in\mathbb{H}_\ell} |\mathbb{E}\, h(F) - \mathbb{E}\, h(N_\Sigma)|,$$

a multi-dimensional generalization of the Kolmogorov distance, where $\ell \in \mathbb{N}$ and $\mathbb{H}_\ell$ is the set of indicator functions of intersections of ℓ closed half-spaces in $\mathbb{R}^m$. The bounds are multivariate counterparts of the second order Poincaré inequalities of [15] and, as such, are expressed in terms of integrated moments of first and second order difference operators. The derived second order Poincaré inequalities for non-smooth test functions, which are of the same order as for smooth test functions, are made possible by new bounds on the derivatives of solutions to the Stein equation for the multivariate normal distribution, which might be of independent interest. We present applications to the multivariate normal approximation of first order Wiener-Itô integrals and of statistics of Boolean models.

Key words and phrases. Stein's method, multivariate normal approximation, second order Poincaré inequality, Malliavin calculus, smoothing, Poisson process

AMS 2010 Subject Classification: Primary 60F05; Central limit and other weak theorems. Secondary 60D05, Geometric probability and stochastic geometry.

[*]Universität Bern, matthias.schulte@stat.unibe.ch

[†]Lehigh University, jey0@lehigh.edu, supported in part by NSF grant DMS-1406410

1 Introduction and main results

1.1 Overview

Roughly speaking, a first order Poincaré inequality for a random variable F measures the closeness of F to its mean. A second order Poincaré inequality [5] measures the closeness of F to a Gaussian random variable, where distance is given by some specified metric on the space of distribution functions. The paper [15] establishes second order Poincaré inequalities for Poisson functionals F, with bounds given in terms of integrated moments of first and second order difference operators, which are an outcome of the research on the Malliavin-Stein method for Poisson functionals in the recent years; see, for example, [7, 22, 30] and the book [21]. The bounds from [15] can be usefully applied to yield presumably optimal rates of normal convergence for many functionals of Poisson processes, including those represented as a sum of stabilizing score functions [14].

The goal of this paper is to establish second order Poincaré inequalities for Poisson functionals in the multivariate setting, providing multivariate counterparts to the univariate results of [15]. The proofs combine Malliavin calculus on Poisson spaces with Stein's method of multivariate normal approximation. Presumably optimal rates of normal convergence depend on good bounds on the right-hand sides of smoothing lemmas. A main contribution of this paper is to provide such bounds via new estimates on derivatives of the solutions to the Stein equation for the multivariate normal distribution, which could be helpful for the multivariate normal approximation of other types of random vectors as well and, thus, might be of independent interest.

We start by making our terms precise and recalling the univariate set-up. Let η be a Poisson process over a measurable space $(\mathbb{X}, \mathcal{F})$ with a σ-finite intensity measure λ (see e.g. [16] for more details on Poisson processes). One can think of η as a random element in the space $\mathbf{N}$ of all σ-finite counting measures equipped with the σ-field generated by the mappings $\nu \mapsto \nu(A)$, $A \in \mathcal{F}$. We call a random variable F a Poisson functional if there is a measurable map $f : \mathbf{N} \to \mathbb{R}$ such that $F = f(\eta)$ almost surely. The map f is called a representative of F. For such a Poisson functional F the difference operator is given by

$$D_x F := f(\eta + \delta_x) - f(\eta), \quad x \in \mathbb{X}, \tag{1.1}$$

where δ_x denotes the Dirac measure of x. We say that F belongs to the domain of the difference operator, i.e., $F \in \operatorname{dom} D$, if $\mathbb{E}\, F^2 < \infty$ and

$$\int_{\mathbb{X}} \mathbb{E}\,(D_x F)^2 \, \lambda(\mathrm{d}x) < \infty. \tag{1.2}$$

Iterating the definition of the difference operator one obtains

$$D^2_{x_1,x_2} F := D_{x_1}(D_{x_2} F) = f(\eta + \delta_{x_1} + \delta_{x_2}) - f(\eta + \delta_{x_1}) - f(\eta + \delta_{x_2}) + f(\eta), \quad x_1, x_2 \in \mathbb{X}.$$

Often one is interested in how close the distribution of F is to that of a standard Gaussian random variable N. To compare two random variables Y and Z or, more precisely, their distributions, one can use the Kolmogorov distance

$$d_K(Y,Z) := \sup_{u\in\mathbb{R}} |\mathbb{P}(Y \leq u) - \mathbb{P}(Z \leq u)|, \tag{1.3}$$

which is the supremum norm of the difference of the distribution functions of Y and Z, or the Wasserstein distance

$$d_W(Y,Z) := \sup_{h\in\mathrm{Lip}(1)} |\mathbb{E}\, h(Y) - \mathbb{E}\, h(Z)|,$$

where $\mathrm{Lip}(1)$ stands for the set of functions $h : \mathbb{R} \to \mathbb{R}$ with Lipschitz constant at most one. Note that the d_K-distance is always defined, while the d_W-distance requires finiteness of $\mathbb{E}\,|Y|$ and $\mathbb{E}\,|Z|$.

When $F \in \operatorname{dom} D$, $\mathbb{E}\, F = 0$, and $\mathrm{Var} F = 1$, the main results of [15] establish the inequalities

$$d_W(F,N) \leq \tau_1 + \tau_2 + \tau_3 \tag{1.4}$$

and

$$d_K(F,N) \leq \tau_1 + \tau_2 + \tau_3 + \tau_4 + \tau_5 + \tau_6, \tag{1.5}$$

where $\tau_1, \dots, \tau_6$ are integrals over moments involving only DF and D^2F (see Subsection 1.2 in [15] for exact formulas). The authors of [15] call (1.4) and (1.5), whose proofs rely on previous Malliavin-Stein bounds in [22] and [7, 30], respectively, second order Poincaré inequalities. The reason for this name is that the 'first order' Poincaré inequality

$$\mathrm{Var} F \leq \int_{\mathbb{X}} \mathbb{E}\, (D_x F)^2 \, \lambda(\mathrm{d}x)$$

for $F \in \operatorname{dom} D$ bounds the variance in terms of the first difference operator, whereas the first and the second difference operator control the closeness to Gaussianity in (1.4) and (1.5). The term second order Poincaré inequality was coined in [5] in a similar Gaussian framework, where one has the first two derivatives instead of the first two difference operators.

For many Poisson functionals F the second order Poincaré inequalities (1.4) and (1.5) may be evaluated since the first two difference operators have a clear interpretation via the operation of adding additional points. This is the advantage of these findings over some other Malliavin-Stein bounds for normal approximation of Poisson functionals (see, for example, [7, 11, 22, 30]), which require the knowledge of the whole chaos expansion of F.

The second order Poincaré inequality (1.5) yields rates of normal approximation for some classic problems in stochastic geometry and some non-linear functionals of Poisson-shot-noise processes [15], as well as for functionals of convex hulls of random samples in a smooth convex body, statistics of nearest neighbors graphs, the number of maximal points in a random sample, and estimators of surface area and volume arising in set approximation [14]. The rates of convergence for these examples are of the same order as in the classical central limit theorem and, thus, presumably optimal.

Often one is not only interested in the behavior of a single Poisson functional but in that of a vector $F = (F_1, \dots, F_m)$ of Poisson functionals $F_1, \dots, F_m$ with $m \in \mathbb{N}$. In this situation, one can compare F with an m-dimensional centered Gaussian random vector N_Σ with covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$. We are not only interested in weak convergence results for a vector of Poisson functionals with N_Σ as limiting distribution, which can be deduced from the univariate case by the Cramer-Wold technique, but in quantitative bounds for the closeness between F and N_Σ. In other words, we seek the multivariate counterparts of (1.4) and (1.5).

In this paper F and N_Σ are compared with respect to several distances based on smooth and non-smooth test functions. One of the main achievements of this paper is to show bounds for both cases that are of the same (presumably optimal) order. In general, it is more intricate to deal with non-smooth test functions when one uses Stein's method for multivariate normal approximation. For some bounds for smooth test functions having the same order as in the univariate case we refer to [6, Chapter 12] and the references therein. For non-smooth test functions, even obtaining the rate $n^{-1/2}$ in the classical central limit theorem for sums of n i.i.d. random vectors via Stein's method is challenging [1, 10]. The abstract multivariate normal approximation results in terms of the dependence structure in [27] and [6, Chapter 12] and in terms of exchangeable pairs in [26] contain additional logarithmic factors compared to what one would expect from the case of smooth test functions or from the univariate case. Recently, these logarithms were removed in [8] and [9], using the dependence structure and Stein couplings, respectively. However, it seems that none of these findings can be applied to systematically achieve the normal approximation bounds for Poisson functionals given by our main results.

1.2 Statement of main results

Let us now give a precise formulation of our results. We start with distances defined in terms of smooth test functions, namely the d_2- and the d_3-distances. Let $\mathcal{H}_m^{(2)}$ be the set of all C^2-functions $h : \mathbb{R}^m \to \mathbb{R}$ such that

$$|h(x) - h(y)| \leq \|x - y\|, \quad x, y \in \mathbb{R}^m, \quad \text{and} \quad \sup_{x \in \mathbb{R}^m} \|\operatorname{Hess} h(x)\|_{op} \leq 1,$$

where $\operatorname{Hess} h$ denotes the Hessian matrix of h and $\|\cdot\|_{op}$ stands for the operator norm of a matrix. By $\mathcal{H}_m^{(3)}$ we denote the class of all C^3-functions $h:\mathbb{R}^m\to\mathbb{R}$ such that the absolute values of the second and third partial derivatives are bounded by one. Using this notation, we define, for m-dimensional random vectors Y and Z,

$$d_2(Y,Z):=\sup_{h\in\mathcal{H}_m^{(2)}}|\mathbb{E}\,h(Y)-\mathbb{E}\,h(Z)|$$

if $\mathbb{E}\,\|Y\|,\mathbb{E}\,\|Z\|<\infty$ and

$$d_3(Y,Z):=\sup_{h\in\mathcal{H}_m^{(3)}}|\mathbb{E}\,h(Y)-\mathbb{E}\,h(Z)|$$

if $\mathbb{E}\,\|Y\|^2,\mathbb{E}\,\|Z\|^2<\infty$.

The paper [22] was the first to combine Stein's method and the Malliavin calculus to obtain normal approximation of Poisson functionals. In [23], the univariate main result of [22] for the d_W-distance is extended to vectors of Poisson functionals and the d_2- and the d_3-distances are considered. Evaluating these multivariate Malliavin-Stein bounds in the same way one evaluates in [15] the univariate bounds from [22] and [7, 30] to derive (1.4) and (1.5), one obtains the following multivariate second order Poincaré inequalities.

Theorem 1.1. *Let $F=(F_1,\dots,F_m)$, $m\in\mathbb{N}$, be a vector of Poisson functionals $F_1,\dots,F_m\in\operatorname{dom}D$ with $\mathbb{E}\,F_i=0$, $i\in\{1,\dots,m\}$. Define*

$$\gamma_1:=\Bigg(\sum_{i,j=1}^m\int_{\mathbb{X}^3}\big(\mathbb{E}\,(D^2_{x_1,x_3}F_i)^2(D^2_{x_2,x_3}F_i)^2\big)^{1/2}\big(\mathbb{E}\,(D_{x_1}F_j)^2(D_{x_2}F_j)^2\big)^{1/2}\,\lambda^3(\mathrm{d}(x_1,x_2,x_3))\Bigg)^{1/2}$$

$$\gamma_2:=\Bigg(\sum_{i,j=1}^m\int_{\mathbb{X}^3}\big(\mathbb{E}\,(D^2_{x_1,x_3}F_i)^2(D^2_{x_2,x_3}F_i)^2\big)^{1/2}\big(\mathbb{E}\,(D^2_{x_1,x_3}F_j)^2(D^2_{x_2,x_3}F_j)^2\big)^{1/2}\,\lambda^3(\mathrm{d}(x_1,x_2,x_3))\Bigg)^{1/2}$$

$$\gamma_3:=\sum_{i=1}^m\int_{\mathbb{X}}\mathbb{E}\,|D_xF_i|^3\,\lambda(\mathrm{d}x)$$

and let $\Sigma=(\sigma_{ij})_{i,j\in\{1,\dots,m\}}\in\mathbb{R}^{m\times m}$ be positive semi-definite. Then,

$$d_3(F,N_\Sigma)\le\frac{m}{2}\sum_{i,j=1}^m|\sigma_{ij}-\operatorname{Cov}(F_i,F_j)|+m\gamma_1+\frac{m}{2}\gamma_2+\frac{m^2}{4}\gamma_3. \tag{1.6}$$

If, additionally, Σ is positive definite, then

$$\begin{aligned}d_2(F,N_\Sigma)\le{}&\|\Sigma^{-1}\|_{op}\|\Sigma\|_{op}^{1/2}\sum_{i,j=1}^m|\sigma_{ij}-\operatorname{Cov}(F_i,F_j)|+2\|\Sigma^{-1}\|_{op}\|\Sigma\|_{op}^{1/2}\gamma_1\\&+\|\Sigma^{-1}\|_{op}\|\Sigma\|_{op}^{1/2}\gamma_2+\frac{\sqrt{2\pi}m^2}{8}\|\Sigma^{-1}\|_{op}^{3/2}\|\Sigma\|_{op}\gamma_3.\end{aligned}\tag{1.7}$$

Note that γ_1, γ_2, and γ_3 have a structure similar to that of τ_1, τ_2, and τ_3 in (1.4) and (1.5) and coincide with them up to some constant factors for $m = 1$.

Let us now compare Theorem 1.1 with related results in the literature. The bounds in the underlying paper [23] are formulated in terms of the difference operator D and the inverse Ornstein-Uhlenbeck generator L^{-1} and do not, in general, readily lend themselves to off-the-shelf use. In contrast, the bounds (1.6) and(1.7) involving only difference operators are often tractable, as seen in our applications section and also in the companion paper [31]. Theorem 8.1 of [11] provides a bound on $d_3(F, N_\Sigma)$, which relies on the findings of [23], though this bound requires knowledge of the entire Wiener-Itô chaos expansion for each of the components of F and consequently may also be less useful than (1.6). When the components of F belong to a special class of Poisson U-statistics, which admit a finite chaos expansion with explicitly known kernels, the paper [17] uses the results of [23] to establish Berry-Esseen bounds for the d_3-distance between F and a Gaussian random vector. In [3], the findings from [23] are generalized by comparing a vector of Poisson functionals with a random vector composed of Gaussian and Poisson random variables.

In [13] multivariate second order Poincaré inequalities for functionals of Rademacher sequences are derived. The considered d_4-distance is based on test functions such that the sup-norms of the first four partial derivatives are bounded by one.

To some extent (1.6) and (1.7) can be seen as multivariate counterparts of (1.4). Indeed, as is the case with d_W, the distances d_2 and d_3 are based on continuous test functions, although the exact definitions involving C^2- and C^3-functions are distinct from the multivariate Wasserstein distance obtained by using test functions $h : \mathbb{R}^m \to \mathbb{R}$ having Lipschitz constants at most one.

The Kolmogorov distance (1.3) is arguably more interesting than the Wasserstein distance (and the d_2- and the d_3-distances for $m = 1$), as it has a clearer interpretation as the supremum norm of the difference of the distribution functions of the involved random variables, though it is often harder to deal with because the underlying test functions are discontinuous. The straightforward multivariate analog to the univariate Kolmogorov distance for two m-dimensional random vectors $Y = (Y_1, \dots, Y_m)$ and $Z = (Z_1, \dots, Z_m)$ would be

$$d_K(Y, Z) := \sup_{u_1,\dots,u_m \in \mathbb{R}} |\mathbb{P}(Y_1 \leq u_1, \dots, Y_m \leq u_m) - \mathbb{P}(Z_1 \leq u_1, \dots, Z_m \leq u_m)|, \tag{1.8}$$

which is again the supremum norm of the difference of the distribution functions of Y and Z. In (1.8) one only takes into account rectangular solids aligned with coordinate planes, so that for a rotation $A \in \mathbb{R}^{m\times m}$ the distance between AY and AZ could be different from the distance between Y and Z. Although convergence in the distance given in (1.8) still implies weak convergence, one would like to have invariance under rotation. To resolve this issue, we define for $\ell \in \mathbb{N}$ and two m-dimensional random vectors Y and Z,

$$d_{\mathbb{H}_\ell}(Y,Z) := \sup_{h\in\mathbb{H}_\ell} |\mathbb{E}\, h(Y) - \mathbb{E}\, h(Z)|,$$

where $\mathbb{H}_\ell$ is the set of indicator functions of intersections of ℓ closed half-spaces in $\mathbb{R}^m$. For $m = \ell = 1$ this is the same as the univariate Kolmogorov distance and for $\ell = m$ it dominates the distance in (1.8), whence $d_{\mathbb{H}_\ell}$ may be viewed as a multi-dimensional generalization of the Kolmogorov distance. By the Cramer-Wold device, convergence in $d_{\mathbb{H}_\ell}$ for $\ell = 1$ (and, thus, for any $\ell \in \mathbb{N}$) implies weak convergence.

For a vector $F = (F_1,\dots,F_m)$, $m \in \mathbb{N}$, of Poisson functionals $F_1,\dots,F_m \in \operatorname{dom} D$ with $\mathbb{E}\, F_i = 0$, $i \in \{1,\dots,m\}$, we use the abbreviations $D_xF := (D_xF_1,\dots,D_xF_m)$ for $x \in \mathbb{X}$, $D^2_{x,y}F := (D^2_{x,y}F_1,\dots,D^2_{x,y}F_m)$ for $x,y \in \mathbb{X}$, and

$$\begin{aligned}
\gamma_4 := \Bigg(\sum_{j,k=1}^m \int_{\mathbb{X}} \mathbb{E}\,(D_xF_j)^4\,\lambda(\mathrm{d}x) + 6\int_{\mathbb{X}^2} \big(\mathbb{E}\,(D^2_{x,y}F_j)^4\big)^{1/2}\big(\mathbb{E}\,(D_xF_k)^4\big)^{1/2}\,\lambda^2(\mathrm{d}(x,y)) \\
+ 3\int_{\mathbb{X}^2} \big(\mathbb{E}\,(D^2_{x,y}F_j)^4\big)^{1/2}\big(\mathbb{E}\,(D^2_{x,y}F_k)^4\big)^{1/2}\,\lambda^2(\mathrm{d}(x,y))\Bigg)^{1/2}
\end{aligned}$$

$$\begin{aligned}
\gamma_5 := \Bigg(\sum_{i,j,k=1}^m \int_{\mathbb{X}} \mathbb{E}\,(D_xF_i)^6\,\lambda(\mathrm{d}x) \\
+ 8\int_{\mathbb{X}^2} \big(\mathbb{E}\,\mathbf{1}\{D^2_{x,y}F \neq \mathbf{0}\}|D_xF_i\,D_xF_j|^3\big)^{2/3}\big(\mathbb{E}\,(D_xF_k)^6\big)^{1/3}\,\lambda^2(\mathrm{d}(x,y)) \\
+ 42\int_{\mathbb{X}^2} \big(\mathbb{E}\,|D^2_{x,y}F_i|^6\big)^{1/3}\,\big(\mathbb{E}\,|D_xF_j|^6\big)^{1/3}\,\big(\mathbb{E}\,(D_xF_k)^6\big)^{1/3}\,\lambda^2(\mathrm{d}(x,y)) \\
+ 42\int_{\mathbb{X}^2} \big(\mathbb{E}\,|D^2_{x,y}F_i|^6\big)^{1/3}\,\big(\mathbb{E}\,|D^2_{x,y}F_j|^6\big)^{1/3}\,\big(\mathbb{E}\,(D_xF_k)^6\big)^{1/3}\,\lambda^2(\mathrm{d}(x,y)) \\
+ 14\int_{\mathbb{X}^2} \big(\mathbb{E}\,|D^2_{x,y}F_i|^6\big)^{1/3}\,\big(\mathbb{E}\,|D^2_{x,y}F_j|^6\big)^{1/3}\,\big(\mathbb{E}\,(D^2_{x,y}F_k)^6\big)^{1/3}\,\lambda^2(\mathrm{d}(x,y))\Bigg)^{1/2},
\end{aligned}$$

where $\mathbf{0}$ stands for the origin in $\mathbb{R}^m$.

The following multivariate second order Poincaré inequality shows that bounds in the $d_{\mathbb{H}_\ell}$-distance closely resemble those for the d_2 and d_3-distances at (1.6) and (1.7).

Theorem 1.2. *Let $F = (F_1,\dots,F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals $F_1,\dots,F_m \in \operatorname{dom} D$ with $\mathbb{E}\, F_i = 0$, $i \in \{1,\dots,m\}$, and let $\Sigma = (\sigma_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$ be positive definite. Then, for any $\ell \in \mathbb{N}$,*

$$d_{\mathbb{H}_\ell}(F, N_\Sigma) \leq 718 m^{47/24}\ell\|\Sigma^{-1}\|_{op} \max\Bigg\{ \sum_{i,j=1}^m |\sigma_{ij} - \operatorname{Cov}(F_i,F_j)|, \gamma_1, \gamma_2, \gamma_4, \frac{\sqrt{\ell}\sqrt{\gamma_5}}{\|\Sigma^{-1}\|_{op}^{1/4}}\Bigg\} \tag{1.9}$$

with γ_1 and γ_2 as in Theorem 1.1 and γ_4 and γ_5 as defined above.

To the best of our knowledge the $d_{\mathbb{H}_\ell}$-distance has never been used before. Instead, the standard multivariate counterpart to the univariate Kolmogorov distance (1.3) is the d_{convex}-distance, defined for two m-dimensional random vectors Y and Z as

$$d_{convex}(Y,Z) := \sup_{h\in\mathcal{I}_m} |\mathbb{E}\, h(Y) - \mathbb{E}\, h(Z)|,$$

where $\mathcal{I}_m$ is the set of all indicator functions of closed convex sets in $\mathbb{R}^m$. Under the additional assumption that the difference operators of the components of F are almost surely bounded, we may establish the following multivariate second order Poincaré inequality for d_{convex}, a counterpart to (1.6), (1.7), and (1.9).

Theorem 1.3. *Let $F = (F_1,\dots,F_m)$, $m\in\mathbb{N}$, be a vector of Poisson functionals $F_1,\dots,F_m \in \operatorname{dom} D$ with $\mathbb{E}\, F_i = 0$, $i\in\{1,\dots,m\}$. Assume that there is a constant $\varrho\in(0,\infty)$ such that*

$$\max_{i\in\{1,\dots,m\}} |D_xF_i| \le \varrho \quad \mathbb{P}\text{-a.s.}, \qquad \lambda\text{-a.e. } x\in\mathbb{X}, \tag{1.10}$$

and let $\Sigma = (\sigma_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$ be positive definite. Then, for any $A\in\mathcal{F}$ with $0<\lambda(A)<\infty$,

$$d_{convex}(F, N_\Sigma) \le 2304 m^3 \|\Sigma^{-1}\|_{op}\gamma \tag{1.11}$$

with

$$\gamma := \max\Bigg\{ \sum_{i,j=1}^m |\sigma_{ij} - \operatorname{Cov}(F_i,F_j)|, \gamma_1, \gamma_2, \gamma_4, \frac{8\sqrt{6}}{3} m^2 \|\Sigma^{-1}\|_{op}^{1/2} \varrho^3 \lambda(A), \\ \frac{m^{3/2}\sqrt{\varrho^4\lambda(A)}}{\|\Sigma^{-1}\|_{op}^{1/4}}, \frac{1}{m\|\Sigma^{-1}\|_{op}\lambda(A)} \int_{\mathbb{X}\setminus A} \mathbb{P}(D_xF\neq 0)\,\lambda(\mathrm{d}x)\Bigg\}$$

and γ_1, γ_2, and γ_4 as in Theorem 1.1 and Theorem 1.2.

Note that d_{convex} is stronger than $d_{\mathbb{H}_\ell}$ in the sense that d_{convex} is always at least $d_{\mathbb{H}_\ell}$ since $\mathbb{H}_\ell \subset \mathcal{I}_m$ for any $\ell\in\mathbb{N}$. Although any closed convex set in $\mathbb{R}^m$ is the intersection of at most countably many closed half-spaces, $d_{\mathbb{H}_\ell}$ and d_{convex} are not equivalent since $\mathrm{d}_{\mathbb{H}_\ell}$ considers only intersections of up to ℓ closed half-spaces. So the relation between Theorem 1.2 and Theorem 1.3 is that the latter concerns a stronger distance, but requires a more restrictive boundedness assumption on the difference operators. The proofs of Theorems 1.2 and 1.3 rely on two innovations, one of which involves bounding second moments instead of sup-norms of solutions to the multivariate Stein equation, whereas the other, in the setting of Theorem 1.2, allows us to remove the boundedness assumption (1.10) needed in Theorem 1.3 by using the particular structure of the test functions from $\mathbb{H}_\ell$.

The condition (1.10) in Theorem 1.3 concurs with existing multivariate results for the d_{convex}-distance or generalizations of it in [6, 8, 9, 26, 27] which all require some boundedness assumptions comparable to almost surely bounded difference operators. The Malliavin-Stein method is used in [19] to establish for the multivariate normal approximation of functionals of Gaussian processes bounds in the d_W-distance. In [12], a similar bound with an additional logarithm is derived for the d_{convex}-distance. Compared to Theorem 1.3, the latter result does not require any boundedness assumptions, but this might stem from the fact that the Malliavin-Stein approach in the Gaussian case involves less terms than in the Poisson case and, in particular, not the one requiring the boundedness assumption (1.10) in our approach. Moreover, we expect that one can use our proof technique to remove the logarithm from the result in [12]. For a subclass of functionals of Gaussian processes, namely multiple Wiener-Itô integrals, one may even establish rates of multivariate normal approximation with respect to the total variation distance [20]. This bound also involves additional logarithmic factors and its proof relies on controlling the relative entropy, an approach which differs from Stein's method.

Clearly, if the random vector N_Σ is replaced by a normal random vector whose covariance matrix consists of entries $\mathrm{Cov}(F_i, F_j)$, then the term $\sum_{i,j=1}^m |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)|$ in the multivariate bounds of our main theorems disappears. In Theorem 1.2 and Theorem 1.3 we require that the covariance matrix Σ of the approximating Gaussian random vector N_Σ is positive definite. Otherwise, N_Σ would be concentrated on some lower-dimensional linear subspace of $\mathbb{R}^m$. If now F belongs to any given lower dimensional subspace of $\mathbb{R}^m$ with probability zero, then we have $d_{\mathbb{H}_\ell}(F, N_\Sigma) \geq 1/2$ and $d_{convex}(F, N_\Sigma) \geq 1/2$. In such situations, one could have weak convergence without convergence in $d_{\mathbb{H}_\ell}$ or d_{convex}.

1.3 Examples and applications

At first sight, the bounds in our general results appear unwieldy. However for many functionals of interest, we may readily bound the integrated moments of difference operators and the terms $\gamma_1, ..., \gamma_5$ remarkably simplify. We illustrate this by four examples, which indicate that our bounds yield presumably optimal rates of convergence.

We start with the following analog to the classical central limit theorem for sums of i.i.d. random vectors, where we consider the sum of a Poisson distributed number of i.i.d. random vectors.

Corollary 1.4. *Given a Poisson distributed random variable Y with mean $s > 0$ and a sequence of i.i.d. centered random vectors $(X_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^m$, which are independent of Y, define*

$$Z_s := \frac{1}{\sqrt{s}} \sum_{n=1}^{Y} X_n \quad \text{and} \quad \Sigma := (\mathrm{Cov}(X_1^{(i)}, X_1^{(j)}))_{i,j\in\{1,\dots,m\}}.$$

(a) *It is the case that*

$$d_3(Z_s, N_\Sigma) \le \frac{m^2}{4} \sum_{i=1}^m \mathbb{E}\,|X_1^{(i)}|^3 \frac{1}{\sqrt{s}}.$$

(b) *If Σ is positive definite,*

$$d_2(Z_s, N_\Sigma) \le \frac{\sqrt{2\pi} m^2}{8} \|\Sigma^{-1}\|_{op}^{3/2} \|\Sigma\|_{op} \sum_{i=1}^m \mathbb{E}\,|X_1^{(i)}|^3 \frac{1}{\sqrt{s}}.$$

(c) *If Σ is positive definite, for any $\ell \in \mathbb{N}$,*

$$\begin{aligned} d_{\mathbb{H}_\ell}(Z_s, N_\Sigma) \le\ & 718 m^{59/24} \ell^{3/2} \max\{\|\Sigma^{-1}\|_{op}, \|\Sigma^{-1}\|_{op}^{3/4}\} \\ & \max\left\{ \left(\sum_{i=1}^m \mathbb{E}\,(X_1^{(i)})^4 \right)^{1/2}, \left(\sum_{i=1}^m \mathbb{E}\,(X_1^{(i)})^6 \right)^{1/4} \right\} \frac{1}{\sqrt{s}}. \end{aligned}$$

(d) *If there is a constant $a \in (0, \infty)$ such that $|X_1^{(i)}| \le a$ $\mathbb{P}$-a.s. for all $i \in \{1, \dots, m\}$ and if Σ is positive definite,*

$$d_{convex}(Z_s, N_\Sigma) \le 15050 m^5 \max\{\|\Sigma^{-1}\|_{op}^{3/4}, \|\Sigma^{-1}\|_{op}^{3/2}\} \max\{a^2, a^3\} \frac{1}{\sqrt{s}}.$$

Here, as well as in Theorems 1.1-1.3, we implicitly assume that the normal approximation bounds all involve finite quantities, as otherwise there is nothing to prove.

Since one can rewrite Z_s as a sum of a fixed number of i.i.d. random vectors, one can also apply the classical multivariate central limit theorem. In [1, 10, 28] corresponding Berry-Esseen inequalities for the d_{convex}-distance are derived, which provide in the case of Corollary 1.4 rates of convergence of the order $1/\sqrt{s}$ as well. These findings are even stronger since they require for the d_{convex}-distance only finite third moments, while we require bounded summands for the d_{convex}-distance and finite sixth moments for the $d_{\mathbb{H}_\ell}$-distance. The stricter assumptions in Corollary 1.4 might come from the fact that the proofs of the underlying results for more general Poisson functionals are not optimized for the considered special case.

Since Z_s is a vector of first order Wiener-Itô integrals, Corollary 1.4 follows from a more general theorem in Subsection 4.1, which is obtained by applying our main results to first order Wiener-Itô integrals.

As a second example we consider the case that one has for some $m \in \mathbb{N}$ a family of vectors $F_s = (F_{1,s}, \dots, F_{m,s})$, $s > 0$, of square integrable Poisson functionals $F_{1,s}, \dots, F_{m,s}$ with underlying Poisson processes η_s, $s > 0$, having intensity measures μ_s, $s > 0$, of the form $\mu_s = s\mu$ with a fixed finite measure μ, e.g., homogenous Poisson processes on the d-dimensional unit cube $[0,1]^d$ with increasing intensity. Moreover, we denote by Σ_s

the covariance matrix of F_s and assume that $(\Sigma_s)_{s>0}$ converges to a matrix $\overline{\Sigma} \in \mathbb{R}^{m\times m}$. Under some additional assumptions on the difference operators our main results imply the following bounds.

Corollary 1.5. *Let F_s, $s>0$, be as above and assume that $\overline{\Sigma}$ is positive definite and that there are constants $a, b \in (0,\infty)$ such that, for $i \in \{1,\dots,m\}$ and $s>0$,*

$$|D_x F_{i,s}| \leq \frac{a}{\sqrt{s}}, \quad \mu\text{-a.e. } x \in \mathbb{X}, \quad \text{and} \quad |D^2_{x_1,x_2} F_{i,s}| \leq \frac{a}{\sqrt{s}}, \quad \mu^2\text{-a.e. } (x_1,x_2) \in \mathbb{X}^2, \quad (1.12)$$

and

$$s \int_{\mathbb{X}} \mathbb{P}(D^2_{x,y} F_{i,s} \neq 0)^{1/4} \, \mu(\mathrm{d}y) \leq b, \quad \mu\text{-a.e. } x \in \mathbb{X}. \quad (1.13)$$

Let $\ell \in \mathbb{N}$. Then, there exist constants $s_0, C_3, C_2, C_{\mathbb{H}_\ell}, C_{convex} \in (0,\infty)$ depending on a, b, $\mu(\mathbb{X})$, $\overline{\Sigma}$, $(\Sigma_s)_{s>0}$ (and ℓ in case of $C_{\mathbb{H}_\ell}$) such that

$$d_3(F_s, N_{\Sigma_s}) \leq \frac{C_3}{\sqrt{s}}, \; d_2(F_s, N_{\Sigma_s}) \leq \frac{C_2}{\sqrt{s}}, \; d_{\mathbb{H}_\ell}(F_s, N_{\Sigma_s}) \leq \frac{C_{\mathbb{H}_\ell}}{\sqrt{s}}, \; d_{convex}(F_s, N_{\Sigma_s}) \leq \frac{C_{convex}}{\sqrt{s}}$$

for $s \geq s_0$.

We observe that the obtained rates of convergence in Corollary 1.5 are of the order $s^{-1/2}$ for all distances. The situation of Corollary 1.5 that one rescales by the square root of the intensity parameter and that the unrescaled difference operators are bounded occurs, for example, in some problems in stochastic geometry such as degree counts or component counts in nearest neighbors graphs (see Subsection 3.1 in [25] as well as Chapter 12.4 of [6] for a colored version arising in multivariate statistics for equality of distributions and Subsection 6.2 in [24], respectively).

The third example is the situation where, before centering, the components of F have representations $s^{-1/2}\sum_{x\in\eta_{sg}\cap A_i} \xi_s^{(i)}(x,\eta_{sg})$, $i \in \{1,\dots,m\}$, with $s \in [1,\infty)$, where η_{sg} is a Poisson process in $\mathbb{R}^d$ whose intensity measure has density sg with respect to the Lebesgue measure, A_i, $i \in \{1,\dots,m\}$, are bounded subsets of $\mathbb{R}^d$, and $\xi_s^{(i)}$, $i \in \{1,\dots,m\}$, are stabilizing score functions. Then the companion paper [31], which can be seen as a multivariate counterpart to some of the findings in [14], shows that the right-hand sides of (1.6), (1.7), (1.9), and (1.11) reduce to $O(\sum_{i,j=1}^m |\sigma_{ij} - \mathrm{Cov}(F_i,F_j)|) + O(s^{-1/2})$ under some assumptions on $(\xi_s^{(i)})_{s\geq 1}$, $i \in \{1,\dots,m\}$, A_i, $i \in \{1,\dots,m\}$, and g. This means that the approximation error consists of a term taking the difference of the covariances into account and a rate of order $s^{-1/2}$, which also occurs in the univariate case (see [14]). In Section 3 of [31], these findings are applied to obtain quantitative multivariate central limit theorems for statistics of k-nearest neighbors graphs and random geometric graphs as well as statistics arising in topological data analysis and entropy estimation.

As a fourth example we mention the intrinsic volumes of Boolean models, a prominent problem from stochastic geometry. Here, our main results lead to the convergence rate $\mathrm{Vol}(W)^{-1/2}$, where $\mathrm{Vol}(W)$ is the volume of the compact convex observation window W, if one compares the vector of intrinsic volumes of the Boolean model in W with a centered Gaussian random vector having exactly the same covariance matrix and if one increases the inradius of W; see Subsection 4.2.

In the last three examples the rates of convergence $s^{-1/2}$ and $\mathrm{Vol}(W)^{-1/2}$, respectively, are comparable to $n^{-1/2}$ in the uni- and multivariate central limit theorems for the i.i.d. case and, thus, presumably optimal.

Among these examples, we will consider the first order Wiener-Itô integrals generalizing the situation of Corollary 1.4 and the intrinsic volumes of Boolean models in more detail in Subsections 4.1 and 4.2, while Corollary 1.5 is a consequence of a theorem derived in Subsection 4.3.

1.4 Proof techniques

Let us now informally comment on the method of proof. Assume we aim to compare an m-dimensional random vector $Y = (Y_1, \dots, Y_m)$ with an m-dimensional centered Gaussian random vector N_I with the identity matrix $I \in \mathbb{R}^{m\times m}$ as covariance matrix (we assume $\Sigma = I$ for simplicity) in terms of a measurable test function $h : \mathbb{R}^m \to \mathbb{R}$. The idea of Stein's method for multivariate normal approximation (see e.g. [6, 10]) is now to use the identity

$$\mathbb{E}\, h(Y) - \mathbb{E}\, h(N_I) = \mathbb{E} \sum_{i=1}^m Y_i \frac{\partial f_h}{\partial y_i}(Y) - \frac{\partial^2 f_h}{\partial y_i^2}(Y),$$

where $f_h : \mathbb{R}^m \to \mathbb{R}$ is a solution of the multivariate Stein equation

$$\sum_{i=1}^m y_i \frac{\partial f}{\partial y_i}(y) - \frac{\partial^2 f}{\partial y_i^2}(y) = h(y) - \mathbb{E}\, h(N_I), \quad y \in \mathbb{R}^m. \tag{1.14}$$

Under some smoothness assumptions on h one can give formulas for f_h (see, for example, Lemma 2.6 in [6]). However for non-smooth h such as indicator functions it appears unclear how to deal with f_h. This problem is resolved by considering instead of h some smoothed C^∞ version $h_{t,I}$ of h, which depends on some smoothing parameter $t \in (0,1)$. Of course one makes some error by replacing the test functions defining the $d_{\mathbb{H}_\ell}$- and d_{convex}-distances by their smoothed versions, but smoothing lemmas allow us to bound this error by some constant multiple of $\sqrt{t}$.

Thus it remains to find upper bounds for $|\mathbb{E}\, h_{t,I}(Y) - \mathbb{E}\, h_{t,I}(N_I)|$ as a function of $t \in (0,1)$. We sketch how this goes as follows. Given $h : \mathbb{R}^m \to \mathbb{R}$ measurable and bounded and $t \in (0,1)$ we introduce the smoothed function

$$h_{t,I}(y) := \int_{\mathbb{R}^m} h(\sqrt{t}z + \sqrt{1-t}y)\, \varphi_I(z)\,\mathrm{d}z, \quad y \in \mathbb{R}^m, \tag{1.15}$$

where φ_I denotes the density of N_I. The function $f_{t,h,I} : \mathbb{R}^m \to \mathbb{R}$ given by

$$f_{t,h,I}(y) := -\frac{1}{2}\int_t^1 \frac{1}{1-s}\int_{\mathbb{R}^m} (h(\sqrt{s}z + \sqrt{1-s}y) - h(z))\, \varphi_I(z)\,\mathrm{d}z\,\mathrm{d}s, \quad y \in \mathbb{R}^m, \tag{1.16}$$

is a solution of the Stein equation (1.14) with h replaced by $h_{t,I}$; see [10, p. 726] and [6, p. 337]. Moreover, when $\|h\|_\infty := \sup_{x\in\mathbb{R}^m} |h(x)| \le 1$, it follows (see e.g. the first display on p. 1498 in [23]) that, for a vector $F = (F_1, \dots, F_m)$, $m \in \mathbb{N}$, of Poisson functionals $F_1, \dots, F_m \in \operatorname{dom} D$ with $\mathbb{E}\, F_i = 0$, $i \in \{1, \dots, m\}$,

$$|\mathbb{E}\, h_{t,I}(F) - \mathbb{E}\, h_{t,I}(N_I)| = \left| \sum_{i=1}^m \mathbb{E}\, \frac{\partial^2 f_{t,h,I}}{\partial y_i^2}(F) - \sum_{k=1}^m \mathbb{E} \int_{\mathbb{X}} D_x \frac{\partial f_{t,h,I}}{\partial y_k}(F)(-D_x L^{-1} F_k)\, \lambda(\mathrm{d}x) \right|,$$

where D_x is the difference operator given in (1.1) and L^{-1} is the inverse Ornstein-Uhlenbeck generator defined in the Appendix.

A main idea behind the proofs of Theorems 1.2 and 1.3 is to show for the right-hand side of the above a bound involving $\sqrt{\sum_{i,j=1}^m \mathbb{E}\left(\frac{\partial^2 f_{t,h,I}}{\partial y_i \partial y_j}(F)\right)^2}$ and then to use

$$\sup_{h\in\mathbb{H}_\ell} \mathbb{E} \sum_{i,j=1}^m \left(\frac{\partial^2 f_{t,h,I}}{\partial y_i \partial y_j}(F)\right)^2 \le M_2 (\log t)^2 d_{\mathbb{H}_{2\ell}}(F, N_I) + 444 m^{23/6} \tag{1.17}$$

for all $t \in (0,1)$, $i, j \in \{1, \dots, m\}$, and $\ell \in \mathbb{N}$ with $M_2 \le m^2$ as well as a similar bound for $h \in \mathcal{I}_m$ (cf. Proposition 2.4). By choosing t appropriately we may deduce the bounds of Theorems 1.2 and 1.3. In the case of Theorem 1.2, one needs a second technique, which makes use of the special form of the constituent test functions defining $d_{\mathbb{H}_\ell}$, to remove the boundedness assumption (1.10) on $D_x F_i$, $i \in \{1, \dots, m\}$, in Theorem 1.3. The inequality (1.17) is not restricted to a vector F of Poisson functionals, but holds for arbitrary random vectors Y in $\mathbb{R}^m$. Thus, we expect that it might be helpful for other applications of Stein's method for multivariate normal approximation as well.

In all our main results we provide explicit constants, which are sometimes very large. In part, this is caused by some generous estimates in our proofs, in order to obtain relatively short bounds valid for all choices of m and ℓ and to simplify the proofs. We expect that one can obtain better constants for many instances if one goes back to our proofs and uses the particular stucture of the functionals and the choices of m and ℓ.

1.5 Structure of the paper

This paper is organized as follows. The next section establishes some smoothing lemmas and bounds on solutions of the multivariate Stein equation, including the afore-mentioned key Proposition 2.4. Section 3, which draws on the auxiliary results of Section 2, is devoted to the proofs of our main results. Section 4 deals with the application of our findings to first order Wiener-Itô integrals and intrinsic volumes of Boolean models. Moreover, we further evaluate our results for the case of marked Poisson processes - a result which will be used in the companion paper [31]. In the Appendix we recall the definitions of the Malliavin operators as well as some results from Malliavin calculus on the Poisson space that are used in Section 3.

2 Smoothing and the multivariate Stein equation

2.1 Smoothing lemmas for the d_{convex}- and the $d_{\mathbb{H}_\ell}$-distance

Let $m \in \mathbb{N}$ be fixed in the sequel. Let φ_Σ denote the density of an m-dimensional centered Gaussian random vector N_Σ with a positive definite covariance matrix $\Sigma = (\sigma_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$. Given measurable and bounded $h : \mathbb{R}^m \to \mathbb{R}$, positive definite $\Sigma \in \mathbb{R}^{m\times m}$, and $t \in (0,1)$ we introduce the smoothed version

$$h_{t,\Sigma}(y) := \int_{\mathbb{R}^m} h(\sqrt{t}z + \sqrt{1-t}y)\,\varphi_\Sigma(z)\,\mathrm{d}z = \mathbb{E}\,h(\sqrt{t}N_\Sigma + \sqrt{1-t}y), \quad y \in \mathbb{R}^m,$$

of h, extending (1.15) to general Σ. The following so-called smoothing lemma (see Lemma 2.11 in [10], Lemma 11.4 in [2] or Lemma 12.1 of [6]) allows one to bound the d_{convex}-distance to the m-dimensional standard Gaussian random vector N_I in terms of smooth test functions.

Lemma 2.1. *For an m-dimensional random vector Y and $t \in (0,1)$,*

$$d_{convex}(Y, N_I) \le \frac{4}{3}\sup_{h\in\mathcal{I}_m} |\mathbb{E}\,h_{t,I}(Y) - \mathbb{E}\,h_{t,I}(N_I)| + \frac{20}{\sqrt{\pi}}m^2\frac{\sqrt{t}}{1-t}.$$

Proof. This is the statement of [10, Lemma 2.11] with $\varepsilon = \sqrt{t}$, $\Delta = 2\sqrt{2/\pi}m^{3/2}$ (which can be deduced from [28, Lemma 1]) and $a_m \le 2\sqrt{2m}$ (which follows from Markov's inequality) there. □

Lemma 2.1 is the starting point for proving the asserted bound (1.11). Lemma 2.1 in fact holds for any m-dimensional centered Gaussian random vector N_Σ with positive definite covariance matrix $\Sigma \in \mathbb{R}^{m\times m}$ as implied by the following result.

Lemma 2.2. *For any regular $\Theta \in \mathbb{R}^{m\times m}$, $\ell \in \mathbb{N}$, positive definite $\Sigma \in \mathbb{R}^{m\times m}$, and m-dimensional random vectors Y and Z,*

$$d_{\mathbb{H}_\ell}(Y,Z) = d_{\mathbb{H}_\ell}(\Theta Y, \Theta Z), \quad d_{convex}(Y,Z) = d_{convex}(\Theta Y, \Theta Z),$$

and

$$\sup_{h\in\mathcal{G}} |\mathbb{E}\, h_{t,\Sigma}(\Theta Y) - \mathbb{E}\, h_{t,\Sigma}(\Theta Z)| = \sup_{h\in\mathcal{G}} |\mathbb{E}\, h_{t,\Theta^{-1}\Sigma(\Theta^{-1})^T}(Y) - \mathbb{E}\, h_{t,\Theta^{-1}\Sigma(\Theta^{-1})^T}(Z)|$$

for $t \in (0,1)$ and $\mathcal{G} = \mathbb{H}_\ell$ or $\mathcal{G} = \mathcal{I}_m$.

Proof. For any $h \in \mathbb{H}_\ell$ (resp. $h \in \mathcal{I}_m$) the functions $h_\Theta : \mathbb{R}^m \ni x \mapsto h(\Theta x)$ and $h_{\Theta^{-1}} : \mathbb{R}^m \ni x \mapsto h(\Theta^{-1}x)$ also belong to $\mathbb{H}_\ell$ (resp. $\mathcal{I}_m$) and

$$h_{t,\Sigma}(\Theta x) = \mathbb{E}\, h(\sqrt{t}N_\Sigma + \sqrt{1-t}\Theta x) = \mathbb{E}\, h_\Theta(\sqrt{t}\Theta^{-1}N_\Sigma + \sqrt{1-t}x) = (h_\Theta)_{t,\Theta^{-1}\Sigma(\Theta^{-1})^T}(x),$$

which yields the desired conclusions. □

One of the assumptions behind Lemma 2.1 is that, for any $\varepsilon > 0$, the class $\mathcal{I}_m$ is closed under the supremum and infimum operations sending $h \in \mathcal{I}_m$ to $h_\varepsilon^+(x) := \sup_{y\in B^m(\mathbf{0},\varepsilon)} h(x+y)$, $x \in \mathbb{R}^m$, and $h_\varepsilon^-(x) := \inf_{y\in B^m(\mathbf{0},\varepsilon)} h(x+y)$, $x \in \mathbb{R}^m$, respectively. Here and elsewhere $B^m(x,r)$ denotes the closed Euclidean ball centered at $x \in \mathbb{R}^m$ and having radius r. The class $\mathbb{H}_\ell$, $\ell \geq 2$, is not closed under these operations, and therefore we may not replace $\mathcal{I}_m$ with $\mathbb{H}_\ell$, $\ell \geq 2$, in Lemma 2.1. To proceed for the $d_{\mathbb{H}_\ell}$-distance similarly as for the d_{convex}-distance, we establish an analogous smoothing lemma, whose proof goes along the lines of the proof of Lemma 11.4 in [2].

Lemma 2.3. *For $\ell \in \mathbb{N}$, $\Sigma \in \mathbb{R}^{m\times m}$ positive definite, an m-dimensional random vector Z, and $t \in (0,1)$,*

$$d_{\mathbb{H}_\ell}(Z, N_\Sigma) \leq 2 \sup_{h\in\mathbb{H}_\ell} |\mathbb{E}\, h_{t,\Sigma}(Z) - \mathbb{E}\, h_{t,\Sigma}(N_\Sigma)| + \frac{24\ell\sqrt{m}}{\sqrt{\pi}}\sqrt{t}.$$

Proof. By definition any $h \in \mathbb{H}_\ell$ can be written as

$$h(x) = \mathbf{1}\{\langle x, u_i\rangle \leq z_i, i \in \{1,\dots,\ell\}\}, \quad x \in \mathbb{R}^m,$$

with $u_1,\dots,u_\ell \in \mathbb{S}^{m-1}$ and $z_1,\dots,z_\ell \in \mathbb{R}$, and where $\mathbb{S}^{m-1}$ is the boundary of $B^m(\mathbf{0},1)$. By Lemma 2.2, we can assume without loss of generality that Σ is the identity matrix $I \in \mathbb{R}^{m\times m}$. For $v \in \mathbb{R}$ and $h \in \mathbb{H}_\ell$ as given above, we define $h^{(v)} : \mathbb{R}^m \ni x \mapsto \mathbf{1}\{\langle x, u_i\rangle \leq z_i + v, i \in \{1,\dots,\ell\}\}$, which is also in $\mathbb{H}_\ell$. Note that, for $x, y \in \mathbb{R}^m$ with $\|x\| \leq |v|$,

$$h^{(-|v|)}(x+y) \leq h(y) \leq h^{(|v|)}(x+y) \tag{2.1}$$

and that, for $x, y \in \mathbb{R}^m$,

$$|h^{(v)}(x+y) - h(y)| \leq \sum_{i=1}^{\ell} \mathbf{1}\{z_i - \|x\| - |v| \leq \langle u_i, y\rangle \leq z_i + \|x\| + |v|\}. \tag{2.2}$$

A straightforward computation shows that, for all $w \in \mathbb{R}$,

$$\sup_{u\in\mathbb{S}^{m-1}, z\in\mathbb{R}} \mathbb{P}(-w \leq \langle u, N_I\rangle - z \leq w) \leq \frac{\sqrt{2}}{\sqrt{\pi}} w. \tag{2.3}$$

Let N_I' denote an independent copy of N_I. Without loss of generality we can assume that Z, N_I, and N_I' are defined on the same probability space and are independent. Note that

$$\mathbb{P}(N_I' \in B^m(\mathbf{0}, 2\sqrt{m})) = 1 - \mathbb{P}(\|N_I'\| \geq 2\sqrt{m}) \geq 1 - \frac{\mathbb{E}\,\|N_I'\|^2}{4m} = \frac{3}{4}. \tag{2.4}$$

Let $h \in \mathbb{H}_\ell$ and assume that $\mathbb{E}\, h(\sqrt{1-t}Z) - \mathbb{E}\, h(\sqrt{1-t}N_I) \geq 0$. Then, it follows from the definition of $h_{t,I}$ and (2.1) that

$$\begin{aligned}
&\mathbb{E}\,(h^{(2\sqrt{mt})})_{t,I}(Z) - \mathbb{E}\,(h^{(2\sqrt{mt})})_{t,I}(N_I) \\
&= \mathbb{E}\left[h^{(2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}Z) - h^{(2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I)\right] \\
&\geq \mathbb{E}\left[\mathbf{1}\{N_I' \in B^m(\mathbf{0}, 2\sqrt{m})\}\big(h(\sqrt{1-t}Z) - h(\sqrt{1-t}N_I) \right.\\
&\qquad\qquad\qquad\left. - (h^{(2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I) - h(\sqrt{1-t}N_I))\big)\right] \\
&\quad + \mathbb{E}\left[\mathbf{1}\{N_I' \notin B^m(\mathbf{0}, 2\sqrt{m})\}\big(h(\sqrt{t}N_I' + \sqrt{1-t}Z) - h(\sqrt{t}N_I' + \sqrt{1-t}N_I) \right.\\
&\qquad\qquad\qquad\left. - (h^{(2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I) - h(\sqrt{t}N_I' + \sqrt{1-t}N_I))\big)\right].
\end{aligned}$$

By independence and (2.4) we have that

$$\mathbb{E}\left[\mathbf{1}\{N_I' \in B^m(\mathbf{0}, 2\sqrt{m})\}\big(h(\sqrt{1-t}Z) - h(\sqrt{1-t}N_I)\big)\right] \geq \frac{3}{4}\mathbb{E}\left[h(\sqrt{1-t}Z) - h(\sqrt{1-t}N_I)\right].$$

The observation that, for $h \in \mathbb{H}_\ell$, $p \in \mathbb{R}$, and $z \in \mathbb{R}^m$, $\mathbb{R}^m \ni x \mapsto h(px+z)$ also belongs to $\mathbb{H}_\ell$ and (2.4) yield

$$\mathbb{E}\left[\mathbf{1}\{N_I' \notin B^m(\mathbf{0}, 2\sqrt{m})\}\big(h(\sqrt{t}N_I' + \sqrt{1-t}Z) - h(\sqrt{t}N_I' + \sqrt{1-t}N_I)\big)\right] \geq -\frac{1}{4} d_{\mathbb{H}_\ell}(Z, N_I).$$

From (2.2) and (2.3) we obtain

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}\{N_I' \in B^m(\mathbf{0}, 2\sqrt{m})\}\big(h^{(2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I) - h(\sqrt{1-t}N_I)\big)\right] \\
&\leq \sum_{i=1}^{\ell} \mathbb{P}(z_i - 4\sqrt{mt} \leq \langle u_i, \sqrt{1-t}N_I\rangle \leq z_i + 4\sqrt{mt}) \leq \frac{4\sqrt{2}}{\sqrt{\pi}} \ell\sqrt{m} \frac{\sqrt{t}}{\sqrt{1-t}}
\end{aligned}$$

and that

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}\{N_I' \notin B^m(\mathbf{0}, 2\sqrt{m})\}(h^{(2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I) - h(\sqrt{t}N_I' + \sqrt{1-t}N_I))\right] \\
&\le \sum_{i=1}^{\ell} \mathbb{P}(z_i - \langle u_i, \sqrt{t}N_I'\rangle - 2\sqrt{mt} \le \langle u_i, \sqrt{1-t}N_I\rangle \le z_i - \langle u_i, \sqrt{t}N_I'\rangle + 2\sqrt{mt}) \\
&\le \frac{2\sqrt{2}}{\sqrt{\pi}} \ell\sqrt{m} \frac{\sqrt{t}}{\sqrt{1-t}}.
\end{aligned}$$

For $t \in (0,1)$ let us use the abbreviation

$$d_{\mathbb{H}_\ell, t}(Z, N_I) := \sup_{h \in \mathbb{H}_\ell} |\mathbb{E}\, h_{t,I}(Z) - \mathbb{E}\, h_{t,I}(N_I)|.$$

The previous estimates may be combined to give

$$\begin{aligned}
&\frac{3}{4}\left|\mathbb{E}\, h(\sqrt{1-t}Z) - \mathbb{E}\, h(\sqrt{1-t}N_I)\right| \\
&\le d_{\mathbb{H}_\ell, t}(Z, N_I) + \frac{1}{4} d_{\mathbb{H}_\ell}(Z, N_I) + \frac{6\sqrt{2}}{\sqrt{\pi}} \ell\sqrt{m} \frac{\sqrt{t}}{\sqrt{1-t}}.
\end{aligned} \tag{2.5}$$

On the other hand, if $\mathbb{E}\, h(\sqrt{1-t}Z) - \mathbb{E}\, h(\sqrt{1-t}N_I) < 0$, we obtain by arguments similar to those above that

$$\begin{aligned}
&\mathbb{E}\,(h^{(-2\sqrt{mt})})_{t,I}(N_I) - \mathbb{E}\,(h^{(-2\sqrt{mt})})_{t,I}(Z) \\
&= \mathbb{E}\left[h^{(-2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I) - h^{(-2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}Z)\right] \\
&\ge \mathbb{E}\Big[\mathbf{1}\{N_I' \in B^m(\mathbf{0}, 2\sqrt{m})\}\big(h(\sqrt{1-t}N_I) - h(\sqrt{1-t}Z) \\
&\qquad\qquad - (h(\sqrt{1-t}N_I) - h^{(-2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I))\big)\Big] \\
&\quad + \mathbb{E}\Big[\mathbf{1}\{N_I' \notin B^m(\mathbf{0}, 2\sqrt{m})\}\big(h(\sqrt{t}N_I' + \sqrt{1-t}N_I) - h(\sqrt{t}N_I' + \sqrt{1-t}Z) \\
&\qquad\qquad - (h(\sqrt{t}N_I' + \sqrt{1-t}N_I) - h^{(-2\sqrt{mt})}(\sqrt{t}N_I' + \sqrt{1-t}N_I))\big)\Big] \\
&\ge \frac{3}{4}\big(\mathbb{E}\, h(\sqrt{1-t}N_I) - \mathbb{E}\, h(\sqrt{1-t}Z)\big) - \frac{1}{4} d_{\mathbb{H}_\ell}(Z, N_I) - \frac{6\sqrt{2}}{\sqrt{\pi}} \ell\sqrt{m} \frac{\sqrt{t}}{\sqrt{1-t}}.
\end{aligned}$$

This implies that (2.5) is still true. Taking the supremum over $h \in \mathbb{H}_\ell$ in (2.5), we obtain

$$d_{\mathbb{H}_\ell}(Z, N_I) \le 2 d_{\mathbb{H}_\ell, t}(Z, N_I) + \frac{12\sqrt{2}}{\sqrt{\pi}} \ell\sqrt{m} \frac{\sqrt{t}}{\sqrt{1-t}}.$$

Since the left-hand side is at most one, we may replace $\sqrt{t}/\sqrt{1-t}$ by $\sqrt{2t}$, which completes the proof. □

2.2 Bounds on the derivatives of solutions to Stein's equation for multivariate normal approximation

We extend the definition of $f_{t,h,I}$ given at (1.16) to include indices with general covariance matrix Σ. This goes as follows. For $h : \mathbb{R}^m \to \mathbb{R}$ measurable and bounded, $\Sigma =$

$(\sigma_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$ positive definite, and $t \in (0,1)$, the function $f_{t,h,\Sigma} : \mathbb{R}^m \to \mathbb{R}$ given by

$$f_{t,h,\Sigma}(y) := \frac{1}{2}\int_t^1 \frac{1}{1-s}\int_{\mathbb{R}^m} (h(\sqrt{s}z+\sqrt{1-s}y)-h(z))\,\varphi_\Sigma(z)\,\mathrm{d}z\,\mathrm{d}s, \quad y\in\mathbb{R}^m,$$

is a solution of the Stein equation

$$h_{t,\Sigma}(y) - \mathbb{E}\,h_{t,\Sigma}(N_\Sigma) = \sum_{i=1}^m y_i \frac{\partial f}{\partial y_i}(y) - \sum_{i,j=1}^m \sigma_{ij}\frac{\partial^2 f}{\partial y_i \partial y_j}(y), \quad y\in\mathbb{R}^m,$$

see [10, p. 726] and [6, p. 337] for $\Sigma = I$ as well as [18, Lemma 1] and [19, Lemma 3.3] for general Σ. Some calculations show that, for $i,j,k\in\{1,\dots,m\}$ and $y\in\mathbb{R}^m$,

$$\frac{\partial f_{t,h,\Sigma}}{\partial y_i}(y) = -\frac{1}{2}\int_t^1 \frac{1}{\sqrt{s}\sqrt{1-s}}\int_{\mathbb{R}^m} h(\sqrt{s}z+\sqrt{1-s}y)\,\frac{\partial \varphi_\Sigma}{\partial y_i}(z)\,\mathrm{d}z\,\mathrm{d}s,$$

$$\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(y) = \frac{1}{2}\int_t^1 \frac{1}{s}\int_{\mathbb{R}^m} h(\sqrt{s}z+\sqrt{1-s}y)\,\frac{\partial^2 \varphi_\Sigma}{\partial y_i\partial y_j}(z)\,\mathrm{d}z\,\mathrm{d}s, \tag{2.6}$$

and

$$\frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(y) = -\frac{1}{2}\int_t^1 \frac{\sqrt{1-s}}{s^{3/2}}\int_{\mathbb{R}^m} h(\sqrt{s}z+\sqrt{1-s}y)\,\frac{\partial^3 \varphi_\Sigma}{\partial y_i\partial y_j\partial y_k}(z)\,\mathrm{d}z\,\mathrm{d}s. \tag{2.7}$$

Let $\Sigma^{1/2}$ be the positive definite matrix in $\mathbb{R}^{m\times m}$ such that $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$ and let $\Sigma^{-1/2} := (\Sigma^{1/2})^{-1}$. By $h\circ\Sigma^{1/2}$ we denote the function $\mathbb{R}^m \ni y \mapsto h(\Sigma^{1/2}y)$. It follows from the definition of $f_{t,h,\Sigma}$ that, for $y\in\mathbb{R}^m$,

$$\begin{aligned} f_{t,h,\Sigma}(y) &= \frac{1}{2}\int_t^1 \frac{1}{1-s}\mathbb{E}\,[h(\sqrt{s}N_\Sigma + \sqrt{1-s}y) - h(N_\Sigma)]\,\mathrm{d}s \\ &= \frac{1}{2}\int_t^1 \frac{1}{1-s}\mathbb{E}\,[h\circ\Sigma^{1/2}(\sqrt{s}N_I + \sqrt{1-s}\Sigma^{-1/2}y) - h\circ\Sigma^{1/2}(N_I)]\,\mathrm{d}s \\ &= f_{t,h\circ\Sigma^{1/2},I}(\Sigma^{-1/2}y). \end{aligned} \tag{2.8}$$

Since $\varphi_\Sigma(z) = \varphi_I(\Sigma^{-1/2}z)/\sqrt{\det(\Sigma)}$ for $z\in\mathbb{R}^m$, we have that, for $i,j,k\in\{1,\dots,m\}$ and $z\in\mathbb{R}^m$,

$$\frac{\partial^3\varphi_\Sigma}{\partial y_i\partial y_j\partial y_k}(z) = \frac{1}{\sqrt{\det(\Sigma)}}\sum_{u,v,w=1}^m (\Sigma^{-1/2})_{ui}(\Sigma^{-1/2})_{vj}(\Sigma^{-1/2})_{wk}\frac{\partial^3\varphi_I}{\partial y_u\partial y_v\partial y_w}(\Sigma^{-1/2}z),$$

which yields together with a short computation

$$\sum_{i,j,k=1}^m \left(\frac{\partial^3\varphi_\Sigma}{\partial y_i\partial y_j\partial y_k}(z)\right)^2 \le \frac{\|\Sigma^{-1}\|_{op}^3}{\det(\Sigma)}\sum_{i,j,k=1}^m \left(\frac{\partial^3\varphi_I}{\partial y_i\partial y_j\partial y_k}(\Sigma^{-1/2}z)\right)^2. \tag{2.9}$$

From the above formulas for the derivatives of $f_{t,h,\Sigma}$ one can deduce that

$$\sup_{y\in\mathbb{R}^m}\left|\frac{\partial^2 f_{t,h,\Sigma}(y)}{\partial y_i\partial y_j}\right| \le m^2\|\Sigma^{-1}\|_{op}\|h\|_\infty|\log t|, \quad t\in(0,1),$$

and

$$\sup_{y\in\mathbb{R}^m}\left|\frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i\partial y_j\partial y_k}(y)\right| \le 6m^3\|\Sigma^{-1}\|_{op}^{3/2}\|h\|_\infty\frac{1}{\sqrt{t}}, \quad t\in(0,1). \tag{2.10}$$

These sup norm bounds on the derivatives of $f_{t,h,\Sigma}$ go hand-in-hand with the following more useful second moment bounds. They are the key to the proofs of our main results, as they will be used to bound the right-hand sides of the smoothing inequalities in Lemma 2.1 and Lemma 2.3.

Proposition 2.4. *Let Y be an m-dimensional random vector, let $\Sigma\in\mathbb{R}^{m\times m}$ be positive definite, and define*

$$M_2 := \frac{1}{4}\sum_{i,j=1}^m\left(\int_{\mathbb{R}^m}\left|\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\right|\mathrm{d}z\right)^2 \le m^2. \tag{2.11}$$

Then,

$$\sup_{h\in\mathbb{H}_\ell}\mathbb{E}\sum_{i,j=1}^m\left(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i\partial y_j}(Y)\right)^2 \le \|\Sigma^{-1}\|_{op}^2(M_2(\log t)^2 d_{\mathbb{H}_{2\ell}}(Y,N_\Sigma)+444m^{23/6})$$

for all $t\in(0,1)$ and $\ell\in\mathbb{N}$ and

$$\sup_{h\in\mathcal{I}_m}\mathbb{E}\sum_{i,j=1}^m\left(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i\partial y_j}(Y)\right)^2 \le \|\Sigma^{-1}\|_{op}^2(M_2(\log t)^2 d_{convex}(Y,N_\Sigma)+444m^{23/6})$$

for all $t\in(0,1)$.

We prepare the proof of Proposition 2.4 with the following lemmas. For $x\in\mathbb{R}^m$ and a Borel set $B\subseteq\mathbb{R}^m$ we define $d(x,B):=\inf_{y\in B}\|x-y\|$.

Lemma 2.5. *For any $\alpha\in(0,1)$,*

$$\sup_{A\subseteq\mathbb{R}^m \text{ convex}}\mathbb{E}\frac{1}{d(N_I,\partial A)^\alpha} \le 1+2\sqrt{\frac{2}{\pi}}m^{3/2}\frac{\alpha}{1-\alpha}.$$

Proof. It is shown in [28, Lemma 1] that for, all convex $A\subseteq\mathbb{R}^m$ and $r>0$,

$$\mathbb{P}(d(N_I,\partial A)\le r)\le 2\sqrt{\frac{2}{\pi}}m^{3/2}r. \tag{2.12}$$

This implies for any convex $A \subseteq \mathbb{R}^m$ that

$$\begin{aligned}
\mathbb{E}\,\frac{1}{d(N_I,\partial A)^\alpha} &= \int_0^\infty \mathbb{P}(d(N_I,\partial A)^{-\alpha} \geq u)\,\mathrm{d}u = \int_0^\infty \mathbb{P}(d(N_I,\partial A) \leq u^{-1/\alpha})\,\mathrm{d}u \\
&\leq 1 + \int_1^\infty \mathbb{P}(d(N_I,\partial A) \leq u^{-1/\alpha})\,\mathrm{d}u \\
&\leq 1 + 2\sqrt{\frac{2}{\pi}}m^{3/2}\int_1^\infty u^{-1/\alpha}\,\mathrm{d}u = 1 + 2\sqrt{\frac{2}{\pi}}m^{3/2}\frac{\alpha}{1-\alpha},
\end{aligned}$$

which completes the proof. □

Lemma 2.6. *For any positive definite* $\Sigma \in \mathbb{R}^{m\times m}$ *and* $i,j \in \{1,\dots,m\}$,

$$\int_{\mathbb{R}^m} \frac{\partial^2 \varphi_\Sigma}{\partial y_i \partial y_j}(z)\,\mathrm{d}z = 0.$$

Proof. As noted at display (12.72) of [6] we have that the integral of the mixed derivative $\frac{\partial^2\varphi_\Sigma}{\partial y_i\partial y_j}(z)$ is the mixed derivative of $x \mapsto \int_{\mathbb{R}^m}\varphi_\Sigma(z+x)\,\mathrm{d}z$ evaluated at $x = \mathbf{0}$. The integral is one, so the derivative vanishes. □

Lemma 2.7. *For all* $h \in \mathcal{I}_m$ *and* $t \in (0,1)$,

$$\max_{i,j\in\{1,\dots,m\}} \mathbb{E}\left(\frac{\partial^2 f_{t,h,I}}{\partial y_i\partial y_j}(N_I)\right)^2 \leq 444m^{11/6}.$$

Proof. We can assume that $h = \mathbf{1}\{\cdot \in A\}$ for some closed convex set $A \subseteq \mathbb{R}^m$. Then, for $i,j \in \{1,\dots,m\}$ and $y \in \mathbb{R}^m$, it follows from (2.6) that

$$\begin{aligned}
\frac{\partial^2 f_{t,h,I}}{\partial y_i\partial y_j}(y) &= \frac{1}{2}\int_t^1 \frac{1}{s}\int_{\mathbb{R}^m} \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}y \in A\}\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\,\mathrm{d}z\,\mathrm{d}s \\
&= \frac{1}{2}\int_t^1 \frac{1}{s}\int_{\mathbb{R}^m} \mathbf{1}\{z \in \frac{1}{\sqrt{s}}(A - \sqrt{1-s}y)\}\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\,\mathrm{d}z\,\mathrm{d}s.
\end{aligned}$$

For $s \in (0,1)$ and $y \in \mathbb{R}^m$ let $r_{s,y} := d(\mathbf{0}, \partial(\frac{1}{\sqrt{s}}(A - \sqrt{1-s}y))) = \frac{1}{\sqrt{s}}d(\sqrt{1-s}y, \partial A)$. If $\mathbf{0} \notin \frac{1}{\sqrt{s}}(A - \sqrt{1-s}y)$, we have

$$\left|\int_{\mathbb{R}^m} \mathbf{1}\{z \in \frac{1}{\sqrt{s}}(A - \sqrt{1-s}y)\}\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\,\mathrm{d}z\right| \leq \int_{\mathbb{R}^m\setminus B^m(\mathbf{0},r_{s,y})}\left|\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\right|\mathrm{d}z.$$

If $\mathbf{0} \in \frac{1}{\sqrt{s}}(A - \sqrt{1-s}y)$, Lemma 2.6 implies that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^m} \mathbf{1}\{z \in \frac{1}{\sqrt{s}}(A - \sqrt{1-s}y)\}\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\,\mathrm{d}z\right| \\
&= \left|\int_{\mathbb{R}^m} \mathbf{1}\{z \notin \frac{1}{\sqrt{s}}(A - \sqrt{1-s}y)\}\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\,\mathrm{d}z\right| \leq \int_{\mathbb{R}^m\setminus B^m(\mathbf{0},r_{s,y})}\left|\frac{\partial^2\varphi_I}{\partial y_i\partial y_j}(z)\right|\mathrm{d}z.
\end{aligned}$$

Since for the density ϕ of a standard Gaussian random variable and $a \in \mathbb{R}$ it holds that

$$|\phi'(a)| = \frac{1}{\sqrt{2\pi}}|a|e^{-a^2/2} \leq \frac{1}{\sqrt{2\pi}} \underbrace{|ae^{-a^2/4}|}_{\leq 1} e^{-a^2/4} \leq \frac{\sqrt{2}}{\sqrt{4\pi}} e^{-a^2/4}$$

and

$$|\phi''(a)| = \frac{1}{\sqrt{2\pi}}|a^2-1|e^{-a^2/2} \leq \frac{1}{\sqrt{2\pi}} \underbrace{|(a^2-1)e^{-a^2/4}|}_{\leq 2} e^{-a^2/4} \leq \frac{2^{3/2}}{\sqrt{4\pi}} e^{-a^2/4},$$

we obtain that

$$\left| \frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z) \right| \leq 2^{3/2} \varphi_{I_{i,j}}(z), \quad z \in \mathbb{R}^m,$$

where $I_{i,j}$ is the identity matrix I where the i-th and the j-th diagonal element are replaced by 2. Consequently, we have

$$\left| \int_{\mathbb{R}^m} \mathbf{1}\{z \in \frac{1}{\sqrt{s}}(A - \sqrt{1-s}y)\} \frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z) \, \mathrm{d}z \right| \leq 2^{3/2} \mathbb{P}(\|N_{I_{i,j}}\| \geq r_{s,y}).$$

The Markov inequality yields

$$\mathbb{P}(\|N_{I_{i,j}}\| \geq r_{s,y}) \leq \frac{\mathbb{E}\,\|N_{I_{i,j}}\|^{1/3}}{r_{s,y}^{1/3}} \leq \frac{s^{1/6}(\mathbb{E}\,\|N_{I_{i,j}}\|^2)^{1/6}}{d(\sqrt{1-s}y, \partial A)^{1/3}} \leq \frac{2^{1/6} m^{1/6} s^{1/6}}{(1-s)^{1/6} d(y, \partial A/\sqrt{1-s})^{1/3}}.$$

Hence, we obtain

$$\left| \frac{\partial^2 f_{t,h,I}}{\partial y_i \partial y_j}(y) \right| \leq 2^{2/3} m^{1/6} \int_t^1 \frac{1}{s^{5/6}(1-s)^{1/6}} \frac{1}{d(y, \partial A/\sqrt{1-s})^{1/3}} \, \mathrm{d}s, \quad y \in \mathbb{R}^m.$$

Now the Cauchy-Schwarz inequality leads to

$$\left(\frac{\partial^2 f_{t,h,I}}{\partial y_i \partial y_j}(y) \right)^2 \leq 2^{4/3} m^{1/3} \int_t^1 \frac{1}{s^{5/6}(1-s)^{1/3}} \, \mathrm{d}s \int_t^1 \frac{1}{s^{5/6}} \frac{1}{d(y, \partial A/\sqrt{1-s})^{2/3}} \, \mathrm{d}s, \quad y \in \mathbb{R}^m.$$

Numerical integration shows that the first integral may be generously bounded by 7 so that we obtain, together with Lemma 2.5,

$$\begin{aligned}
\mathbb{E} \left(\frac{\partial^2 f_{t,h,I}}{\partial y_i \partial y_j}(N_I) \right)^2 &\leq 7 \cdot 2^{4/3} m^{1/3} \int_t^1 \frac{1}{s^{5/6}} \mathbb{E} \, \frac{1}{d(N_I, \partial A/\sqrt{1-s})^{2/3}} \, \mathrm{d}s \\
&\leq 7 \cdot 2^{4/3} m^{1/3} \int_t^1 \frac{1}{s^{5/6}} \, \mathrm{d}s \sup_{A' \subseteq \mathbb{R}^m \text{ convex}} \mathbb{E} \, \frac{1}{d(N_I, \partial A')^{2/3}} \\
&\leq 42 \cdot 2^{4/3} m^{1/3} (1 + 4\sqrt{2/\pi} m^{3/2}) \\
&\leq 444 m^{11/6},
\end{aligned}$$

which completes the proof. □

Proof of Proposition 2.4. First we prove the assertion for the special case $\Sigma = I$. For $i, j \in \{1, \ldots, m\}$ we have that

$$
\begin{aligned}
&\mathbb{E}\left(\frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(Y)\right)^2 \\
&= \mathbb{E}\left(\frac{1}{2}\int_t^1 \frac{1}{s}\int_{\mathbb{R}^m} h(\sqrt{s}z + \sqrt{1-s}Y)\frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z)\,\mathrm{d}z\,\mathrm{d}s\right)^2 \\
&= \frac{1}{4}\int_t^1 \int_t^1 \frac{1}{s_1 s_2}\int_{\mathbb{R}^m}\int_{\mathbb{R}^m} \mathbb{E}\, h(\sqrt{s_1}z_1 + \sqrt{1-s_1}Y)h(\sqrt{s_2}z_2 + \sqrt{1-s_2}Y) \\
&\qquad\qquad \frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z_1)\frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z_2)\,\mathrm{d}z_2\,\mathrm{d}z_1\,\mathrm{d}s_2\,\mathrm{d}s_1 \\
&= \frac{1}{4}\int_t^1 \int_t^1 \frac{1}{s_1 s_2}\int_{\mathbb{R}^m}\int_{\mathbb{R}^m} \mathbb{E}\, h(\sqrt{s_1}z_1 + \sqrt{1-s_1}N_I)h(\sqrt{s_2}z_2 + \sqrt{1-s_2}N_I) \\
&\qquad\qquad \frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z_1)\frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z_2)\,\mathrm{d}z_2\,\mathrm{d}z_1\,\mathrm{d}s_2\,\mathrm{d}s_1 \\
&\quad + \frac{1}{4}\int_t^1 \int_t^1 \frac{1}{s_1 s_2}\int_{\mathbb{R}^m}\int_{\mathbb{R}^m} \Big(\mathbb{E}\, h(\sqrt{s_1}z_1 + \sqrt{1-s_1}Y)h(\sqrt{s_2}z_2 + \sqrt{1-s_2}Y) \\
&\qquad\qquad - \mathbb{E}\, h(\sqrt{s_1}z_1 + \sqrt{1-s_1}N_I)h(\sqrt{s_2}z_2 + \sqrt{1-s_2}N_I)\Big) \\
&\qquad\qquad \frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z_1)\frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(z_2)\,\mathrm{d}z_2\,\mathrm{d}z_1\,\mathrm{d}s_2\,\mathrm{d}s_1 \\
&= \mathbb{E}\left(\frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(N_I)\right)^2 + R_{ij},
\end{aligned}
$$

where R_{ij} denotes the second four-fold integral in the penultimate equation. It follows from Lemma 2.7 that

$$
\mathbb{E}\sum_{i,j=1}^m \left(\frac{\partial^2 \varphi_I}{\partial y_i \partial y_j}(N_I)\right)^2 \leq 444 m^{23/6}.
$$

For $h \in \mathbb{H}_\ell$, $z_1, z_2 \in \mathbb{R}^m$, and $s_1, s_2 \in (0,1)$ we have that

$$
h_{z_1,z_2,s_1,s_2} : \mathbb{R}^m \ni y \mapsto h(\sqrt{s_1}z_1 + \sqrt{1-s_1}y)h(\sqrt{s_2}z_2 + \sqrt{1-s_2}y)
$$

is the indicator function of the intersection of 2ℓ closed half-spaces and, thus, an element of $\mathbb{H}_{2\ell}$. This yields that

$$
|\mathbb{E}\, h_{z_1,z_2,s_1,s_2}(Y) - \mathbb{E}\, h_{z_1,z_2,s_1,s_2}(N_I)| \leq d_{\mathbb{H}_{2\ell}}(Y, N_I)
$$

and, consequently, recalling the definition of M_2 at (2.11),

$$
\sum_{i,j=1}^m |R_{ij}| \leq M_2(\log t)^2 d_{\mathbb{H}_{2\ell}}(Y, N_I).
$$

For $h \in \mathcal{I}_m$ we have that $h_{z_1,z_2,s_1,s_2} \in \mathcal{I}_m$, whence

$$\sum_{i,j=1}^{m} |R_{ij}| \leq M_2 (\log t)^2 d_{convex}(Y, N_I).$$

Combining the previous estimates completes the proof of Proposition 2.4 for the special case $\Sigma = I$.

For a positive definite $\Sigma \in \mathbb{R}^{m\times m}$ it follows from (2.8) that, for $y \in \mathbb{R}^m$,

$$\text{Hess}\, f_{t,h,\Sigma}(y) = \Sigma^{-1/2}\, \text{Hess}\, f_{t,h\circ\Sigma^{1/2},I}(\Sigma^{-1/2} y)\Sigma^{-1/2}.$$

Using the Hilbert-Schmidt norm $\|A\|_{H.S.} := \sqrt{\sum_{i,j=1}^m a_{ij}^2}$ of a matrix $A = (a_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$ and the relation that $\|AB\|_{H.S.} \leq \|A\|_{op}\|B\|_{H.S}$ for $A, B \in \mathbb{R}^{m\times m}$, we obtain that

$$\begin{aligned}
\mathbb{E} \sum_{i,j=1}^{m} \left(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(Y) \right)^2 &= \mathbb{E}\, \|\, \text{Hess}\, f_{t,h,\Sigma}(Y)\|_{H.S.}^2 \\
&= \mathbb{E}\, \|\Sigma^{-1/2}\, \text{Hess}\, f_{t,h\circ\Sigma^{1/2},I}(\Sigma^{-1/2}Y)\Sigma^{-1/2}\|_{H.S.}^2 \\
&\leq \|\Sigma^{-1/2}\|_{op}^4 \mathbb{E}\, \|\, \text{Hess}\, f_{t,h\circ\Sigma^{1/2},I}(\Sigma^{-1/2}Y)\|_{H.S.}^2 \\
&= \|\Sigma^{-1}\|_{op}^2 \mathbb{E} \sum_{i,j=1}^{m} \left(\frac{\partial^2 f_{t,h\circ\Sigma^{1/2},I}}{\partial y_i \partial y_j}(\Sigma^{-1/2}Y) \right)^2.
\end{aligned}$$

Now the special case proven above (for $\Sigma = I$) and the observation that, by Lemma 2.2, $d_{\mathbb{H}_\ell}(\Sigma^{-1/2}Y, N_I) = d_{\mathbb{H}_\ell}(Y, N_\Sigma)$ and $d_{convex}(\Sigma^{-1/2}Y, N_I) = d_{convex}(Y, N_\Sigma)$, respectively, complete the proof. $\square$

3 Proofs of the main results

Throughout this section we assume that the reader is familiar with Malliavin calculus on the Poisson space. The Appendix provides the essential definitions and properties of Malliavin operators needed in the sequel.

3.1 Proof of Theorem 1.1

The starting point for the proofs for the d_3- and the d_2-distance are the following quantitative bounds for the normal approximation of Poisson functionals, which were derived in [23, Theorem 4.2] and [23, Theorem 3.3] by a combination of Malliavin calculus with the interpolation method and Stein's method, respectively (see also [4, Section 6]). For a definition of the inverse Ornstein-Uhlenbeck generator L^{-1} we refer to [15, 23] or the Appendix.

Proposition 3.1. *Let $F = (F_1, \dots, F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals $F_1, \dots, F_m \in \operatorname{dom} D$ with $\mathbb{E}\, F_i = 0$, $i \in \{1, \dots, m\}$, let $\Sigma = (\sigma_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$ be positive semi-definite, and put*

$$\beta_1 := \sqrt{\sum_{i,j=1}^{m} \mathbb{E}\left(\sigma_{ij} - \int_{\mathbb{X}} D_x F_i(-D_x L^{-1} F_j)\, \lambda(\mathrm{d}x)\right)^2}$$

$$\beta_2 := \int_{\mathbb{X}} \mathbb{E}\left(\sum_{i=1}^{m} |D_x F_i|\right)^2 \sum_{j=1}^{m} |D_x L^{-1} F_j|\, \lambda(\mathrm{d}x).$$

Then,

$$d_3(F, N_\Sigma) \leq \frac{m}{2}\beta_1 + \frac{1}{4}\beta_2.$$

If, additionally, Σ is positive definite, then

$$d_2(F, N_\Sigma) \leq \|\Sigma^{-1}\|_{op} \|\Sigma\|_{op}^{1/2} \beta_1 + \frac{\sqrt{2\pi}}{8} \|\Sigma^{-1}\|_{op}^{3/2} \|\Sigma\|_{op} \beta_2.$$

The main difficulty in evaluating these bounds is to control the behavior of the terms involving the inverse Ornstein-Uhlenbeck generator L^{-1}, which will be done in the same way as in [15]. The following proposition collects two estimates from [15, Lemma 3.4 and Proposition 4.1], which will play a crucial role in the sequel.

Proposition 3.2. *(a) For a square integrable Poisson functional F and $p \geq 1$,*

$$\mathbb{E}\, |D_x L^{-1} F|^p \leq \mathbb{E}\, |D_x F|^p, \quad \lambda\text{-a.e. } x \in \mathbb{X}$$

and

$$\mathbb{E}\, |D^2_{x,y} L^{-1} F|^p \leq \mathbb{E}\, |D^2_{x,y} F|^p, \quad \lambda^2\text{-a.e. } (x, y) \in \mathbb{X}^2.$$

(b) For $F, G \in \operatorname{dom} D$ with $\mathbb{E}\, F = \mathbb{E}\, G = 0$,

$$\begin{aligned}
&\mathbb{E}\left(\operatorname{Cov}(F, G) - \int_{\mathbb{X}} D_x F(-D_x L^{-1} G)\, \lambda(\mathrm{d}x)\right)^2 \\
&\leq 3 \int_{\mathbb{X}^3} \left[\mathbb{E}\, (D^2_{x_1,x_3} F)^2 (D^2_{x_2,x_3} F)^2\right]^{1/2} \left[\mathbb{E}\, (D_{x_1} G)^2 (D_{x_2} G)^2\right]^{1/2} \lambda^3(\mathrm{d}(x_1, x_2, x_3)) \\
&\quad + \int_{\mathbb{X}^3} \left[\mathbb{E}\, (D_{x_1} F)^2 (D_{x_2} F)^2\right]^{1/2} \left[\mathbb{E}\, (D^2_{x_1,x_3} G)^2 (D^2_{x_2,x_3} G)^2\right]^{1/2} \lambda^3(\mathrm{d}(x_1, x_2, x_3)) \\
&\quad + \int_{\mathbb{X}^3} \left[\mathbb{E}\, (D^2_{x_1,x_3} F)^2 (D^2_{x_2,x_3} F)^2\right]^{1/2} \left[\mathbb{E}\, (D^2_{x_1,x_3} G)^2 (D^2_{x_2,x_3} G)^2\right]^{1/2} \lambda^3(\mathrm{d}(x_1, x_2, x_3)).
\end{aligned}$$

Combining Proposition 3.1 and Proposition 3.2 yields the proof of Theorem 1.1, which goes as follows.

Proof of Theorem 1.1. From the triangle inequality we obtain that

$$\beta_1 \le \sum_{i,j=1}^{m} |\sigma_{ij} - \operatorname{Cov}(F_i, F_j)| + \sqrt{\sum_{i,j=1}^{m} \mathbb{E}\left(\operatorname{Cov}(F_i, F_j) - \int_{\mathbb{X}} D_x F_i(-D_x L^{-1} F_j)\, \lambda(\mathrm{d}x)\right)^2}.$$

Now an application of Proposition 3.2(b) yields that, for $i, j \in \{1, \dots, m\}$,

$$\begin{aligned}
&\mathbb{E}\left(\operatorname{Cov}(F_i, F_j) - \int_{\mathbb{X}} D_x F_i(-D_x L^{-1} F_j)\, \lambda(\mathrm{d}x)\right)^2 \\
&\le 3 \int_{\mathbb{X}^3} \left[\mathbb{E}\,(D^2_{x_1,x_3} F_i)^2 (D^2_{x_2,x_3} F_i)^2\right]^{1/2} \left[\mathbb{E}\,(D_{x_1} F_j)^2 (D_{x_2} F_j)^2\right]^{1/2} \lambda^3(\mathrm{d}(x_1, x_2, x_3)) \\
&\quad + \int_{\mathbb{X}^3} \left[\mathbb{E}\,(D_{x_1} F_i)^2 (D_{x_2} F_i)^2\right]^{1/2} \left[\mathbb{E}\,(D^2_{x_1,x_3} F_j)^2 (D^2_{x_2,x_3} F_j)^2\right]^{1/2} \lambda^3(\mathrm{d}(x_1, x_2, x_3)) \\
&\quad + \int_{\mathbb{X}^3} \left[\mathbb{E}\,(D^2_{x_1,x_3} F_i)^2 (D^2_{x_2,x_3} F_i)^2\right]^{1/2} \left[\mathbb{E}\,(D^2_{x_1,x_3} F_j)^2 (D^2_{x_2,x_3} F_j)^2\right]^{1/2} \lambda^3(\mathrm{d}(x_1, x_2, x_3))
\end{aligned}$$

so that

$$\beta_1 \le \sum_{i,j=1}^{m} |\sigma_{ij} - \operatorname{Cov}(F_i, F_j)| + 2\gamma_1 + \gamma_2. \tag{3.1}$$

It follows from Hölder's inequality and Proposition 3.2(a) that

$$\begin{aligned}
\beta_2 &\le m \int_{\mathbb{X}} \sum_{i=1}^{m} \mathbb{E}\left[|D_x F_i|^3\right]^{2/3} \sum_{j=1}^{m} \mathbb{E}\left[|D_x L^{-1} F_j|^3\right]^{1/3} \lambda(\mathrm{d}x) \\
&\le m \int_{\mathbb{X}} \sum_{i=1}^{m} \mathbb{E}\left[|D_x F_i|^3\right]^{2/3} \sum_{j=1}^{m} \mathbb{E}\left[|D_x F_j|^3\right]^{1/3} \lambda(\mathrm{d}x) \\
&\le m \int_{\mathbb{X}} m^{1/3} \left(\sum_{i=1}^{m} \mathbb{E}\left[|D_x F_i|^3\right]\right)^{2/3} m^{2/3} \left(\sum_{j=1}^{m} \mathbb{E}\left[|D_x F_j|^3\right]\right)^{1/3} \lambda(\mathrm{d}x) \\
&= m^2 \int_{\mathbb{X}} \sum_{i=1}^{m} \mathbb{E}\left[|D_x F_i|^3\right] \lambda(\mathrm{d}x) = m^2 \gamma_3.
\end{aligned}$$

Now Proposition 3.1 completes the proof of Theorem 1.1. □

3.2 Proofs of Theorem 1.2 and Theorem 1.3

Throughout this subsection we use several Malliavin operators, namely the already introduced difference operator D, the inverse Ornstein-Uhlenbeck generator L^{-1} and the Skorohod integral δ. Recall that we denote the domain of D by $\operatorname{dom} D$ and we define $\operatorname{dom} \delta$ similarly. For definitions we refer, for example, to [15, Section 2] or to the Appendix.

Proof of Theorem 1.2. In the following we can assume that $\gamma_1, \gamma_2, \gamma_4, \gamma_5 < \infty$ since otherwise there is nothing to prove. Let $h : \mathbb{R}^m \to \mathbb{R}$ be measurable with $\|h\|_\infty \leq 1$. As noted in Section 1.4, it follows from p. 1498 in [23] that

$$
\begin{aligned}
&|\mathbb{E}\, h_{t,\Sigma}(F) - \mathbb{E}\, h_{t,\Sigma}(N_\Sigma)| \\
&= \Bigg| \sum_{i,j=1}^m \sigma_{ij} \mathbb{E}\, \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) - \sum_{k=1}^m \mathbb{E} \int_{\mathbb{X}} D_x \frac{\partial f_{t,h,\Sigma}}{\partial y_k}(F)(-D_x L^{-1} F_k)\, \lambda(\mathrm{d}x) \Bigg|.
\end{aligned}
$$

The fundamental theorem of calculus yields

$$
\begin{aligned}
&\sum_{k=1}^m \mathbb{E} \int_{\mathbb{X}} D_x \frac{\partial f_{t,h,\Sigma}}{\partial y_k}(F)(-D_x L^{-1} F_k)\, \lambda(\mathrm{d}x) \\
&= \sum_{k=1}^m \mathbb{E} \int_{\mathbb{X}} \int_0^1 \sum_{j=1}^m \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F + u D_x F) D_x F_j (-D_x L^{-1} F_k)\, \mathrm{d}u\, \lambda(\mathrm{d}x) \\
&= \sum_{j,k=1}^m \mathbb{E} \int_{\mathbb{X}} \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F) D_x F_j (-D_x L^{-1} F_k)\, \lambda(\mathrm{d}x) \\
&\quad + \sum_{j,k=1}^m \mathbb{E} \int_{\mathbb{X}} \int_0^1 \bigg(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F + u D_x F) - \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F) \bigg) D_x F_j (-D_x L^{-1} F_k)\, \mathrm{d}u\, \lambda(\mathrm{d}x) \\
&=: J_1 + J_2.
\end{aligned}
$$

The idea of the proof is to decompose J_2 into J_{21} and J_{22} and to show that the terms $\Big| J_1 - \sum_{i,j=1}^m \sigma_{ij} \mathbb{E}\, \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) \Big|$ and $|J_{21}|$ are each bounded by a product involving $\sqrt{\mathbb{E} \sum_{i,j=1}^m \big(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) \big)^2}$ and to then apply Proposition 2.4.

Recalling the definition of β_1 in Proposition 3.1 and applying the Cauchy-Schwarz inequality we obtain

$$
\begin{aligned}
&\bigg| J_1 - \sum_{i,j=1}^m \sigma_{ij} \mathbb{E}\, \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) \bigg| \\
&\leq \sqrt{\mathbb{E} \sum_{i,j=1}^m \bigg(\sigma_{ij} - \int_{\mathbb{X}} D_x F_j (-D_x L^{-1} F_k)\, \lambda(\mathrm{d}x) \bigg)^2} \sqrt{\mathbb{E} \sum_{i,j=1}^m \bigg(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) \bigg)^2} \qquad (3.2) \\
&= \beta_1 \sqrt{\mathbb{E} \sum_{i,j=1}^m \bigg(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) \bigg)^2}.
\end{aligned}
$$

From now on we assume that $h \in \mathbb{H}_\ell$. Now Proposition 2.4 leads to

$$
\sqrt{\mathbb{E} \sum_{i,j=1}^m \bigg(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F) \bigg)^2} \leq \|\Sigma^{-1}\|_{op} \big(\sqrt{M_2} |\log t| \sqrt{d_{\mathbb{H}_{2\ell}}(F, N_\Sigma)} + 22 m^{23/12} \big). \qquad (3.3)
$$

Combining inequalities (3.1)-(3.3) yields

$$\begin{aligned}
&\Big|J_1 - \sum_{i,j=1}^m \sigma_{ij}\mathbb{E}\,\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_i \partial y_j}(F)\Big| \\
&\le \|\Sigma^{-1}\|_{op}\big(\sqrt{M_2}|\log t|\sqrt{d_{\mathbb{H}_{2\ell}}(F,N_\Sigma)} + 22m^{23/12}\big)\Bigg(\sum_{i,j=1}^m |\sigma_{ij} - \mathrm{Cov}(F_i,F_j)| + 2\gamma_1 + \gamma_2\Bigg).
\end{aligned} \tag{3.4}$$

This concludes the estimation related to J_1.

Further applications of the fundamental theorem of calculus yield

$$\begin{aligned}
J_2 &= \sum_{j,k=1}^m \mathbb{E}\int_{\mathbb{X}}\int_0^1 \Bigg(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F+uD_xF) - \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F)\Bigg) D_xF_j(-D_xL^{-1}F_k)\,\mathrm{d}u\,\lambda(\mathrm{d}x) \\
&= \sum_{i,j,k=1}^m \mathbb{E}\int_{\mathbb{X}}\int_0^1\int_0^1 \frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F+vuD_xF)uD_xF_iD_xF_j(-D_xL^{-1}F_k)\,\mathrm{d}v\,\mathrm{d}u\,\lambda(\mathrm{d}x) \\
&= \sum_{i,j,k=1}^m \mathbb{E}\int_{\mathbb{X}}\int_0^1\int_0^1 \frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F+vD_xF)uD_xF_iD_xF_j(-D_xL^{-1}F_k)\,\mathrm{d}v\,\mathrm{d}u\,\lambda(\mathrm{d}x) \\
&\quad + \sum_{i,j,k=1}^m \mathbb{E}\int_{\mathbb{X}}\int_0^1\int_0^1 \Bigg(\frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F+vuD_xF) - \frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F+vD_xF)\Bigg) \\
&\qquad\qquad uD_xF_iD_xF_j(-D_xL^{-1}F_k)\,\mathrm{d}v\,\mathrm{d}u\,\lambda(\mathrm{d}x) \\
&= \frac{1}{2}\sum_{j,k=1}^m \mathbb{E}\int_{\mathbb{X}} \Bigg(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F+D_xF) - \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F)\Bigg) D_xF_j(-D_xL^{-1}F_k)\,\lambda(\mathrm{d}x) \\
&\quad + \sum_{i,j,k=1}^m \mathbb{E}\int_{\mathbb{X}}\int_0^1\int_0^1 \Bigg(\frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F+vuD_xF) - \frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F+vD_xF)\Bigg) \\
&\qquad\qquad uD_xF_iD_xF_j(-D_xL^{-1}F_k)\,\mathrm{d}v\,\mathrm{d}u\,\lambda(\mathrm{d}x) \\
&=: J_{2,1} + J_{2,2}.
\end{aligned}$$

We can rewrite $J_{2,1}$ as

$$J_{2,1} = \frac{1}{2}\sum_{j,k=1}^m \mathbb{E}\int_{\mathbb{X}} D_x\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F)D_xF_j(-D_xL^{-1}F_k)\,\lambda(\mathrm{d}x).$$

All third partial derivatives of $f_{t,h,\Sigma}$ are bounded by some constant (recall (2.10)), and thus

$$\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F) \in \mathrm{dom}\,D, \quad j,k \in \{1,\dots,m\}.$$

From Lemma A.3 and the computation for $\mathbb{E}\,\delta(DF_j(-DL^{-1}F_k))^2$ below, one deduces that $DF_j(-DL^{-1}F_k) \in \mathrm{dom}\,\delta$. It follows from integration by parts (see Lemma A.2) and the Cauchy-Schwarz inequality that

$$
\begin{aligned}
|J_{2,1}| &= \frac{1}{2}\bigg| \sum_{j,k=1}^m \mathbb{E}\, \frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F) \delta(DF_j(-DL^{-1}F_k))\, \lambda(\mathrm{d}x) \bigg| \\
&\leq \frac{1}{2} \bigg(\mathbb{E} \sum_{j,k=1}^m \Big(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F) \Big)^2 \bigg)^{1/2} \bigg(\sum_{j,k=1}^m \mathbb{E}\, \delta(DF_j(-DL^{-1}F_k))^2 \bigg)^{1/2}.
\end{aligned}
$$

Concerning the first factor Proposition 2.4 implies that

$$
\bigg(\mathbb{E} \sum_{j,k=1}^m \Big(\frac{\partial^2 f_{t,h,\Sigma}}{\partial y_j \partial y_k}(F) \Big)^2 \bigg)^{1/2} \leq \|\Sigma^{-1}\|_{op} \big(\sqrt{M_2} |\log t| \sqrt{d_{\mathbb{H}_{2\ell}}(F, N_\Sigma)} + 22 m^{23/12} \big).
$$

For the summands in the second factor it follows from Lemma A.3 that

$$
\begin{aligned}
&\mathbb{E}\, \delta(DF_j(-DL^{-1}F_k))^2 \\
&\leq \int_{\mathbb{X}} \mathbb{E}\, (D_x F_j)^2 (-D_x L^{-1} F_k)^2\, \lambda(\mathrm{d}x) + \int_{\mathbb{X}^2} \mathbb{E}\, \big(D_y (D_x F_j (-D_x L^{-1} F_k)) \big)^2\, \lambda^2(\mathrm{d}(x,y)) \\
&\leq \frac{1}{2} \int_{\mathbb{X}} \mathbb{E}\, (D_x F_j)^4 + \mathbb{E}\, (-D_x L^{-1} F_k)^4\, \lambda(\mathrm{d}x) \\
&\quad + 3 \int_{\mathbb{X}^2} \mathbb{E}\, (D^2_{x,y} F_j)^2 (-D_x L^{-1} F_k)^2 + \mathbb{E}\, (D_x F_j)^2 (-D^2_{x,y} L^{-1} F_k)^2 \\
&\qquad\qquad + \mathbb{E}\, (D^2_{x,y} F_j)^2 (-D^2_{x,y} L^{-1} F_k)^2\, \lambda^2(\mathrm{d}(x,y)),
\end{aligned}
$$

where we used the arithmetic geometric mean inequality $a_1 a_2 \leq \frac{1}{2}(a_1^2 + a_2^2)$ for $a_1, a_2 \in (0, \infty)$ as well as Lemma A.1 and Jensen's inequality. It follows from Proposition 3.2 (a) and the Cauchy-Schwarz inequality that

$$
\begin{aligned}
&\mathbb{E}\, \delta(DF_j(-DL^{-1}F_k))^2 \\
&\leq \frac{1}{2} \int_{\mathbb{X}} \mathbb{E}\, (D_x F_j)^4 + \mathbb{E}\, (D_x F_k)^4\, \lambda(\mathrm{d}x) \\
&\quad + 3 \int_{\mathbb{X}^2} \big(\mathbb{E}\, (D^2_{x,y} F_j)^4 \big)^{1/2} \big(\mathbb{E}\, (D_x F_k)^4 \big)^{1/2} + \big(\mathbb{E}\, (D_x F_j)^4 \big)^{1/2} \big(\mathbb{E}\, (D^2_{x,y} F_k)^4 \big)^{1/2} \\
&\qquad\qquad + \big(\mathbb{E}\, (D^2_{x,y} F_j)^4 \big)^{1/2} \big(\mathbb{E}\, (D^2_{x,y} F_k)^4 \big)^{1/2}\, \lambda^2(\mathrm{d}(x,y)).
\end{aligned}
$$

Since $\gamma_4 < \infty$, the right-hand side is finite, which implies that assumptions (A.2) and (A.3) are satisfied and, thus, justifies the previous applications of Lemma A.2 and Lemma A.3. Combining the previous estimates yields

$$
|J_{2,1}| \leq \frac{1}{2} \|\Sigma^{-1}\|_{op} (\sqrt{M_2} |\log t| \sqrt{d_{\mathbb{H}_{2\ell}}(F, N_\Sigma)} + 22 m^{23/12}) \gamma_4. \tag{3.5}
$$

The bound for $|J_{2,2}|$ is a bit more involved and goes as follows. First, note that the triangle inequality and (2.7) imply that

$$
\begin{aligned}
|J_{2,2}| &\leq \sum_{i,j,k=1}^{m} \mathbb{E} \int_{\mathbb{X}} \int_0^1 \int_0^1 \left| \frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F + vuD_xF) - \frac{\partial^3 f_{t,h,\Sigma}}{\partial y_i \partial y_j \partial y_k}(F + vD_xF) \right| \\
&\qquad\qquad\qquad\qquad\qquad\qquad u|D_xF_i\, D_xF_j\, D_xL^{-1}F_k| \,\mathrm{d}v \,\mathrm{d}u \,\lambda(\mathrm{d}x) \\
&\leq \sum_{i,j,k=1}^{m} \mathbb{E} \int_{\mathbb{X}} \int_0^1 \int_0^1 \frac{1}{2} \int_t^1 \int_{\mathbb{R}^m} \frac{\sqrt{1-s}}{s^{3/2}} \\
&\qquad\qquad \left| h(\sqrt{s}z + \sqrt{1-s}(F + vuD_xF)) - h(\sqrt{s}z + \sqrt{1-s}(F + vD_xF)) \right| \\
&\qquad\qquad \left| \frac{\partial^3 \varphi_\Sigma}{\partial y_i \partial y_j \partial y_k}(z) \right| u|D_xF_i\, D_xF_j\, D_xL^{-1}F_k| \,\mathrm{d}z \,\mathrm{d}s \,\mathrm{d}v \,\mathrm{d}u \,\lambda(\mathrm{d}x).
\end{aligned}
$$

Using the abbreviation

$$
\begin{aligned}
U_{ijk} := \sup_{\substack{z\in\mathbb{R}^m,\\ s,u,v\in[0,1]}} \mathbb{E} \int_{\mathbb{X}} &\left| h(\sqrt{s}z + \sqrt{1-s}(F + vuD_xF)) - h(\sqrt{s}z + \sqrt{1-s}(F + vD_xF)) \right| \\
&\qquad |D_xF_i\, D_xF_j\, D_xL^{-1}F_k| \,\lambda(\mathrm{d}x).
\end{aligned}
$$

for $i, j, k \in \{1, \ldots, m\}$ and the Cauchy-Schwarz inequality, we obtain that

$$
\begin{aligned}
|J_{2,2}| &\leq \frac{1}{2\sqrt{t}} \sum_{i,j,k=1}^{m} \int_{\mathbb{R}^m} \left| \frac{\partial^3 \varphi_\Sigma}{\partial y_i \partial y_j \partial y_k}(z) \right| \mathrm{d}z \, U_{ijk} \\
&\leq \frac{1}{2\sqrt{t}} \int_{\mathbb{R}^m} \left(\sum_{i,j,k=1}^{m} \left(\frac{\partial^3 \varphi_\Sigma}{\partial y_i \partial y_j \partial y_k}(z) \right)^2 \right)^{1/2} \mathrm{d}z \left(\sum_{i,j,k=1}^{m} U_{ijk}^2 \right)^{1/2}.
\end{aligned}
$$

We have by (2.9) and substitution that

$$
\begin{aligned}
&\int_{\mathbb{R}^m} \left(\sum_{i,j,k=1}^{m} \left(\frac{\partial^3 \varphi_\Sigma}{\partial y_i \partial y_j \partial y_k}(z) \right)^2 \right)^{1/2} \mathrm{d}z \\
&\leq \frac{\|\Sigma^{-1}\|_{op}^{3/2}}{\sqrt{\det(\Sigma)}} \int_{\mathbb{R}^m} \left(\sum_{i,j,k=1}^{m} \left(\frac{\partial^3 \varphi_I}{\partial y_i \partial y_j \partial y_k}(\Sigma^{-1/2}z) \right)^2 \right)^{1/2} \mathrm{d}z = M_3 \|\Sigma^{-1}\|_{op}^{3/2}
\end{aligned}
$$

with

$$
M_3 := \int_{\mathbb{R}^m} \left(\sum_{i,j,k=1}^{m} \left(\frac{\partial^3 \varphi_I}{\partial y_i \partial y_j \partial y_k}(z) \right)^2 \right)^{1/2} \mathrm{d}z
$$

so that

$$
|J_{2,2}| \leq M_3 \|\Sigma^{-1}\|_{op}^{3/2} \frac{1}{2\sqrt{t}} \left(\sum_{i,j,k=1}^{m} U_{ijk}^2 \right)^{1/2}. \tag{3.6}
$$

The Cauchy-Schwarz inequality yields that

$$
\begin{aligned}
M_3 &= \int_{\mathbb{R}^m} \Bigg(\sum_{i,j,k=1}^{m} \bigg(\frac{\partial^3 \varphi_I}{\partial y_i \partial y_j \partial y_k}(z) \frac{1}{\varphi_I(z)} \bigg)^2 \Bigg)^{1/2} \varphi_I(z) \, \mathrm{d}z \\
&\leq \Bigg(\sum_{i,j,k=1}^{m} \int_{\mathbb{R}^m} \bigg(\frac{\partial^3 \varphi_I}{\partial y_i \partial y_j \partial y_k}(z) \frac{1}{\varphi_I(z)} \bigg)^2 \varphi_I(z) \, \mathrm{d}z \Bigg)^{1/2}.
\end{aligned}
$$

Together with the observation that, for a standard Gaussian random variable N with density ϕ,

$$
\begin{aligned}
\mathbb{E}\,[(\phi'(N)/\phi(N))^2] &= \mathbb{E}\,[N^2] = 1 \\
\mathbb{E}\,[(\phi''(N)/\phi(N))^2] &= \mathbb{E}\,[(N^2-1)^2] = \mathbb{E}\,[N^4 - 2N^2 + 1] = 2 \\
\mathbb{E}\,[(\phi'''(N)/\phi(N))^2] &= \mathbb{E}\,[(N^3-3N)^2] = \mathbb{E}\,[N^6 - 6N^4 + 9N^2] = 6
\end{aligned}
$$

this implies that

$$
M_3 \leq \sqrt{6} m^{3/2}. \tag{3.7}
$$

Next we bound U_{ijk} for fixed $i,j,k \in \{1,\dots,m\}$. Since $h \in \mathbb{H}_\ell$, it can be written as

$$
h(x) = \mathbf{1}\{x \in \bigcap_{n=1}^{\ell} H_n\}, \quad x \in \mathbb{R}^m,
$$

with some closed half-spaces $H_1,\dots,H_\ell$ in $\mathbb{R}^m$. With s and z fixed, notice that the convexity of H_n and H_n^c, $n \in \{1,\dots,\ell\}$, yields for all $u,v \in [0,1]$,

$$
\begin{aligned}
&\big| h(\sqrt{s}z + \sqrt{1-s}(F + uvD_xF)) - h(\sqrt{s}z + \sqrt{1-s}(F + vD_xF)) \big| \\
&\leq \sum_{n=1}^{\ell} |\mathbf{1}\{\sqrt{s}z + \sqrt{1-s}(F + uvD_xF) \in H_n\} - \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}(F + vD_xF) \in H_n\}| \\
&\leq \sum_{n=1}^{\ell} |\mathbf{1}\{\sqrt{s}z + \sqrt{1-s}(F + D_xF) \in H_n\} - \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}| \\
&= \sum_{n=1}^{\ell} |D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}|.
\end{aligned}
$$

Denoting by $u_1,\dots,u_\ell \in \mathbb{S}^{d-1}$ the outward pointing unit normal vectors of $H_1,\dots,H_\ell$, we see that, for $n \in \{1,\dots,\ell\}$,

$$
\begin{aligned}
|D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}| &= |D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}|\, \mathbf{1}\{\langle u_n, D_xF\rangle \leq 0\} \\
&\quad + |D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}|\, \mathbf{1}\{\langle u_n, D_xF\rangle > 0\} \\
&= D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}\, \mathbf{1}\{\langle u_n, D_xF\rangle \leq 0\} \\
&\quad - D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}\, \mathbf{1}\{\langle u_n, D_xF\rangle > 0\} \\
&= D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}(2\mathbf{1}\{\langle u_n, D_xF\rangle \leq 0\} - 1).
\end{aligned}
$$

Let $g(x) := |D_xF_i\, D_xF_j\, D_xL^{-1}F_k|$ for $x \in \mathbb{X}$. Combining the previous estimates with the integration by parts formula in Lemma A.4 (whose assumptions (A.2) and (A.3) are discussed before (3.8) below) and the Cauchy-Schwarz-inequality, we see that

$$
\begin{aligned}
&\mathbb{E} \int_{\mathbb{X}} \big|h(\sqrt{s}z + \sqrt{1-s}(F + uvD_xF)) - h(\sqrt{s}z + \sqrt{1-s}(F + vD_xF))\big|\, g(x)\, \lambda(\mathrm{d}x) \\
&\le \sum_{n=1}^{\ell} \mathbb{E} \int_{\mathbb{X}} D_x \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}\, (2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1)\, g(x)\, \lambda(\mathrm{d}x) \\
&= \sum_{n=1}^{\ell} \mathbb{E}\, \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}F \in H_n\}\, \delta((2\mathbf{1}\{\langle u_n, DF\rangle \le 0\} - 1)\, g) \\
&\le \sum_{n=1}^{\ell} \big(\mathbb{E}\, \delta((2\mathbf{1}\{\langle u_n, DF\rangle \le 0\} - 1)\, g)^2\big)^{1/2}.
\end{aligned}
$$

For $n \in \{1, \dots, \ell\}$ Lemma A.3 leads to

$$
\begin{aligned}
T_n &:= \mathbb{E}\, \delta((2\mathbf{1}\{\langle u_n, DF\rangle \le 0\} - 1)\, g)^2 \\
&\le \mathbb{E} \int_{\mathbb{X}} (2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1)^2\, g(x)^2\, \lambda(\mathrm{d}x) \\
&\quad + \mathbb{E} \int_{\mathbb{X}^2} \big(D_y\big((2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1)\, g(x)\big)\big)^2\, \lambda^2(\mathrm{d}(x,y)).
\end{aligned}
$$

By the product formula in Lemma A.1 and the fact that $|D_yG + G| \le 1$, $y \in \mathbb{X}$, for any Poisson functional G with values in $\{-1, 1\}$, we have

$$
\begin{aligned}
&\big|D_y\big((2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1)\, g(x)\big)\big| \\
&= \big|D_y(2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1)\, g(x) \\
&\quad + \big(D_y(2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1) + 2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1\big) D_yg(x)\big| \\
&\le 2|D_y\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\}|\, g(x) + \big|D_yg(x)\big|.
\end{aligned}
$$

Since

$$
|D_y\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\}| \le \mathbf{1}\{D^2_{x,y}F \ne \mathbf{0}\}, \quad x, y \in \mathbb{X},
$$

and

$$
|D_yg(x)| \le |D_y(D_xF_i\, D_xF_j\, D_xL^{-1}F_k)|, \quad x, y \in \mathbb{X},
$$

we obtain that

$$
\begin{aligned}
T_n &\le \mathbb{E} \int_{\mathbb{X}} (D_xF_i)^2 (D_xF_j)^2 (D_xL^{-1}F_k)^2\, \lambda(\mathrm{d}x) \\
&\quad + 8\mathbb{E} \int_{\mathbb{X}^2} \mathbf{1}\{D^2_{x,y}F \ne 0\} (D_xF_i\, D_xF_j\, D_xL^{-1}F_k)^2\, \lambda^2(\mathrm{d}(x,y)) \\
&\quad + 2\mathbb{E} \int_{\mathbb{X}^2} (D_y(D_xF_i\, D_xF_j\, D_xL^{-1}F_k))^2\, \lambda^2(\mathrm{d}(x,y)) \\
&=: S_1 + S_2 + S_3.
\end{aligned}
$$

The arithmetic mean geometric mean inequality and Proposition 3.2 (a) yield

$$\begin{aligned} S_1 &\le \frac{1}{3}\int_{\mathbb{X}} \mathbb{E}\,(D_xF_i)^6 + \mathbb{E}\,(D_xF_j)^6 + \mathbb{E}\,(D_xL^{-1}F_k)^6\,\lambda(\mathrm{d}x) \\ &\le \frac{1}{3}\int_{\mathbb{X}} \mathbb{E}\,(D_xF_i)^6 + \mathbb{E}\,(D_xF_j)^6 + \mathbb{E}\,(D_xF_k)^6\,\lambda(\mathrm{d}x). \end{aligned}$$

It follows from Hölder's inequality and Proposition 3.2 (a) that

$$\begin{aligned} S_2 &\le 8\int_{\mathbb{X}^2} \big(\mathbb{E}\,\mathbf{1}\{D^2_{x,y}F \neq \mathbf{0}\}|D_xF_i\,D_xF_j|^3\big)^{2/3}\big(\mathbb{E}\,(D_xL^{-1}F_k)^6\big)^{1/3}\,\lambda^2(\mathrm{d}(x,y)) \\ &\le 8\int_{\mathbb{X}^2} \big(\mathbb{E}\,\mathbf{1}\{D^2_{x,y}F \neq \mathbf{0}\}|D_xF_i\,D_xF_j|^3\big)^{2/3}\big(\mathbb{E}\,(D_xF_k)^6\big)^{1/3}\,\lambda^2(\mathrm{d}(x,y)). \end{aligned}$$

Using the iterated version of the product formula for D_x at Lemma A.1 (iterated since we have the product of three factors), we obtain

$$\begin{aligned} S_3 \le 14\int_{\mathbb{X}^2} &\mathbb{E}\,(D^2_{x,y}F_i\,D_xF_j\,D_xL^{-1}F_k)^2 + \mathbb{E}\,(D_xF_i\,D^2_{x,y}F_j\,D_xL^{-1}F_k)^2 \\ &+ \mathbb{E}\,(D_xF_i\,D_xF_j\,D^2_{x,y}L^{-1}F_k)^2 + \mathbb{E}\,(D^2_{x,y}F_i\,D^2_{x,y}F_j\,D_xL^{-1}F_k)^2 \\ &+ \mathbb{E}\,(D^2_{x,y}F_i\,D_xF_j\,D^2_{x,y}L^{-1}F_k)^2 + \mathbb{E}\,(D_xF_i\,D^2_{x,y}F_j\,D^2_{x,y}L^{-1}F_k)^2 \\ &+ \mathbb{E}\,(D^2_{x,y}F_i\,D^2_{x,y}F_j\,D^2_{x,y}L^{-1}F_k)^2\,\lambda^2(\mathrm{d}(x,y)). \end{aligned}$$

Separating the factors involving L^{-1} by Hölder's inequality and applying Proposition 3.2 (a) to them, we have

$$\begin{aligned} S_3 \le 14\int_{\mathbb{X}^2} &\big(\mathbb{E}\,|D^2_{x,y}F_i\,D_xF_j|^3\big)^{2/3}\big(\mathbb{E}\,(D_xF_k)^6\big)^{1/3} + \big(\mathbb{E}\,|D_xF_i\,D^2_{x,y}F_j|^3\big)^{2/3}\big(\mathbb{E}\,D_xF_k)^6\big)^{1/3} \\ &+ \big(\mathbb{E}\,|D_xF_i\,D_xF_j|^3\big)^{2/3}\big(\mathbb{E}\,(D^2_{x,y}F_k)^6\big)^{1/3} + \big(\mathbb{E}\,|D^2_{x,y}F_i\,D^2_{x,y}F_j|^3\big)^{2/3}\big(\mathbb{E}\,(D_xF_k)^6\big)^{1/3} \\ &+ \big(\mathbb{E}\,|D^2_{x,y}F_i\,D_xF_j|^3\big)^{2/3}\big(\mathbb{E}\,(D^2_{x,y}F_k)^6\big)^{1/3} + \big(\mathbb{E}\,|D_xF_i\,D^2_{x,y}F_j|^3\big)^{2/3}\big(\mathbb{E}\,D^2_{x,y}F_k)^6\big)^{1/3} \\ &+ \big(\mathbb{E}\,|D^2_{x,y}F_i\,D^2_{x,y}F_j|^3\big)^{2/3}\big(\mathbb{E}\,(D^2_{x,y}F_k)^6\big)^{1/3}\,\lambda^2(\mathrm{d}(x,y)). \end{aligned}$$

Note that $\gamma_5 < \infty$ (otherwise there is nothing to prove) implies that (A.2) and (A.3) are satisfied for $\mathbb{X} \ni x \mapsto (2\mathbf{1}\{\langle u_n, D_xF\rangle \le 0\} - 1)\,g(x)$ and justifies the applications of Lemma A.4 and Lemma A.3. From (3.6) and the estimates for S_1, S_2, and S_3, we obtain

$$|J_{2,2}| \le M_3\|\Sigma^{-1}\|^{3/2}_{op}\,\ell\frac{1}{2\sqrt{t}}\gamma_5. \tag{3.8}$$

Combining (3.4), (3.5), and (3.8) yields

$$\begin{aligned} |\mathbb{E}\,h_{t,\Sigma}(F) - \mathbb{E}\,h_{t,\Sigma}(N_\Sigma)| \le \|\Sigma^{-1}\|_{op}(\sqrt{M_2}|\log t|&\sqrt{d_{\mathbb{H}_{2\ell}}(F,N_\Sigma)} + 22m^{23/12}) \\ &\times \Big(\sum_{i,j=1}^{m}|\sigma_{ij} - \mathrm{Cov}(F_i,F_j)| + 2\gamma_1 + \gamma_2 + \frac{1}{2}\gamma_4\Big) \\ + M_3\|\Sigma^{-1}\|^{3/2}_{op}\,\ell\frac{1}{2\sqrt{t}}\gamma_5. \end{aligned}$$

From Lemma 2.3 as well as (2.11) and (3.7) we obtain

$$
\begin{aligned}
d_{\mathbb{H}_\ell}(F,N_\Sigma) \leq 2\|\Sigma^{-1}\|_{op}(\sqrt{M_2}|\log t|\sqrt{d_{\mathbb{H}_{2\ell}}(F,N_\Sigma)}+22m^{23/12}) \\
\times \Big(\sum_{i,j=1}^m |\sigma_{ij}-\operatorname{Cov}(F_i,F_j)|+2\gamma_1+\gamma_2+\frac{1}{2}\gamma_4\Big) \\
&+ M_3\|\Sigma^{-1}\|_{op}^{3/2}\ell\frac{1}{\sqrt{t}}\gamma_5+\frac{24\ell\sqrt{m}}{\sqrt{\pi}}\sqrt{t} \\
&\leq m\bigg(9(|\log t|\sqrt{d_{\mathbb{H}_{2\ell}}(F,N_\Sigma)}+22m^{11/12})\tilde{\gamma}_\ell+\sqrt{6}\sqrt{m}\|\Sigma^{-1}\|_{op}^{3/2}\frac{\ell\gamma_5}{\sqrt{t}}\bigg) \\
&+\frac{24\ell\sqrt{m}}{\sqrt{\pi}}\sqrt{t}
\end{aligned}
$$

with

$$
\tilde{\gamma}_\ell := \|\Sigma^{-1}\|_{op}\max\bigg\{\sum_{i,j=1}^m |\sigma_{ij}-\operatorname{Cov}(F_i,F_j)|,\gamma_1,\gamma_2,\gamma_4,\frac{\sqrt{\ell}\sqrt{\gamma_5}}{\|\Sigma^{-1}\|_{op}^{1/4}}\bigg\}.
$$

We can assume that $\tilde{\gamma}_\ell\in(0,1)$ since otherwise the desired inequality (1.9) is obviously true. Choosing $\sqrt{t}=\tilde{\gamma}_\ell$ and using $\ell\|\Sigma^{-1}\|_{op}^{3/2}\gamma_5/\tilde{\gamma}_\ell\leq\sqrt{\ell}\|\Sigma^{-1}\|_{op}^{3/4}\sqrt{\gamma_5}\leq\tilde{\gamma}_\ell$, we obtain

$$
d_{\mathbb{H}_\ell}(F,N_\Sigma)\leq m\bigg(9(2|\log\tilde{\gamma}_\ell|\sqrt{d_{\mathbb{H}_{2\ell}}(F,N_\Sigma)}+22m^{11/12})\tilde{\gamma}_\ell+\sqrt{6}\sqrt{m}\tilde{\gamma}_\ell\bigg)+\frac{24\ell\sqrt{m}}{\sqrt{\pi}}\tilde{\gamma}_\ell. \quad (3.9)
$$

The elementary bound $d_{\mathbb{H}_{2\ell}}(F,N_\Sigma)\leq 1$ implies

$$
\begin{aligned}
d_{\mathbb{H}_\ell}(F,N_\Sigma) &\leq m\bigg(9(2|\log\tilde{\gamma}_\ell|+22m^{11/12})\tilde{\gamma}_\ell+\sqrt{6}\sqrt{m}\tilde{\gamma}_\ell\bigg)+\frac{24\ell\sqrt{m}}{\sqrt{\pi}}\tilde{\gamma}_\ell \\
&\leq m^{23/12}\ell\tilde{\gamma}_\ell(214+18|\log\tilde{\gamma}_\ell|).
\end{aligned}
$$

Together with $\tilde{\gamma}_\ell\leq\tilde{\gamma}_{2\ell}\leq\sqrt{2}\tilde{\gamma}_\ell$, $\sup_{u\in[0,\infty)}u^2\exp(-u)\leq 1$ and $\sup_{u\in[0,\infty)}u^3\exp(-u)\leq 3/2$ this implies that

$$
\begin{aligned}
|\log\tilde{\gamma}_\ell|\sqrt{d_{\mathbb{H}_{2\ell}}(F,N_\Sigma)} &\leq |\log\tilde{\gamma}_\ell|\sqrt{m^{23/12}(2\ell)\tilde{\gamma}_{2\ell}(214+18|\log\tilde{\gamma}_{2\ell}|)} \\
&\leq |\log\tilde{\gamma}_\ell|\sqrt{2^{3/2}m^{23/12}\ell\tilde{\gamma}_\ell(214+18+18|\log\tilde{\gamma}_\ell|)} \\
&\leq 2^{3/4}\sqrt{m^{23/12}\ell}\sup_{u\in[0,\infty)}\sqrt{u^2\exp(-u)(232+18u)} \\
&\leq 28m^{23/24}\sqrt{\ell}.
\end{aligned}
$$

Putting this in (3.9) yields $d_{\mathbb{H}_\ell}(F,N_\Sigma)\leq 718m^{47/24}\ell\gamma_\ell$, completing the proof. □

Proof of Theorem 1.3. The only part of the proof of Theorem 1.2 that does not apply to $h\in\mathcal{I}_m$ are the upper bounds for U_{ijk}, $i,j,k\in\{1,\dots,m\}$, which rely on the assumption $h\in\mathbb{H}_\ell$. The inequalities (3.4) and (3.5) still hold if the $d_{\mathbb{H}_{2\ell}}$-distance is replaced by the d_{convex}-distance. In the following we derive an upper bound for U_{ijk} with fixed $i,j,k \in \{1,\dots,m\}$, in the case that $h \in \mathcal{I}_m$.

By assumption we have $|D_xF_i|, |D_xF_j| \le \varrho$ $\mathbb{P}$-a.s. for λ-a.e. $x \in \mathbb{X}$. Next we prove $|D_xL^{-1}F_k| \le \varrho$ $\mathbb{P}$-a.s. for λ-a.e. $x \in \mathbb{X}$. By Proposition 3.2(a) and (1.10), for λ-a.e. $x \in \mathbb{X}$ the inequality

$$\mathbb{E}\,|D_xL^{-1}F_k|^p \le \mathbb{E}\,|D_xF_k|^p \le \varrho^p \tag{3.10}$$

is true for all $p \in \mathbb{N}$. If for such an $x \in \mathbb{X}$, $|D_xL^{-1}F_k| \le \varrho$ $\mathbb{P}$-a.s. does not hold, there is an $\varepsilon > 0$ such that $q := \mathbb{P}(|D_xL^{-1}F_k| \ge \varrho + \varepsilon) > 0$. This implies that, for any $p \in \mathbb{N}$,

$$\mathbb{E}\,|D_xL^{-1}F_k|^p \ge q(\varrho+\varepsilon)^p,$$

which contradicts (3.10) for p sufficiently large.

Now assume that $h = \mathbf{1}\{\cdot \in K\}$ with $K \subseteq \mathbb{R}^m$ closed and convex and let $s,u,v \in [0,1]$ and $z \in \mathbb{R}^m$ be fixed. Then, we have that

$$\begin{aligned}
&\big|h(\sqrt{s}z + \sqrt{1-s}(F + vuD_xF)) - h(\sqrt{s}z + \sqrt{1-s}(F + vD_xF))\big| \\
&= \big|\mathbf{1}\{\sqrt{s}z + \sqrt{1-s}(F + vuD_xF) \in K\} - \mathbf{1}\{\sqrt{s}z + \sqrt{1-s}(F + vD_xF) \in K\}\big| \\
&= \big|\mathbf{1}\{F + vuD_xF \in (K - \sqrt{s}z)/\sqrt{1-s}\} - \mathbf{1}\{F + vD_xF \in (K - \sqrt{s}z)/\sqrt{1-s}\}\big|.
\end{aligned}$$

Let $K_{s,z} := (K - \sqrt{s}z)/\sqrt{1-s}$. Together with (1.10) we see that, for λ-a.e. $x \in \mathbb{X}$, $\mathbb{P}$-a.s.,

$$\begin{aligned}
&\big|h(\sqrt{s}z + \sqrt{1-s}(F + vuD_xF)) - h(\sqrt{s}z + \sqrt{1-s}(F + vD_xF))\big| \\
&\le \mathbf{1}\{d(F, K_{s,z}) \le \sqrt{m}\varrho\} - \mathbf{1}\{F \in K_{s,z}, d(F, \partial K_{s,z}) \ge \sqrt{m}\varrho\}.
\end{aligned}$$

Note that $\mathbb{R}^m \ni v \mapsto \mathbf{1}\{d(v, K_{s,z}) \le \sqrt{m}\varrho\}$ and $\mathbb{R}^m \ni v \mapsto \mathbf{1}\{v \in K_{s,z}, d(v, \partial K_{s,z}) \ge \sqrt{m}\varrho\}$ are both indicator functions of closed convex sets. This implies that, for λ-a.e. $x \in \mathbb{X}$,

$$\begin{aligned}
&\mathbb{E}\,\big|h(\sqrt{s}z + \sqrt{1-s}(F + vuD_xF)) - h(\sqrt{s}z + \sqrt{1-s}(F + vD_xF))\big| \\
&\le \mathbb{E}\left[\mathbf{1}\{d(N_\Sigma, K_{s,z}) \le \sqrt{m}\varrho\} - \mathbf{1}\{N_\Sigma \in K_{s,z}, d(N_\Sigma, \partial K_{s,z}) \ge \sqrt{m}\varrho\}\right] + 2d_{convex}(F, N_\Sigma).
\end{aligned}$$

It follows from (2.12) that, for $u \ge 0$,

$$\begin{aligned}
&\sup_{B \subseteq \mathbb{R}^m \text{ convex}} \mathbb{E}\left[\mathbf{1}\{d(N_\Sigma, B) \le u\} - \mathbf{1}\{N_\Sigma \in B, d(N_\Sigma, \partial B) \ge u\}\right] \\
&= \sup_{B \subseteq \mathbb{R}^m \text{ convex}} \mathbb{P}(d(N_\Sigma, \partial B) \le u) = \sup_{B \subseteq \mathbb{R}^m \text{ convex}} \mathbb{P}(d(N_I, \partial B) \le u) \\
&\le 2\sqrt{\frac{2}{\pi}} m^{3/2} u.
\end{aligned}$$

This implies that, for λ-a.e. $x \in \mathbb{X}$,

$$\mathbb{E}\left|h(\sqrt{s}z+\sqrt{1-s}(F+vuD_xF))-h(\sqrt{s}z+\sqrt{1-s}(F+vD_xF))\right|$$
$$\leq 2\sqrt{\frac{2}{\pi}}m^2\varrho+2d_{convex}(F,N_\Sigma).$$

On the other hand, we have that

$$\mathbb{E}\left|h(\sqrt{s}z+\sqrt{1-s}(F+vuD_xF))-h(\sqrt{s}z+\sqrt{1-s}(F+vD_xF))\right|\leq \mathbb{P}(D_xF\neq \mathbf{0})$$

for $x\in\mathbb{X}$. Altogether, we obtain that

$$U_{ijk}\leq\bigg(\Big(2d_{convex}(F,N_\Sigma)+2\sqrt{\frac{2}{\pi}}m^2\varrho\Big)\lambda(A)+\int_{\mathbb{X}\setminus A}\mathbb{P}(D_xF\neq 0)\,\lambda(\mathrm{d}x)\bigg)\varrho^3.$$

Thus, it follows from (3.6) that

$$|J_{2,2}|\leq M_3\|\Sigma^{-1}\|_{op}^{3/2}\frac{m^{3/2}}{2\sqrt{t}}\bigg(\Big(2d_{convex}(F,N_\Sigma)+2\sqrt{\frac{2}{\pi}}m^2\varrho\Big)\lambda(A)+\int_{\mathbb{X}\setminus A}\mathbb{P}(D_xF\neq 0)\,\lambda(\mathrm{d}x)\bigg)\varrho^3. \tag{3.11}$$

In the sequel we denote the integral in the last inequality by γ_6.

In light of Lemma 2.2, we may now substitute the bounds (3.4), (3.5) (with $d_{\mathbb{H}_{2\ell}}$ replaced by d_{convex} there), and (3.11) in Lemma 2.1 to obtain

$$\begin{aligned}d_{convex}(F,N_\Sigma)\leq \frac{4}{3}\bigg(&\|\Sigma^{-1}\|_{op}(\sqrt{M_2}|\log t|\sqrt{d_{convex}(F,N_\Sigma)}+22m^{23/12})\\&\times\Big(\sum_{i,j=1}^m|\sigma_{ij}-\mathrm{Cov}(F_i,F_j)|+2\gamma_1+\gamma_2+\frac{1}{2}\gamma_4\Big)\\&+\frac{M_3m^{3/2}\|\Sigma^{-1}\|_{op}^{3/2}\varrho^3}{\sqrt{t}}\bigg(\Big(d_{convex}(F,N_\Sigma)+\sqrt{\frac{2}{\pi}}m^2\varrho\Big)\lambda(A)+\frac{\gamma_6}{2}\bigg)\bigg)\\&+\frac{20}{\sqrt{\pi}}m^2\frac{\sqrt{t}}{1-t}.\end{aligned}$$

Define $\tilde{\gamma}:=m\|\Sigma^{-1}\|_{op}\gamma$ with γ as in Theorem 1.3 and use (2.11) and (3.7) so that

$$\begin{aligned}d_{convex}(F,N_\Sigma)\leq\ & 6(|\log t|\sqrt{d_{convex}(F,N_\Sigma)}+22m^{11/12})\tilde{\gamma}\\&+\frac{4\sqrt{6}m^3\|\Sigma^{-1}\|_{op}^{3/2}\varrho^3}{3\sqrt{t}}\bigg(\Big(d_{convex}(F,N_\Sigma)+\sqrt{\frac{2}{\pi}}m^2\varrho\Big)\lambda(A)+\frac{\gamma_6}{2}\bigg)\\&+\frac{20}{\sqrt{\pi}}m^2\frac{\sqrt{t}}{1-t}.\end{aligned}$$

We can assume that $\tilde{\gamma}\in(0,1/2)$ as otherwise the desired conclusion (1.11) becomes trivial. Putting $\sqrt{t}=\tilde{\gamma}$ and noting

$$\frac{4\sqrt{6}m^3\|\Sigma^{-1}\|_{op}^{3/2}\varrho^3\lambda(A)}{3\tilde{\gamma}}\leq\frac{1}{2}\quad\text{and}\quad\frac{4\sqrt{6}\sqrt{\frac{2}{\pi}}m^5\|\Sigma^{-1}\|_{op}^{3/2}\varrho^4\lambda(A)}{3\tilde{\gamma}}\leq\frac{8}{\sqrt{3\pi}}\tilde{\gamma}$$

as well as $\tilde{\gamma}/(1-\tilde{\gamma}^2) \leq 4\tilde{\gamma}/3$ yield

$$d_{convex}(F, N_\Sigma) \leq 6(2|\log \tilde{\gamma}|\sqrt{d_{convex}(F, N_\Sigma)} + 22m^{11/12})\tilde{\gamma} + \frac{1}{2}d_{convex}(F, N_\Sigma) + \frac{8}{\sqrt{3\pi}}\tilde{\gamma} + \frac{\tilde{\gamma}}{4} + \frac{80}{3\sqrt{\pi}}m^2\tilde{\gamma}.$$

Consequently, we have

$$d_{convex}(F, N_\Sigma) \leq 4\max\left\{\left(132m^{11/12}+\frac{8}{\sqrt{3\pi}}+\frac{1}{4}+\frac{80}{3\sqrt{\pi}}m^2\right)\tilde{\gamma}, 12|\log \tilde{\gamma}|\sqrt{d_{convex}(F, N_\Sigma)}\tilde{\gamma}\right\}$$

and, thus,

$$d_{convex}(F, N_\Sigma) \leq \max\left\{600m^2\tilde{\gamma}, 2304(\log \tilde{\gamma})^2\tilde{\gamma}^2\right\}.$$

The observation that $\sup_{u\in(0,1/2]}(\log u)^2 u \leq 1$ completes the proof of Theorem 1.3. □

4 Applications

4.1 Multivariate normal approximation of first order Wiener-Itô integrals

In this subsection we apply our main results to first order Wiener-Itô integrals with respect to the Poisson process η (as considered before). For $f \in L^1(\lambda) \cap L^2(\lambda)$ one can define the Wiener-Itô integral $I_1(f)$ of f as

$$I_1(f) := \int_{\mathbb{X}} f(x)\,\eta(\mathrm{d}x) - \int_{\mathbb{X}} f(x)\,\lambda(\mathrm{d}x).$$

If η is a proper Poisson process, i.e., it has almost surely a representation $\eta = \sum_{i\in I}\delta_{X_i}$ with a countable collection $(X_i)_{i\in I}$ of random elements of $\mathbb{X}$, this can be rewritten as

$$I_1(f) = \sum_{i\in I} f(X_i) - \int_{\mathbb{X}} f(x)\,\lambda(\mathrm{d}x).$$

Using approximation arguments in $L^2(\mathbb{P})$, one can extend the above definition to integrands $f \in L^2(\lambda)$. Note that, for all $f, g \in L^2(\lambda)$,

$$\mathbb{E}\, I_1(f) = 0 \quad \text{and} \quad \mathbb{E}\, I_1(f)I_1(g) = \int_{\mathbb{X}} f(x)g(x)\,\lambda(\mathrm{d}x). \tag{4.1}$$

For an exact definition and more details on first order Wiener-Itô integrals with respect to Poisson processes we refer to [16, Subsection 12.1].

Corollary 4.1. *Let $F = (I_1(f_1), \dots, I_1(f_m))$ with $f_1, \dots, f_m \in L^2(\lambda)$ and $m \in \mathbb{N}$ and let $\Sigma = (\sigma_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$ be positive semi-definite.*

(a) *It is the case that*

$$d_3(F, N_\Sigma) \le \frac{m}{2} \sum_{i,j=1}^m \Big|\sigma_{ij} - \int_{\mathbb{X}} f_i(x) f_j(x)\, \lambda(\mathrm{d}x)\Big| + \frac{m^2}{4} \sum_{i=1}^m \int_{\mathbb{X}} |f_i(x)|^3\, \lambda(\mathrm{d}x).$$

(b) *If Σ is positive definite,*

$$\begin{aligned} d_2(F, N_\Sigma) \le \|\Sigma^{-1}\|_{op} \|\Sigma\|_{op}^{1/2} \sum_{i,j=1}^m \Big|\sigma_{ij} - \int_{\mathbb{X}} f_i(x) f_j(x)\, \lambda(\mathrm{d}x)\Big| \\ + \frac{\sqrt{2\pi} m^2}{8} \|\Sigma^{-1}\|_{op}^{3/2} \|\Sigma\|_{op} \sum_{i=1}^m \int_{\mathbb{X}} |f_i(x)|^3\, \lambda(\mathrm{d}x). \end{aligned}$$

(c) *If Σ is positive definite, for any $\ell \in \mathbb{N}$,*

$$\begin{aligned} d_{\mathbb{H}_\ell}(F, N_\Sigma) \le 718 m^{59/24} \ell^{3/2} \max\{\|\Sigma^{-1}\|_{op}, \|\Sigma^{-1}\|_{op}^{3/4}\} \\ \max\Bigg\{ \sum_{i,j=1}^m \Big|\sigma_{ij} - \int_{\mathbb{X}} f_i(x) f_j(x)\, \lambda(\mathrm{d}x)\Big|, \\ \Bigg(\sum_{i=1}^m \int_{\mathbb{X}} f_i(x)^4\, \lambda(\mathrm{d}x)\Bigg)^{1/2}, \Bigg(\sum_{i=1}^m \int_{\mathbb{X}} f_i(x)^6\, \lambda(\mathrm{d}x)\Bigg)^{1/4} \Bigg\}. \end{aligned}$$

(d) *If Σ is positive definite, if there exists a $\varrho \in (0, \infty)$ with $|f_i(x)| \le \varrho$ for λ-a.e. $x \in \mathbb{X}$ and $i \in \{1, \dots, m\}$, and if $\lambda(\mathbb{X}) < \infty$, then*

$$\begin{aligned} d_{convex}(F, N_\Sigma) \le 15050 m^5 \max\{\|\Sigma^{-1}\|_{op}^{3/4}, \|\Sigma^{-1}\|_{op}^{3/2}\} \\ \max\Bigg\{ \sum_{i,j=1}^m \Big|\sigma_{ij} - \int_{\mathbb{X}} f_i(x) f_j(x)\, \lambda(\mathrm{d}x)\Big|, \varrho^3 \lambda(\mathbb{X}), \varrho^2 \sqrt{\lambda(\mathbb{X})} \Bigg\}. \end{aligned}$$

Proof. It follows from (4.1) that, for $i, j \in \{1, \dots, m\}$,

$$\mathrm{Cov}(I_1(f_i), I_1(f_j)) = \int_{\mathbb{X}} f_i(x) f_j(x)\, \lambda(\mathrm{d}x).$$

Moreover, it is well-known (see, for example, Eqn. (2.6) in [15]) that, for $f \in L^2(\lambda)$ and $x, x_1, x_2 \in \mathbb{X}$,

$$D_x I_1(f) = f(x) \quad \text{and} \quad D^2_{x_1, x_2} I_1(f) = 0.$$

This implies that $\gamma_1 = \gamma_2 = 0$, $\gamma_3 = \sum_{i=1}^m \int_{\mathbb{X}} |f_i(x)|^3\, \lambda(\mathrm{d}x)$,

$$\gamma_4 = \sqrt{m} \Bigg(\sum_{i=1}^m \int_{\mathbb{X}} f_i(x)^4\, \lambda(\mathrm{d}x)\Bigg)^{1/2} \quad \text{and} \quad \gamma_5 = m \Bigg(\sum_{i=1}^m \int_{\mathbb{X}} f_i(x)^6\, \lambda(\mathrm{d}x)\Bigg)^{1/2}.$$

Now (a) and (b) are immediate consequences of Theorem 1.1, while Theorem 1.2 yields (c). Part (d) follows from Theorem 1.3 with $A = \mathbb{X}$ and $\gamma_4 \le m \varrho^2 \sqrt{\lambda(\mathbb{X})}$. $\square$

The idea of the following proof of Corollary 1.4 is to show that it is only a special case of Corollary 4.1.

Proof of Corollary 1.4 . Let $\mathbb{X} = \mathbb{R}^d$ (equipped with its Borel σ-field) and $\lambda(\cdot) = s\mathbb{P}(X_1 \in \cdot)$, i.e., λ is s times the probability measure of X_1. For $i \in \{1, \dots, m\}$ let us denote by π_i the projection $\mathbb{R}^m \ni (y_1, \dots, y_m) \mapsto y_i$. Then we have that

$$Z_s = (I_1(\pi_1/\sqrt{s}), \dots, I_1(\pi_m/\sqrt{s})).$$

Together with the observation that, for $i \in \{1, \dots, m\}$, $p \in (0, \infty)$, and $s > 0$,

$$\int_{\mathbb{X}} |\pi_i(x)/\sqrt{s}|^p \, \lambda(\mathrm{d}x) = \mathbb{E}\, |X_1^{(i)}|^p s^{1-p/2},$$

we see that conclusions (a) and (b) of Corollary 1.4 follow from conclusions (a) and (b) of Corollary 4.1, with $p = 3$, conclusion (c) follows by considering $p \in \{4, 6\}$, whereas conclusion (d) follows from its counterpart in Corollary 4.1. $\square$

4.2 Multivariate central limit theorems for intrinsic volumes of Boolean models

In the following, we derive quantitative multivariate central limit theorems for Boolean models, which extend previous findings in [11] and [16, Chapter 22]. Their proofs rely on the general bounds from Subsection 1.2 as well as arguments from [11] and [16, Chapter 22].

We denote by $\mathcal{K}^d$ the set of compact convex sets in $\mathbb{R}^d$. For a probability measure $\mathbb{Q}$ on $\mathcal{K}^d$ such that $\mathbb{Q}(\{\emptyset\}) = 0$ and $\gamma > 0$ let η be a Poisson process on $\mathbb{R}^d \times \mathcal{K}^d$ with intensity measure $\gamma\lambda_d \otimes \mathbb{Q}$, where λ_d is the Lebesgue measure on $\mathbb{R}^d$. Note that η is a stationary Poisson process in $\mathbb{R}^d$ with independent marks in $\mathcal{K}^d$ distributed according to $\mathbb{Q}$. A random compact convex set Z_0 distributed according to $\mathbb{Q}$ is called the typical grain. From η we construct the random closed set

$$Z := \bigcup_{(x,K)\in\eta} (x + K),$$

which is called the Boolean model. For more details on Boolean models and further references we refer to [29].

In the sequel we study the intersection of the Boolean model with a compact convex observation window $W \in \mathcal{K}^d$. Note that $Z \cap W$ almost surely belongs to the convex ring $\mathcal{R}^d$, i.e., the set of all finite unions of elements from $\mathcal{K}^d$, if $\mathbb{E}\, V_i(Z_0) < \infty$ for $i \in \{1, \dots, d\}$.

Questions of interest include finding the fraction of W covered by Z and the surface area of $Z \cap W$. We address both problems simultaneously by considering the behavior of

$$\mathcal{V}(Z \cap W) := (V_0(Z \cap W), V_1(Z \cap W), \dots, V_d(Z \cap W)),$$

where $V_0, V_1, \dots, V_d : \mathcal{R}^d \to \mathbb{R}$ are the intrinsic volumes (see, for example, [29, Section 14.2] for a definition via the Steiner formula and additive extensions). In particular, for $K \in \mathcal{R}^d$, $V_d(K)$ is the volume of K, $V_{d-1}(K)$ is half the surface area of K (if K is the closure of its interior), and $V_0(K)$ is the Euler characteristic of K.

Let us denote by $r(K)$ the inradius of $K \in \mathcal{K}^d$. In [11, Theorem 3.1] it is shown that there exists a matrix $\Sigma = (\sigma_{i,j})_{i,j \in \{0,\dots,d\}} \in \mathbb{R}^{(d+1)\times(d+1)}$ such that

$$\Sigma(W) := \frac{1}{V_d(W)} (\mathrm{Cov}(V_i(Z \cap W), V_j(Z \cap W)))_{i,j \in \{0,\dots,d\}} \to \Sigma \quad \text{as} \quad r(W) \to \infty$$

if $\mathbb{E}\, V_i(Z_0)^2 < \infty$ for $i \in \{1, \dots, d\}$. If, additionally, $\mathbb{P}(V_d(Z_0) > 0) > 0$, the asymptotic covariance matrix Σ is even positive definite (see [11, Theorem 4.1]). We describe the asymptotic behavior of $\mathcal{V}(Z \cap W)$ as $r(W) \to \infty$ with respect to d_3, d_2, and $d_{\mathbb{H}_\ell}$.

Theorem 4.2. *(a) If $\mathbb{E}\, V_i(Z_0)^3 < \infty$ for $i \in \{1, \dots, d\}$, there exists a constant $C_1 \in (0, \infty)$ depending on d, γ, and $\mathbb{Q}$ such that*

$$d_3\left(\frac{\mathcal{V}(Z \cap W) - \mathbb{E}\,\mathcal{V}(Z \cap W)}{\sqrt{V_d(W)}}, N_\Sigma\right) \le C_1 \frac{1}{r(W)^{\min\{1,d/2\}}}$$

for all $W \in \mathcal{K}^d$ with $r(W) \ge 1$.

(b) If $\mathbb{E}\, V_i(Z_0)^3 < \infty$ for $i \in \{1, \dots, d\}$ and $\mathbb{P}(V_d(Z_0) > 0) > 0$, there exists a constant $C_2 \in (0, \infty)$ depending on d, γ, and $\mathbb{Q}$ such that

$$d_2\left(\frac{\mathcal{V}(Z \cap W) - \mathbb{E}\,\mathcal{V}(Z \cap W)}{\sqrt{V_d(W)}}, N_\Sigma\right) \le C_2 \frac{1}{r(W)^{\min\{1,d/2\}}}$$

for all $W \in \mathcal{K}^d$ with $r(W) \ge 1$.

(c) Let $\ell \in \mathbb{N}$. If $\mathbb{E}\, V_i(Z_0)^7 < \infty$ for $i \in \{1, \dots, d\}$ and $\mathbb{P}(V_d(Z_0) > 0) > 0$, there exists a constant $C_{3,\ell} \in (0, \infty)$ depending on ℓ, d, γ, and $\mathbb{Q}$ such that

$$d_{\mathbb{H}_\ell}\left(\frac{\mathcal{V}(Z \cap W) - \mathbb{E}\,\mathcal{V}(Z \cap W)}{\sqrt{V_d(W)}}, N_\Sigma\right) \le C_{3,\ell} \frac{1}{r(W)^{\min\{1,d/2\}}}$$

for all $W \in \mathcal{K}^d$ with $r(W) \ge 1$.

(d) If N_Σ is replaced by $N_{\Sigma(W)}$, the assertions (a)-(c) hold with the rate $1/\sqrt{V_d(W)}$.

Note that Theorem 4.2(a) is a special case of [11, Theorem 9.1], while (b) and (c) extend it to different distances, in particular, the non-smooth $d_{\mathbb{H}_\ell}$-distance. The findings of [11] as well as the univariate results in [16] consider so-called geometric functionals, which include the intrinsic volumes. Theorem 4.2 could be also generalized to these functionals, but for the sake of simplicity we consider only the intrinsic volumes. Since our proof of Theorem 4.2 is based on second order Poincaré inequalities, it does not require dealing with the whole chaos expansion as in [11]. For previous results on volume and surface area of Boolean models we refer the reader to [11]. Theorem 4.2 indicates that the slow convergence of $\Sigma(W)$ to W weakens the rate of convergence for $d \geq 3$ (see also [11, Remark 9.5]). The rate of convergence $1/\sqrt{V_d(W)}$ for the distance to $N_{\Sigma(W)}$ is comparable to $1/\sqrt{n}$ in the classical central limit theorem for sums of n i.i.d. random vectors and, thus, presumably optimal.

We prepare the proof of Theorem 4.2 by two lemmas. In the sequel, we use the Wills functional $\overline{V}(K) := \sum_{i=0}^{d} \kappa_{d-i} V_i(K)$ for $K \in \mathcal{K}^d$, where κ_{d-i} is the volume of the $(d-i)$-dimensional unit ball. We write the difference operator D with respect to pairs of points and compact convex sets.

Lemma 4.3. *There exists a constant $C \in (0,\infty)$ only depending on d, γ, and $\mathbb{Q}$ such that, for $x, x_1, x_2 \in \mathbb{R}^d$, $K, K_1, K_2 \in \mathcal{K}^d$, $i \in \{0,\dots,d\}$, and $m \in \{1,\dots,6\}$,*

$$\mathbb{E}\,|D_{(x,K)} V_i(Z \cap W)|^m \leq C^m \overline{V}(K \cap W)^m$$

and

$$\mathbb{E}\,|D_{(x_1,K_1),(x_2,K_2)} V_i(Z \cap W)|^m \leq C^m \overline{V}(K_1 \cap K_2 \cap W)^m.$$

Proof. For $m \in \{2,3\}$ this is shown in [16] in Proposition 22.4 in connection with (22.30) and (22.31) (see also [11, Lemma 3.3]), but the proof can be extended to the remaining m. □

Moreover, we will use the following translative integral formula from [16, Proposition 22.5] and [11, Lemma 3.4].

Lemma 4.4. *For all $K, L \in \mathcal{K}^d$,*

$$\int_{\mathbb{R}^d} \overline{V}((x+K) \cap L)\,\mathrm{d}x \leq \overline{V}(K)\overline{V}(L).$$

Proof of Theorem 4.2. We deduce Theorem 4.2 from Theorem 1.1 and Theorem 1.2 by bounding $\gamma_1,\dots,\gamma_5$ from Subsection 1.2 as follows. We denote by $\tilde{\gamma}_1,\dots,\tilde{\gamma}_5$ the corresponding terms without the normalization $1/\sqrt{V_d(W)}$ of the functionals. Without loss of generality we can assume that $\gamma = 1$. In the sequel let $(Z_n)_{n\in\mathbb{N}}$ be independent copies of the typical grain Z_0. It follows from the Cauchy-Schwarz inequality, Lemma 4.3, the monotonicity and the translation invariance of the Wills functional (i.e., $\overline{V}(K) \leq \overline{V}(L)$ for $K, L \in \mathcal{K}^d$ with $K \subseteq L$ and $\overline{V}(K+x) = \overline{V}(K)$ for $K \in \mathcal{K}^d$ and $x \in \mathbb{R}^d$) and Lemma 4.4 that

$$\begin{aligned}
\tilde{\gamma}_1^2 &= (d+1)^2 C^4 \mathbb{E} \int_{(\mathbb{R}^d)^3} \overline{V}((x_1+Z_1)\cap(x_3+Z_3)\cap W)\overline{V}((x_2+Z_1)\cap(x_3+Z_3)\cap W) \\
&\qquad\qquad\qquad \overline{V}((x_1+Z_1)\cap W)\overline{V}((x_2+Z_2)\cap W)\,\mathrm{d}(x_1,x_2,x_3) \\
&\leq (d+1)^2 C^4 \mathbb{E} \int_{(\mathbb{R}^d)^3} \overline{V}((x_1+Z_1)\cap(x_3+Z_3)\cap W)\overline{V}((x_2+Z_1)\cap(x_3+Z_3)\cap W) \\
&\qquad\qquad\qquad \overline{V}(Z_1)\overline{V}(Z_2)\,\mathrm{d}(x_1,x_2,x_3) \\
&\leq (d+1)^2 C^4 \mathbb{E} \int_{\mathbb{R}^d} \overline{V}(Z_1)^2\overline{V}(Z_2)^2\overline{V}((x+Z_3)\cap W)^2\,\mathrm{d}x \\
&\leq (d+1)^2 C^4 \mathbb{E} \int_{\mathbb{R}^d} \overline{V}(Z_1)^2\overline{V}(Z_2)^2\overline{V}(Z_3)\overline{V}((x+Z_3)\cap W)\,\mathrm{d}x \\
&\leq (d+1)^2 C^4 \mathbb{E}\,\overline{V}(Z_1)^2\mathbb{E}\,\overline{V}(Z_2)^2\mathbb{E}\,\overline{V}(Z_3)^2\overline{V}(W) \\
&\leq (d+1)^2 C^4 (\mathbb{E}\,\overline{V}(Z_0)^2)^3\overline{V}(W)
\end{aligned}$$

and

$$\begin{aligned}
\tilde{\gamma}_2^2 &\leq (d+1)^2 C^4 \mathbb{E} \int_{(\mathbb{R}^d)^3} \overline{V}((x_1+Z_1)\cap(x_3+Z_3)\cap W)^2\overline{V}((x_2+Z_2)\cap(x_3+Z_3)\cap W)^2 \\
&\qquad\qquad\qquad \mathrm{d}(x_1,x_2,x_3) \\
&\leq (d+1)^2 C^4 \mathbb{E} \int_{(\mathbb{R}^d)^3} \overline{V}(Z_1)^2\overline{V}(Z_2)^2\overline{V}((x+Z_3)\cap W)^2\,\mathrm{d}x \\
&\leq (d+1)^2 C^4 \mathbb{E}\,\overline{V}(Z_1)^2\overline{V}(Z_2)^2\overline{V}(Z_3)^2\overline{V}(W) \\
&= (d+1)^2 C^4 (\mathbb{E}\,\overline{V}(Z_0)^2)^2\overline{V}(W).
\end{aligned}$$

Hence, we see that γ_1 and γ_2 are at most of the order $\sqrt{\overline{V}(W)}/V_d(W)$. From the same arguments as above we obtain that, for $k \in \mathbb{N}$,

$$\mathbb{E} \int_{\mathbb{R}^d} \overline{V}((x+Z_0)\cap W)^k\,\mathrm{d}x \leq \mathbb{E}\,\overline{V}(Z_0)^{k-1} \int_{\mathbb{R}^d} \overline{V}((x+Z_0)\cap W)\,\mathrm{d}x \leq \mathbb{E}\,\overline{V}(Z_0)^k\overline{V}(W), \quad (4.2)$$

whence γ_3 is at most of order $\overline{V}(W)/V_d(W)^{3/2}$. We can also show that

$$\begin{aligned}
&\mathbb{E} \int_{(\mathbb{R}^d)^2} \overline{V}((x_1+Z_1)\cap(x_2+Z_2)\cap W)^2 \\
&\qquad \big(\overline{V}((x_1+Z_1)\cap(x_2+Z_2)\cap W)^2 + \overline{V}((x_1+Z_1)\cap W)^2\big)\mathrm{d}(x_1,x_2) \\
&\leq 2\mathbb{E} \int_{(\mathbb{R}^d)^2} \overline{V}((x_1+Z_1)\cap(x_2+Z_2)\cap W)\overline{V}(Z_2)\overline{V}(Z_1)^2\mathrm{d}(x_1,x_2) \\
&\leq 2\mathbb{E}\,\overline{V}(Z_1)^3\overline{V}(Z_2)^2\overline{V}(W)
\end{aligned}$$

so that together with (4.2) γ_4 is at most of order $\sqrt{\overline{V}(W)}/V_d(W)$. Moreover, we have that

$$\begin{aligned}
&\mathbb{E}\int_{(\mathbb{R}^d)^2} \overline{V}((x_1+Z_1)\cap(x_2+Z_2)\cap W)^2 \\
&\qquad\qquad \big(\overline{V}((x_1+Z_1)\cap(x_2+Z_2)\cap W)^4+\overline{V}((x_1+Z_1\cap W))^4\big)\mathrm{d}(x_1,x_2) \\
&\le 2\mathbb{E}\int_{(\mathbb{R}^d)^2} \overline{V}((x_1+Z_1)\cap(x_2+Z_2)\cap W)\overline{V}(Z_2)\overline{V}(Z_1)^4\mathrm{d}(x_1,x_2) \\
&\le 2\mathbb{E}\,\overline{V}(Z_1)^5\overline{V}(Z_2)^2\overline{V}(W).
\end{aligned} \tag{4.3}$$

From [11, Lemma 3.2] or [16, Lemma 22.6] it follows that, for $i\in\{0,\dots,d\}$, $x_1,x_2\in\mathbb{R}^d$, and $K_1,K_2\in\mathcal{K}^d$,

$$D^2_{(x_1,K_1),(x_2,K_2)}V_i(Z\cap W)=V_i(Z\cap(x_1+K_1)\cap(x_2+K_2)\cap W)-V_i((x_1+K_1)\cap(x_2+K_2)\cap W).$$

Consequently, we have that

$$\begin{aligned}
\mathbf{1}\{D^2_{(x_1,K_1),(x_2,K_2)}\mathcal{V}(Z\cap W)\neq 0\} &\le \mathbf{1}\{(x_1+K_1)\cap(x_2+K_2)\cap W\neq\emptyset\} \\
&\le \overline{V}((x_1+K_1)\cap(x_2+K_2)\cap W).
\end{aligned}$$

Together with similar arguments as above this yields that, for $i,j,k\in\{0,\dots,d\}$,

$$\begin{aligned}
&\int_{(\mathcal{K}^d)^2}\int_{(\mathbb{R}^d)^2}\big(\mathbb{E}\,\mathbf{1}\{D^2_{(x_1,K_1),(x_2,K_2)}\mathcal{V}(Z\cap W)\neq 0\}|D_{(x_1,K_1)}V_i(Z\cap W)D_{(x_1,K_1)}V_j(Z\cap W)|^3\big)^{2/3} \\
&\qquad\qquad \big(\mathbb{E}\,(D_{(x_1,K_1)}V_k(Z\cap W))^6\big)^{1/3}\,\mathrm{d}(x_1,x_2)\,\mathbb{Q}^2(\mathrm{d}(K_1,K_2)) \\
&\le \int_{(\mathcal{K}^d)^2}\int_{(\mathbb{R}^d)^2}\overline{V}((x_1+K_1)\cap(x_2+K_2)\cap W)\big(\mathbb{E}\,|D_{(x_1,K_1)}V_i(Z\cap W)D_{(x_1,K_1)}V_j(Z\cap W)|^3\big)^{2/3} \\
&\qquad\qquad \big(\mathbb{E}\,(D_{(x_1,K_1)}V_k(Z\cap W))^6\big)^{1/3}\,\mathrm{d}(x_1,x_2)\,\mathbb{Q}^2(\mathrm{d}(K_1,K_2)) \\
&\le C^6\mathbb{E}\int_{(\mathbb{R}^d)^2}\overline{V}((x_1+Z_1)\cap(x_2+Z_2)\cap W)\overline{V}((x_1+Z_1)\cap W)^6\,\mathrm{d}(x_1,x_2) \\
&\le C^6\mathbb{E}\,\overline{V}(Z_1)^7\overline{V}(Z_2)\overline{V}(W).
\end{aligned}$$

Combining this with (4.2) and (4.3) yields that γ_5 is at most of the order $\sqrt{\overline{V}(W)}/V_d(W)^{3/2}$. By [11, Lemma 3.7], there exists a dimension dependent constant $C_d\in(0,\infty)$ such that

$$\frac{\overline{V}(W)}{V_d(W)}\le C_d \quad \text{for all} \quad W\in\mathcal{K}^d \quad \text{with} \quad r(W)\ge 1.$$

This implies that γ_1, γ_2, γ_3, γ_4 and $\sqrt{\gamma_5}$ have at most the order $1/\sqrt{V_d(K)}$. It is known from [11, Theorem 3.1] that there exists a constant $C_\Sigma\in(0,\infty)$ only depending on d, γ, and $\mathbb{Q}$ such that

$$\left|\frac{\operatorname{Cov}(V_i(Z\cap W),V_j(Z\cap W))}{V_d(W)}-\sigma_{i,j}\right|\le C_\Sigma\frac{1}{r(W)}$$

for $i, j \in \{0, \dots, d\}$ and $W \in \mathcal{K}^d$ with $r(W) \geq 1$. Now Theorem 1.1 and Theorem 1.2 complete the proof. $\square$

4.3 Multivariate normal approximation for functionals of marked Poisson processes

In this subsection we establish a consequence of Theorem 1.1, Theorem 1.2, and Theorem 1.3, which can be seen as a multivariate version of Proposition 1.4 and Theorem 6.1 in [15]. This result will be used heavily in the companion paper [31], in order to deduce rates of normal approximation for Poisson functionals which may be expressed as sums of stabilizing score functions. We work in the context of marked Poisson processes, where $(\mathbb{M}, \mathcal{F}_{\mathbb{M}}, \lambda_{\mathbb{M}})$ denotes the probability space of marks. Let $\widehat{\mathbb{X}} := \mathbb{X} \times \mathbb{M}$, put $\widehat{\mathcal{F}}$ to be the product σ-field of $\mathcal{F}$ and $\mathcal{F}_{\mathbb{M}}$, and let $\widehat{\lambda}$ be the product measure of λ and $\lambda_{\mathbb{M}}$. Here, $(\mathbb{X}, \mathcal{F}, \lambda)$ is as before. For a given point $x \in \mathbb{X}$ we denote by M_x the corresponding random mark, which has distribution $\lambda_{\mathbb{M}}$ and which is independent of everything else.

Let $F = (F_1, \dots, F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals $F_1, \dots, F_m \in \operatorname{dom} D$ with $\mathbb{E}\, F_i = 0$, $i \in \{1, \dots, m\}$. Define for all $c, p \in (0, \infty)$,

$$\Gamma_1(c,p) := c^{\frac{2}{4+p}} \Bigg(\sum_{i=1}^m \int_{\mathbb{X}} \bigg(\int_{\mathbb{X}} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} F_i \neq 0)^{\frac{p}{16+4p}}\, \lambda(\mathrm{d}x_2) \bigg)^2 \lambda(\mathrm{d}x_1) \Bigg)^{1/2}$$

$$\Gamma_2(c,p) := c^{\frac{3}{4+p}} \sum_{i=1}^m \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)} F_i \neq 0)^{\frac{1+p}{4+p}}\, \lambda(\mathrm{d}x)$$

$$\Gamma_3(c,p) := c^{\frac{2}{4+p}} \Bigg(\sum_{i=1}^m 9 \int_{\mathbb{X}^2} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} F_i \neq 0)^{\frac{p}{8+2p}}\, \lambda^2(\mathrm{d}(x_1,x_2)) + \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)} F_i \neq 0)^{\frac{p}{4+p}}\, \lambda(\mathrm{d}x) \Bigg)^{1/2}$$

$$\Gamma_4(c,p) := c^{\frac{3}{4+p}} \Bigg(\sum_{i=1}^m 106 \int_{\mathbb{X}^2} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} F_i \neq 0)^{\frac{p-2}{12+3p}}\, \lambda^2(\mathrm{d}(x_1,x_2)) + \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)} F_i \neq 0)^{\frac{p-2}{4+p}}\, \lambda(\mathrm{d}x) \Bigg)^{1/2}.$$

Theorem 4.5. *Let $F = (F_1, \dots, F_m)$, $m \in \mathbb{N}$, be a vector of Poisson functionals $F_1, \dots, F_m \in \operatorname{dom} D$ with $\mathbb{E}\, F_i = 0$, $i \in \{1, \dots, m\}$, and assume that there are constants $c, p \in (0, \infty)$ such that*

$$\mathbb{E}\, |D_{(x,M_x)} F_i|^{4+p} \leq c, \quad \lambda\text{-a.e. } x \in \mathbb{X}, \tag{4.4}$$

and

$$\mathbb{E}\, |D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} F_i|^{4+p} \leq c, \quad \lambda^2\text{-a.e. } (x_1,x_2) \in \mathbb{X}^2, \tag{4.5}$$

for all $i \in \{1, \dots, m\}$.

(a) *For positive semi-definite* $\Sigma = (\sigma_{ij})_{i,j\in\{1,\dots,m\}} \in \mathbb{R}^{m\times m}$,

$$d_3(F, N_\Sigma) \le \frac{m}{2}\sum_{i,j=1}^{m} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)| + \frac{3m^{3/2}}{2}\Gamma_1(c,p) + \frac{m^2}{4}\Gamma_2(c,p).$$

(b) *For positive definite* $\Sigma \in \mathbb{R}^{m\times m}$,

$$\begin{aligned} d_2(F, N_\Sigma) \le{}& \|\Sigma^{-1}\|_{op}\|\Sigma\|_{op}^{1/2}\sum_{i,j=1}^{m} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)| + 3\|\Sigma^{-1}\|_{op}\|\Sigma\|_{op}\sqrt{m}\Gamma_1(c,p) \\ &+ \frac{\sqrt{2\pi}}{8}\|\Sigma^{-1}\|_{op}^{3/2}\|\Sigma\|_{op} m^2\Gamma_2(c,p). \end{aligned}$$

(c) *Let* $\Sigma \in \mathbb{R}^{m\times m}$ *be positive definite and assume that* $p > 2$. *For any* $\ell \in \mathbb{N}$,

$$\begin{aligned} d_{\mathbb{H}_\ell}(F, N_\Sigma) \le 718 m^{65/24}\ell\|\Sigma^{-1}\|_{op}\max\Bigg\{&\sum_{i,j\in\{1,\dots,m\}} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)|, \Gamma_1(c,p), \\ &\Gamma_3(c,p), \frac{\sqrt{\ell}\sqrt{\Gamma_4(c,p)}}{\|\Sigma^{-1}\|_{op}^{1/4}}\Bigg\}. \end{aligned}$$

(d) *Let* $\Sigma \in \mathbb{R}^{m\times m}$ *be positive definite and assume that* $p > 2$ *and that there is a constant* $\varrho \in (0,\infty)$ *such that, for* $i \in \{1,\dots,m\}$ *and* $\widehat{\lambda}$*-a.e.* $\widehat{x} \in \widehat{\mathbb{X}}$, $|D_{\widehat{x}}F_i| \le \varrho$ $\mathbb{P}$*-a.s. Then,*

$$\begin{aligned} d_{convex}(F, N_\Sigma) \le 2304 m^5\|\Sigma^{-1}\|_{op}\max\Bigg\{&\sum_{i,j\in\{1,\dots,m\}} |\sigma_{ij} - \mathrm{Cov}(F_i, F_j)|, \Gamma_1(c,p), \\ &\Gamma_3(c,p), \frac{8\sqrt{6}}{3}\|\Sigma^{-1}\|_{op}^{1/2}\varrho^3\lambda(A), \frac{\sqrt{\varrho^4\lambda(A)}}{\|\Sigma^{-1}\|_{op}^{1/4}}, \\ &\frac{1}{m\|\Sigma^{-1}\|_{op}\lambda(A)}\int_{\mathbb{X}\setminus A}\mathbb{P}(D_{(x,M_x)}F \neq 0)\,\lambda(\mathrm{d}x)\Bigg\} \end{aligned}$$

for any $A \in \mathcal{F}$ *with* $0 < \lambda(A) < \infty$.

Proof. Obviously, Theorem 1.1, Theorem 1.2, and Theorem 1.3 can be also applied to marked Poisson processes. By combining the product form of $\widehat{\lambda}$ with the Cauchy-Schwarz inequality we obtain that

$$\begin{aligned} &\int_{\widehat{\mathbb{X}}^3}\left[\mathbb{E}\,(D^2_{\widehat{x}_1,\widehat{x}_3}F_i)^2(D^2_{\widehat{x}_2,\widehat{x}_3}F_i)^2\right]^{1/2}\left[\mathbb{E}\,(D_{\widehat{x}_1}F_j)^2(D_{\widehat{x}_2}F_j)^2\right]^{1/2}\widehat{\lambda}^3(\mathrm{d}(\widehat{x}_1,\widehat{x}_2,\widehat{x}_3)) \\ &= \int_{\mathbb{X}^3}\int_{\mathbb{M}^3}\left[\mathbb{E}\,(D^2_{(x_1,m_1),(x_3,m_3)}F_i)^2(D^2_{(x_2,m_2),(x_3,m_3)}F_i)^2\right]^{1/2}\left[\mathbb{E}\,(D_{(x_1,m_1)}F_j)^2(D_{(x_2,m_2)}F_j)^2\right]^{1/2} \\ &\qquad \lambda^3_{\mathbb{M}}(\mathrm{d}(m_1,m_2,m_3))\,\lambda^3(\mathrm{d}(x_1,x_2,x_3)) \end{aligned}$$

$$\le \int_{\mathbb{X}^3} \left[\int_{\mathbb{M}^3} \mathbb{E}\, (D^2_{(x_1,m_1),(x_3,m_3)} F_i)^2 (D^2_{(x_2,m_2),(x_3,m_3)} F_i)^2 \, \lambda^3_{\mathbb{M}}(\mathrm{d}(m_1,m_2,m_3)) \right]^{1/2}$$
$$\left[\int_{\mathbb{M}^3} \mathbb{E}\, (D_{(x_1,m_1)} F_j)^2 (D_{(x_2,m_2)} F_j)^2 \, \lambda^3_{\mathbb{M}}(\mathrm{d}(m_1,m_2,m_3)) \right]^{1/2} \lambda^3(\mathrm{d}(x_1,x_2,x_3))$$
$$= \int_{\mathbb{X}^3} \left[\mathbb{E}\, (D^2_{(x_1,M_{x_1}),(x_3,M_{x_3})} F_i)^2 (D^2_{(x_2,M_{x_2}),(x_3,M_{x_3})} F_i)^2\right]^{1/2} \left[\mathbb{E}\, (D_{(x_1,M_{x_1})} F_j)^2 (D_{(x_2,M_{x_2})} F_j)^2\right]^{1/2}$$
$$\lambda^3(\mathrm{d}(x_1,x_2,x_3)).$$

Since we can apply the same arguments to the other terms, we see that the bounds from Theorem 1.1, Theorem 1.2, and Theorem 1.3 are still valid if we integrate with respect to λ and always replace x_i by (x_i, M_{x_i}), where M_{x_i} is an independent random mark. We denote the corresponding versions of $\gamma_1, \dots, \gamma_5$ by $\hat{\gamma}_1, \dots, \hat{\gamma}_5$. For $i \in \{1, \dots, m\}$ and $q \in (0, 4+p)$ it follows from (4.4), (4.5), and Hölder's inequality that

$$\mathbb{E}\, |D_{(x,M_x)} F_i|^q \le c^{\frac{q}{4+p}} \mathbb{P}(D_{(x,M_x)} F_i \neq 0)^{\frac{4+p-q}{4+p}}, \quad \lambda\text{-a.e. } x \in \mathbb{X},$$

and

$$\mathbb{E}\, |D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} F_i|^q \le c^{\frac{q}{4+p}} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} F_i \neq 0)^{\frac{4+p-q}{4+p}}, \quad \lambda^2\text{-a.e. } (x_1,x_2) \in \mathbb{X}^2.$$

Moreover, for $i, j \in \{1, \dots, m\}$, $p > 2$, and λ^2-a.e. $(x,y) \in \mathbb{X}^2$,

$$\begin{aligned} \mathbb{E}\, \mathbf{1}\{D^2_{(x,M_x),(y,M_y)} F \neq 0\} |D_{(x,M_x)} F_i \, D_{(x,M_x)} F_j|^3 &\le c^{\frac{6}{4+p}} \mathbb{P}(D^2_{(x,M_x),(y,M_y)} F \neq 0)^{\frac{p-2}{4+p}} \\ &\le c^{\frac{6}{4+p}} \sum_{u=1}^m \mathbb{P}(D^2_{(x,M_x),(y,M_y)} F_u \neq 0)^{\frac{p-2}{4+p}}. \end{aligned}$$

Applying Hölder's inequality to separate expectations of products and using these inequalities, one obtains that

$$\hat{\gamma}_1 \le \sqrt{m}\Gamma_1(c,p),\ \hat{\gamma}_2 \le \sqrt{m}\Gamma_1(c,p),\ \hat{\gamma}_3 \le \Gamma_2(c,p),\ \hat{\gamma}_4 \le \sqrt{m}\Gamma_3(c,p), \text{ and } \hat{\gamma}_5 \le m^{3/2}\Gamma_4(c,p).$$

Combining these estimates with the marked versions of Theorem 1.1, Theorem 1.2 and Theorem 1.3 described above completes the proof of Theorem 4.5. $\square$

Finally, we remark that we may deduce Corollary 1.5 from Theorem 4.5 as follows. First we derive a version of Theorem 4.5 for $p = \infty$. Assume that $\lambda(\mathbb{X}) < \infty$ and that there exists a $\varrho \in (0, \infty)$ with

$$|D_{(x,M_x)} F_i| \le \varrho \quad \mathbb{P}\text{-a.s.}, \quad \lambda\text{-a.e. } x \in \mathbb{X},$$

and

$$|D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} F_i| \le \varrho \quad \mathbb{P}\text{-a.s.}, \quad \lambda^2\text{-a.e. } (x_1,x_2) \in \mathbb{X}^2,$$

for $i \in \{1, \dots, m\}$. Then, for any $p \in (0, \infty)$, the assumptions (4.4) and (4.5) of Theorem 4.5 are satisfied with $c = \varrho^{4+p}$. The dominated convergence theorem yields that, for $j \in \{1, \dots, 4\}$, $\lim_{p\to\infty} \Gamma_j(\varrho^{4+p}, p) = \Gamma_j(\varrho)$, where $\Gamma_j(\varrho)$ is obtained from $\Gamma_j(c, p)$ by replacing $c^{1/(4+p)}$ by ϱ and the exponents of the probabilities by their limits for $p \to \infty$. Consequently, the bounds from Theorem 4.5 hold with $\Gamma_j(\varrho)$ instead of $\Gamma_j(c, p)$ for $j \in \{1, \dots, 4\}$. For $\varrho = a/\sqrt{s}$ and $\lambda = s\mu$ and without marks, the assumptions (1.12) and (1.13) show that $\Gamma_j(\varrho)$, $j \in \{1, \dots, 4\}$, are all of order $s^{-1/2}$, which together with $\Sigma = \Sigma_s$, yields the conclusion of Corollary 1.5.

References

[1] R.N. Bhattacharya and S.P. Holmes (2010), An Exposition of Götze's Estimation of the Rate of Convergence in the Multivariate Central Limit Theorem, Chapter 7: An Application of Stein's Method, in: Normal Approximation and Asymptotic Expansions (eds. R.N. Bhattacharya and R.R. Rao), SIAM, Philadelphia.

[2] R.N. Bhattacharya and R.R. Rao (1976), *Normal Approximation and Asymptotic Expansions*, Wiley.

[3] S. Bourguin and G. Peccati (2014), Portmanteau inequalities on the Poisson space: mixed regimes and multidimensional clustering, *Electron. J. Probab.* **19**, 1–42.

[4] S. Bourguin and G. Peccati (2016), The Malliavin-Stein method on the Poisson space, in: Stochastic Analysis for Poisson Point Processes (eds. G. Peccati and M. Reitzner, Bocconi & Springer), 185–228.

[5] S. Chatterjee (2009), Fluctuations of eigenvalues and second order Poincaré inequalities, *Probab. Theory Related Fields* **143**, 1–40.

[6] L. Chen, L. Goldstein, and Q.-M. Shao (2011), *Normal Approximation by Stein's Method*, Springer.

[7] P. Eichelsbacher and C. Thäle (2014), New Berry-Esseen bounds for non-linear functionals of Poisson random measures, *Electron. J. Probab.* **19**, no. 102.

[8] X. Fang (2016), A multivariate CLT for bounded decomposable random vectors with the best known rate, *J. Theoretical Prob.* **29**, 1510–1523.

[9] X. Fang and A. Röllin (2015), Rates of convergence for multivariate normal approximation with applications to dense graphs and doubly indexed permutation statistics, *Bernoulli* **21**, 2157–2189.

[10] F. Götze (1991), On the rate of convergence in the multivariate CLT, *Ann. Probab.* **19**, 724–739.

[11] D. Hug, G. Last and M. Schulte (2016), Second order properties and central limit theorems for geometric functionals of Boolean models, *Ann. Appl. Probab.* **26**, 73–135.

[12] Y. T. Kim and H. S. Park (2015), Kolmogorov distance for multivariate normal approximation, *Korean J. Math.* **23**, 1–10.

[13] K. Krokowski and C. Thäle (2017), Multivariate central limit theorems for Rademacher functionals with applications, *Electron. J. Probab.* **22**, no. 87.

[14] R. Lachièze-Rey, M. Schulte, and J. E. Yukich (2017), Normal approximation for sums of stabilizing functionals, preprint, arxiv:1702.00726.

[15] G. Last, G. Peccati and M. Schulte (2016), Normal approximations on the Poisson space: Mehler's formula, second order Poincaré inequalities and stabilization, *Probab. Theory and Related Fields* **165**, 667–723.

[16] G. Last and M. Penrose (2017), *Lectures on the Poisson Process*, Cambridge University Press.

[17] G. Last, M. Penrose, M. Schulte and C. Thäle (2014), Moments and central limit theorems for some multivariate Poisson functionals, *Adv. Appl. Prob.* **46**, 348–364.

[18] E. Meckes (2009), On Stein's method for multivariate normal approximation, in: High Dimensional Probability V: The Luminy Volume, 153–178.

[19] I. Nourdin, G. Peccati, and A. Réveillac (2010), Multivariate normal approximation using Stein's method and Malliavin calculus, *Ann. Inst. Henri Poincaré Probab. Stat.* **46**, 45–58.

[20] I. Nourdin, G. Peccati, and Y. Swan (2014), Entropy and the fourth moment phenomenon, *J. Funct. Anal.* **266**, 3170–3207.

[21] G. Peccati and M. Reitzner (eds.) (2016), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer.

[22] G. Peccati, J. L. Sole, M. S. Taqqu, and F. Utzet (2010), Stein's method and normal approximation of Poisson functionals, *Ann. Probab.* **38**, 443–478.

[23] G. Peccati and C. Zheng (2010), Multi-dimensional Gaussian fluctuations on the Poisson space, *Electron. J. Probab.* **15**, 1487–1527.

[24] M. Penrose and J. E. Yukich (2001), Central limit theorems for some graphs in computational geometry, *Ann. Appl. Probab.* **11**, 1005–1041.

[25] M. Penrose and J. E. Yukich (2005), Normal approximation in geometric probability, in: Stein's method and applications, 37–58, Singapore Univ. Press.

[26] G. Reinert and A. Röllin (2009), Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition, *Ann. Probab.* **37**, 2150–2173.

[27] Y. Rinott and V. Rotar (1996), A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph related statistics, *J. Multivariate Analysis* **56**, 333–350.

[28] V.V. Sazonov (1968): On the multi-dimensional central limit theorem, *Sankhyā Ser. A* **30**, 181–204.

[29] R. Schneider and W. Weil (2008), *Stochastic and Integral Geometry*, Springer.

[30] M. Schulte (2016), Normal approximation of Poisson functionals in Kolmogorov distance, *J. Theoret. Probab.* **29**, 96–117.

[31] M. Schulte and J. E. Yukich (2018), Rates of multivariate normal approximation for statistics in geometric probability, in preparation.

A Appendix: Malliavin calculus on the Poisson space

We recall the definitions of the Malliavin operators as well as some of their relations. For more details we refer to, for example, [15, Section 2].

We start with a pathwise product formula for the difference operator.

Lemma A.1. *For Poisson functionals F and G and $x \in \mathbb{X}$,*

$$D_x(FG) = (D_xF)G + F(D_xG) + (D_xF)(D_xG) \quad \mathbb{P}\text{-}a.s.$$

For $n \in \mathbb{N}$ let us denote by $I_n(g)$ the multiple Wiener-Itô integral of $g \in L^2(\lambda^n)$. Note that for $g \in L^2(\lambda^n)$, $n \in \mathbb{N}$, and $h \in L^2(\lambda^m)$, $m \in \mathbb{N}$,

$$\mathbb{E}\, I_n(g)I_m(h) = \mathbf{1}\{n = m\}n! \int_{\mathbb{X}^n} g(x)h(x)\, \lambda^n(\mathrm{d}x). \tag{A.1}$$

Any square integrable Poisson functional F has a so-called Wiener-Itô chaos expansion

$$F = \mathbb{E}\, F + \sum_{n=1}^{\infty} I_n(f_n),$$

where the functions $f_n \in L^2(\lambda^n)$, $n \in \mathbb{N}$, are symmetric and λ^n-a.e. uniquely defined and the right-hand side converges in $L^2(\mathbb{P})$. Together with (A.1) one sees that

$$\mathrm{Var} F = \sum_{n=1}^{\infty} n! \|f_n\|_n^2,$$

where $\| \cdot \|_n$ denotes the usual norm in $L^2(\lambda^n)$ for $n \in \mathbb{N}$.

If $F \in \operatorname{dom} D$ (see (1.2)), the difference operator defined in (1.1) satisfies the identity

$$D_x F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(x, \cdot)) \quad \mathbb{P}\text{-a.s.}$$

for λ-a.e. $x \in \mathbb{X}$. Here, $f_n(x, \cdot)$ denotes the function in $n-1$ variables one obtains after fixing the first argument to be x. Moreover, $F \in \operatorname{dom} D$ is equivalent to

$$\sum_{n=1}^{\infty} n\, n! \|f_n\|_n^2 < \infty.$$

The inverse Ornstein-Uhlenbeck generator of F is given by

$$L^{-1} F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f)$$

and is the pseudo-inverse of the Ornstein-Uhlenbeck generator L, which we do not need for our purposes. Next we present the definition of the Skorohod integral δ. We say that a random function $g : \mathbb{X} \to \mathbb{R}$ depending only on η such that

$$\mathbb{E} \int_{\mathbb{X}} g(x)^2 \, \lambda(\mathrm{d}x) < \infty \tag{A.2}$$

belongs to $\operatorname{dom} \delta$ if

$$g(x) = g_0(x) + \sum_{n=1}^{\infty} I_n(g_n(x, \cdot))$$

for λ-a.e. $x \in \mathbb{X}$ with functions $g_n \in L^2(\lambda^{n+1})$, $n \in \mathbb{N} \cup \{0\}$, such that

$$\sum_{n=0}^{\infty} (n+1)! \|\tilde{g}_n\|_{n+1}^2 < \infty.$$

Here, $\tilde{g}_n \in L^2(\lambda^{n+1})$ denotes the symmetrization

$$\tilde{g}_n(x_1, \dots, x_{n+1}) = \frac{1}{(n+1)!} \sum_{\pi \in \Pi(n+1)} g_n(x_{\pi(1)}, \dots, x_{\pi(n+1)})$$

of g_n, where $\Pi(n+1)$ stands for the set of all permutations of $\{1, \dots, n+1\}$. For $g \in \operatorname{dom} \delta$ the Skorohod integral $\delta(g)$ is defined as

$$\delta(g) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{g}_n),$$

i.e., δ maps a random function to a random variable. The difference operator and the Skorohod integral are adjoint operators in the sense that they satisfy the following well-known integration by parts formula.

Lemma A.2. *For $F \in \operatorname{dom} D$ and $g \in \operatorname{dom} \delta$,*

$$\mathbb{E} \int_{\mathbb{X}} D_x F g(x)\, \lambda(\mathrm{d}x) = \mathbb{E}\, F\delta(g).$$

The following lemma (see [15, Proposition 2.3 and Corollary 2.4]) provides a criterion for g belonging to $\operatorname{dom} \delta$ and an upper bound for the second moment of $\delta(g)$.

Lemma A.3. *Let g be a random function depending only on η and satisfying* (A.2) *and*

$$\mathbb{E} \int_{\mathbb{X}^2} (D_y g(x))^2\, \lambda^2(\mathrm{d}(x,y)) < \infty. \tag{A.3}$$

Then, $g \in \operatorname{dom} D$ and

$$\mathbb{E}\, \delta(g)^2 \leq \mathbb{E} \int_{\mathbb{X}} g(x)^2\, \lambda(\mathrm{d}x) + \mathbb{E} \int_{\mathbb{X}^2} (D_y g(x))^2\, \lambda^2(\mathrm{d}(x,y)).$$

In addition to Lemma A.2 we also make use of the following integration by parts formula involving indicator functions, which need not belong to $\operatorname{dom} D$. It is a consequence of [15, Lemma 2.2 and Proposition 2.3].

Lemma A.4. *Let F be a Poisson functional, let $A \subseteq \mathbb{R}$ be measurable, and let g be a random function depending only on η and satisfying* (A.2) *and* (A.3). *Assume that $D_x \mathbf{1}\{F \in A\} g(x) \geq 0$ for λ-a.e. $x \in \mathbb{X}$. Then, $g \in \operatorname{dom} \delta$ and*

$$\mathbb{E} \int_{\mathbb{X}} D_x \mathbf{1}\{F \in A\} g(x)\, \lambda(\mathrm{d}x) = \mathbb{E}\, \mathbf{1}\{F \in A\} \delta(g).$$