“Stein’s method and Malliavin calculus on the Poisson space”

by Giovanni PECCATI (Luxembourg University)

Université Paris-Est: November 18, 2010

Based on three joint works

Stein's method and normal approximation of Poisson functionals

(Peccati, Solé, Taqqu and Utzet, 2010)

Multidimensional Gaussian fluctuations on the Poisson space

(Peccati and Zheng, 2010)

Invariance principles for homogeneous sums: universality of Wiener chaos

(Nourdin, Peccati and Reinert, 2010)

[Overview, I]

In a recent series of papers, Nourdin and Peccati (2009, 2010), Nourdin, Peccati and Réveillac (2010) and Nourdin, Peccati and Reinert (2009) have shown that one can combine **Malliavin calculus** with **Stein's method** on a Gaussian space, in order to obtain **explicit bounds** in the **normal** and **Gamma approximations** of **non-linear functionals** of the underlying Gaussian field. This is related to previous works by Nualart and myself (2005) and Nualart and Ortiz-Latorre (2007).

Several applications, for instance: (i) first ever Berry-Esséen bounds in the **Breuer-Major CLT** for functionals of fractional Brownian motion, (ii) exact Berry-Esséen asymptotics for **continuous-time Toepliz quadratic forms**, (iii) generalized **second order Poincaré inequalities**, (iv) **Gaussian polymers**, (v) non-asymptotic **confidence intervals** in Hurst estimation.

[Overview, II]

In this talk: we present the results proved by Peccati, Solé, Taqqu and Utzet (2010), showing how to combine **Stein's method** and **Malliavin calculus on a Poisson space** (in the form of Nualart and Vives, 1989), in order to assess the normal approximation of random variables **in the Wasserstein distance**. This refines some findings by Peccati and Taqqu (2007, 2008ab) obtained by using **decoupling techniques**.

Motivation: e.g., some CLTs for non-linear functionals of **Lévy moving averages,** like Ornstein-Uhlenbeck Lévy processes. These results have applications to prior specification in **Bayesian survival analysis** (Peccati and Prünster, 2008; de Blasi, Peccati and Prünster, 2009).

[Overview, III]

In a final twist, we will show that some of our findings are indeed special instances of a very general **universality principle** for homogeneous sums, related to the notion of "**coordinate influence**" and to invariance principles by Rotar' (1979) and Mossel, O'Donnell and Oleszkiewicz (2010).

[General framework]

Let $(A, \mathcal{A}, \mu)$ be a Borel measure space, with μ a positive measure, σ-finite, non-atomic and s.t. $\mu(A) = \infty$.

We denote by $\hat{N} = \left\{ \hat{N}(B) : B \in \mathcal{A} : \mu(B) < \infty \right\}$ a **centered Poisson random measure** controlled by μ, that is, $\hat{N}$ is a random signed measure such that: **(i)** $\hat{N}(B)$ has a **centered Poisson distribution** with parameter $\mu(B)$, and **(ii)** for every B, C s.t. $B \cap C = \varnothing$, $\hat{N}(B)$ and $\hat{N}(C)$ are **independent**.

We denote by $L^2\left(\sigma\left(\hat{N}\right)\right)$ the space of all real-valued square-integrable functionals of $\hat{N}$.

[Chaos]

For every $d \geq 1$, denote by $L_s^2\left(\mu^d\right)$ the space of (deterministic) symmetric functions on A^d, that are square-integrable w.r.t. μ^d.

The following **Wiener-Itô chaotic decomposition** is well-known: every $F \in L^2\left(\sigma\left(\hat{N}\right)\right)$ admits a unique representation of the type

$$\boxed{F = E\left(F\right) + \sum_{d=1}^{\infty} I_d\left(f_d\right) \quad \text{(convergence in } L^2\text{)},}$$

where $f_d \in L_s^2\left(\mu^d\right)$, and I_d stands for a **multiple Wiener-Itô integral** of order d.

One can think at the r.v. $I_d\left(f_d\right)$ as a d-fold iterated integral w.r.t. $\hat{N}$, over a region of A^d "not containing diagonals".

[The problem]

Given square-integrable random variables F, G, we define the **Wasserstein distance**, between the laws of F and G, as

$$d_W(F,G) = \sup_{|h'|\leq 1} |E[h(F)] - E[h(G)]| .$$

Let $Z \sim \mathbf{N}(0,1)$.

Problem. Given $F \in L^2\left(\sigma\left(\hat{N}\right)\right)$ such that $E(F) = 0$, provide an **estimation** (e.g., an **upper bound**) of the quantity

$$\boxed{d_W(F,Z)}$$

Idea. Use **Stein's method** and **Malliavin calculus**.

[Stein's method in a nutshell, I: Stein's Lemma]

(Ref. : Stein 1972, 1986)

Stein's Lemma. *A random variable* F *has a* $\mathbf{N}(0,1)$ *distribution if and only if for every smooth function* g

$$E\left[g'(F) - Fg(F)\right] = 0.$$

Heuristically, Stein's Lemma suggests that, if F is such that

$$E\left[g'(F) - Fg(F)\right] \simeq 0$$

for a "**sufficiently large**" class of smooth functions g, then the law of F must be **close to Gaussian**.

[Stein's method in a nutshell, II: Stein equations]

Formally: fix $h \in \mathbf{Lip}(1)$, take $Z \sim \mathbf{N}(0,1)$, and introduce the **Stein equation**

$$\boxed{g'(x) - xg(x) = h(x) - E[h(Z)], \quad x \in \mathbf{R}} \qquad (\star)$$

(the unknown being g).

Classic estimates by Stein (1986) yield that there exists a solution to $(\star)$, say g_h, such that

$$\boxed{\left|g_h'\right| \leq 1 \quad \text{and} \quad \left|g_h''\right| \leq 2}$$

[Stein's method in a nutshell, III: Stein's bounds]

The previous results show that, if $Z \sim \mathbf{N}(0,1)$, then

$$\begin{aligned} d_W(F,Z) &= \sup_{|h'|\le 1} |E[h(F)] - E[h(Z)]| \\ &= \sup_{|h'|\le 1} |E\{h(F) - E[h(Z)]\}| \\ &= \sup_{|h'|\le 1} \left|E\left[g_h'(F) - Fg_h(F)\right]\right| \\ &\le \sup_{\substack{|g'|\le 1 \\ |g''|\le 2}} \left|E\left[g'(F) - Fg(F)\right]\right|. \end{aligned}$$

which is known as the **Stein's bound** on the Wasserstein distance.

How to assess a quantity such as $\sup_g |E[g'(F) - Fg(F)]|$?

[The idea]

Our principal idea is to use **infinite-dimensional integration by parts formulae** (from Malliavin calculus) in order to write, for F centered and sufficiently smooth,

$$\begin{aligned} E\left[Fg\left(F\right)\right] &= E\left[g'\left(F\right)\times H\left(F\right)\right]+R_0\left(g;F\right) \\ \left|R_0\left(g;F\right)\right| &\leq R\left(F\right), \end{aligned}$$

where H and R **do not depend** on g. In this way, for $Z\sim\mathbf{N}\left(0,1\right)$,

$$d_W\left(F,Z\right)\leq \sup_{\substack{|g'|\leq 1\\ |g''|\leq 2}}\left|E\left[g'\left(F\right)-Fg\left(F\right)\right]\right|\leq E\left|1-H\left(F\right)\right|+R\left(F\right),$$

i.e.

$$\boxed{d_W\left(F,Z\right)\leq E\left|1-H\left(F\right)\right|+R\left(F\right)\leq\sqrt{E\left|1-H\left(F\right)\right|^2}+R\left(F\right).}$$

To explicitly write H and R, one must introduce **Malliavin operators**.

[Derivatives and Integrals]

We write $L^2(\mu \times P) = \left\{ u(t,\omega) : E \int_A u(t)^2 \mu(dt) < \infty \right\}$.

The **derivative operator** D is such that $\mathbf{dom}(D) \subset L^2\left(\sigma\left(\hat{N}\right)\right)$. For $F = E(F) + \sum_{d \geq 1} I_d(f_d) \in \mathbf{dom}(D)$

$$\boxed{D_t F = \sum_{d=1}^{\infty} d \times I_{d-1}(f_d(t,\cdot)), \quad t \in A.}$$

Alternatively, $D_t F(\omega) = F(\omega + \varepsilon_t) - F(\omega)$ (ε_t = Dirac mass at t).

We denote by δ the **adjoint** of D, called the **Skorohod integral**. One has that $\mathbf{dom}(\delta) \subset L^2(\mu \times P)$ and $\delta : \mathbf{dom}(\delta) \to L^2\left(\sigma\left(\hat{N}\right)\right)$. For every $u \in \mathbf{dom}(\delta)$ and $F \in \mathbf{dom}(D)$

$$\boxed{E[\delta(u) F] = E \int_A [u(t) D_t F] \mu(dt)}$$

[Ornstein-Uhlenbeck operators, I]

The **generator of the Ornstein-Uhlenbeck semigroup**, noted L, is such that $\mathbf{dom}\,(L) \subset L^2\left(\sigma\left(\hat{N}\right)\right)$, for $F = E\,(F) + \sum_{d\geq 1} I_d\,(f_d) \in \mathbf{dom}\,(L)$

$$\boxed{LF = -\sum_{d=1}^{\infty} d \times I_d\,(f_d)\,.}$$

The **pseudo-inverse** of L, noted L^{-1}, is such that, for every $F = E\,(F) + \sum_{d\geq 1} I_d\,(f_d) \in L^2\left(\sigma\left(\hat{N}\right)\right)$

$$\boxed{L^{-1}F = -\sum_{d=1}^{\infty} \frac{1}{d} \times I_d\,(f_d).}$$

Plainly, $L^{-1}F = L^{-1}\,(F - E\,(F))$, and

$$\boxed{LL^{-1}F = F - E\,(F)\,.}$$

[Ornstein-Uhlenbeck operators, II]

One fundamental relation is the following

$$\boxed{\delta D = -L,}$$

yielding that, for every centered $F \in L^2\left(\sigma\left(\hat{N}\right)\right)$

$$\boxed{F = LL^{-1}F = -\delta DL^{-1}F = \delta\left(-DL^{-1}F\right).}$$

[The crucial computation]

Observe first that, if $|g'| \leq 1$ and $|g''| \leq 2$ and $F \in \mathbf{dom}(D)$, then (by using the **difference representation** of D)

$$D_t g(F) = g'(F) D_t F + R(D_t F), \quad |R(y)| \leq y^2.$$

It follows that, if $F \in \mathbf{dom}(D)$ is also centered,

$$\begin{aligned} E[Fg(F)] &= E\left(\delta\left(-DL^{-1}F\right) g(F)\right) \\ &= E\int_A \left[-D_t L^{-1}F \times D_t g(F)\right] \mu(dt) \\ &= E\left[g'(F)\int_A \left[-D_t L^{-1}F \times D_t F\right] \mu(dt)\right] \\ &\quad +E\left[\int_A \left[-D_t L^{-1}F \times R(D_t F)\right] \mu(dt)\right] \end{aligned}$$

[A general bound]

If $Z \sim \mathbf{N}(0,1)$, then

$$
\begin{aligned}
d_W(F,Z) &\leq \sup_{|g'|\leq 1;|g''|\leq 2} \left| E\left[g'(F) - Fg(F)\right]\right| \\
&\leq E\left|1 - \left\langle DF, -DL^{-1}F\right\rangle_\mu\right| + E\int_A [D_tF]^2 \left|D_tL^{-1}F\right| \mu(dt).
\end{aligned}
$$

If $F = I_d(f)$, then $L^{-1}F = -\frac{1}{d}F$, and

$$
\begin{aligned}
\left\langle DF, -DL^{-1}F\right\rangle_\mu &= \frac{1}{d}\|DF\|_\mu^2 \\
\int_A [D_tF]^2 \left|D_tL^{-1}F\right| \mu(dt) &= \frac{1}{d}\int_A |D_tF|^3 \mu(dt).
\end{aligned}
$$

[First example]

Consider a function $h \in L^2(\mu)$, as well as the stochastic integral

$$F = I_1(h) = \int_A h(x)\,\hat{N}(dx).$$

Observe: $EF^2 = \|h\|_\mu^2$. Then, for $Z \sim \mathbf{N}(0,1)$,

$$d_W(F,Z) \leq \left|1 - \|h\|_\mu^2\right| + \int_A |h(x)|^3\,\mu(dx).$$

For instance, if $\left\{\hat{N}_t : t \geq 0\right\}$ is a centered Poisson process with intensity 1,

$$d_W\left(\hat{N}_t/\sqrt{t}, Z\right) \leq \frac{1}{\sqrt{t}}$$

(not surprising, but consistent with usual Berry-Esséen estimates).

[The example of double integrals]

Consider a function $h \in L_s^2\left(\mu^2\right)$, and define the **double stochastic integral**

$$F = I_2(h) = \int_A \int_A h(x,y)\, \hat{N}(dx)\, \hat{N}(dy) \quad \text{(no diagonals)}.$$

One has $EF^2 = 2\,\|h\|_{\mu^2}^2$. For $Z \sim \mathbf{N}(0,1)$,

$$\begin{aligned} d_W(F,Z) \;\; &\leq \;\; \left|1 - 2\,\|h\|_{\mu}^2\right| + C_1\sqrt{\int_A \int_A \left[h *_1^1 h\ \ (x,y)\right]^2 \mu(dx)\,\mu(dy)} \\ &\quad + C_2\sqrt{\int_A \left[h *_2^1 h\ \ (x)\right]^2 \mu(dx)} \\ &\quad + C_3\sqrt{\int_A \int_A h(x,y)^4\, \mu(dx)\,\mu(dy)}, \end{aligned}$$

where C_1, C_2, C_3 are explicit universal constants.

[Contractions]

Here,

$$h \ast_1^1 h(x,y) = \int_A h(t,x)\, h(t,y)\, \mu(dt),$$

is the usual **contraction** of h with itself, whereas

$$h \ast_2^1 h(x) = \int_A h(t,x)\, h(t,x)\, \mu(dt)$$

is a **degenerate** contraction, obtained as the restriction of $h \ast_1^1 h$ on $\{x = y\}$.

From the previous bound, one deduces a **CLT**: if $EI_2(h_k)^2 \to 1$, $\left\| h_k \ast_1^1 h_k \right\|_{\mu^2} \to 0$, $\left\| h_k \ast_2^1 h_k \right\|_{\mu} \to 0$ and $\int h_k^4 d\mu^2 \to 0$, then

$$I_2(h_k) \xrightarrow{\mathbf{LAW}} Z \sim \mathbf{N}(0,1).$$

[Decoupling]

The **same CLT** was obtained (without explicit bounds) by Peccati and Taqqu (2007, 2008ab) by a **decoupling technique**, named the **principle of conditioning** (**POC** – Jakubowski, 1987). The POC implies that to establish the weak convergence of random variables of the type

$$\sum_{ij} h(i,j)\,\xi_i \xi_j, \quad \{\xi_i\} \text{ independent}$$

it is sufficient to consider

$$\sum_{ij} h(i,j)\,\xi_i \xi_j', \quad \left\{\xi_i'\right\} \text{ independent copy of } \{\xi_i\}\,.$$

Therefore, one can study $I_2(h_k)$ by first considering $\int_{A^2} h_k d\hat{N} d\hat{N}'$, where: $\hat{N}'$ independent copy of $\hat{N}$.

[Higher order integrals]

If

$$F = I_d(h) = \int_{A^d} h(x_1, ..., x_d)\, \hat{N}(dx_1) \cdots \hat{N}(dx_d) \quad \text{(no diagonals)},$$

then a similar bounds holds: for $Z \sim \mathbf{N}(0,1)$

$$\begin{aligned} d_W(F,Z) \ \leq \ & \left|1 - EF^2\right| + \ C \times \sum_{r=1}^{d-1} \|h \star_r^r h\|_{\mu^{2d-2r}} + \\ & + \textbf{norms of degenerate contractions}\,(r = 1, ..., d-1) \\ & + \sqrt{\int_{A^d} h^4 d\mu^d}. \end{aligned}$$

These bounds yield **explicit sufficient conditions for CLTs**. It seems unfeasible to obtain the same CLTs by decoupling.

[Applications in Bayesian survival analysis]

These results can be naturally applied in order to establish CLTs for objects of the type

$$\int_0^T M_t^p dt, \quad T \to \infty,$$

where M is a **Lévy moving average**, that is

$$M_t = \int_0^t k\,(t-s)\,dL_s, \quad L = \text{ Lévy process.}$$

For instance $k\,(t-s) = \sqrt{2\lambda}\,\text{exp}\,[-\lambda\,(t-s)]$ gives the **Ornstein-Uhlebeck Lévy processes.** **CLTs** of this type are useful in **prior specification in Bayesian non-parametric survival analysis**, where moving averages model **random hazard rates**. **See**: Peccati and Prünster (2008; prior), de Blasi, Peccati and Prünster (2008; posterior)

[Multidimensional case]

In Peccati and Zheng (2010), one can find multidimensional extensions of these results. These results partially generalize findings known in the Gaussian world, namely that **inside Wiener chaoses componentwise convergence to Gaussian implies joint convergence**.

By replacing Stein's method with the so-called "smart path method" (popular in spin glasses), we can deal with covariance matrices that are not necessarily positive definite.

The problem of lifting these findings from a finite-dimensional framework to a functional level is still open.

[Comparing Gaussian and Poisson chaos, I]

See: Nualart and Peccati (2005; i–iii), Nualart and Ortiz (2007; iv) and Nourdin and Peccati (2009, estimate); W = Gaussian field.

Theorem. *Suppose $F_k = I_d^W(h_k)$, is such that $EF_k^2 = 1$. Then, the following are equivalent:*

(i) $F_k \xrightarrow{\text{LAW}} Z \sim \mathbf{N}(0,1)$; **(ii)** $EF_k^4 \to 3$;

(iii) $\|h_k *_r^r h_k\| \to 0$, $r = 1, ..., d-1$; **(iv)** $\frac{1}{d}\|DF_k\|^2 \to 1$ *in* L^2.

Moreover,

$$d_W(F_k, Z) \le E\left|1 - \frac{1}{d}\|DF_k\|^2\right| \le \sqrt{\frac{\left|EF_k^4 - 3\right|}{3}}.$$

[Comparing Gaussian and Poisson chaos, II]

Here is what we can deduce from our results on Poisson functionals.

Theorem. *Suppose* $F_k = I_d^{\hat{N}}(h_k)$*, is such that* $EF_k^2 = 1$*. If* **(a)** $\|h *_r^r h\| \to 0$*,* $r = 1, ..., d-1$*,* **(b)** *all degenerate contractions converge to zero, and* **(c)** $\int h_k^4 d\mu^d \to 0$*, then*

$$F_k \xrightarrow{\mathbf{LAW}} Z \sim \mathbf{N}(0,1).$$

Moreover,

$$d_W(F_k, Z) \leq E\left|1 - \frac{1}{d}\|DF_k\|_\mu^2\right| + \frac{1}{d}\int |D_t F_k|^3 \mu(dt).$$

[Homogeneous sums]

Let $\hat{N} = \left\{ \hat{N}(i) : i \geq 1 \right\}$ be a sequence of i.i.d. of centered Poisson r.v.'s with parameter 1 (for simplicity). For $d \geq 2$, consider the sequence d-**homogeneous sum**

$$Q_d\left(K, \hat{N}\right) = \sum_{1 \leq i_1, \ldots, i_d \leq K} f_K(i_1, \ldots, i_d)\, \hat{N}(i_1) \cdots \hat{N}(i_d), \quad K \geq 1,$$

where $EQ_d\left(K, f_K, \hat{N}\right)^2 = 1$ and each f_K "vanishes on diagonals".

Let $G = \{G(i) : i \geq 1\}$ be an i.i.d. centered standard Gaussian sequence, and define analogously

$$Q_d(K, G) = \sum_{1 \leq i_1, \ldots, i_d \leq K} f_K(i_1, \ldots, i_d)\, G(i_1) \cdots G(i_d), \quad K \geq 1.$$

[Bounds for homogeneous sums]

In this framework, our results allow to deduce the (simpler) bounds: for $Z \sim \mathbf{N}(0,1)$,

$$\begin{aligned} d_W\left(Q_d\left(K,\hat{N}\right),Z\right) &\le C_1 \sum_{r=1}^{d-1} \|f_K *_r^r f_K\| \\ d_W\left(Q_d\left(K,G\right),Z\right) &\le C_2 \sum_{r=1}^{d-1} \|f_K *_r^r f_K\| \end{aligned}$$

In view of of the criterion by Nualart and Peccati (2005), one infers that the following implication holds, as $K \to \infty$:

$$\text{if } \; Q_d(K,G) \overset{\mathbf{LAW}}{\longrightarrow} Z, \; \text{ then } \; Q_d\left(K,\hat{N}\right) \overset{\mathbf{LAW}}{\longrightarrow} Z.$$

Is this an instance of a more general phenomenon?

[Answer: Universality of the Gaussian Wiener chaos]

The answer is: **Yes**.

Theorem (Nourdin Peccati Reinert). *With the above notation and for* $d \geq 2$*, suppose that*

$$Q_d(K,G) \xrightarrow{\mathbf{LAW}} Z \sim \mathbf{N}(0,1).$$

Then, for every sequence of independent r.v.'s $X = \{X(i) : i \geq 1\}$ *such that* **(a)** $EX(i) = 0$, **(b)** $EX(i)^2 = 1$ *and* **(c)** $\sup_i E|X(i)|^{2+\varepsilon}$ *(some* $\varepsilon > 0$*)*

$$Q_d(K,X) \xrightarrow{\mathbf{LAW}} Z \quad (\textit{in the Kolmogorov distance}).$$

If the $\{X(i)\}$ *are also identically distributed, then the requirement* **(c)** *can be dropped (at the cost of the Kolmogorov distance).*

NB: the conclusion is **false** for $d = 1$! **Also**: multidimensional versions and "Gamma versions".

[Influence functions and Lindeberg method]

Idea of the proof. Due to a remarkable invariance principle by Rotar' (1979) and Mossel, O'Donnel and Oleszkiewicz (2009), one knows that the "**maximum over influence functions**"

$$\tau_K = \max_i \sum_{a_1,\ldots,a_{d-1}} f_K\left(i, a_1, \ldots, a_{d-1}\right)^2$$

measures the distance between the laws of $Q_d(K,G)$ and $Q_d(K,X)$. This fact can be established by using the "Lindeberg method" for probabilistic approximations. One also knows that

$$\left\| f_K *_{d-1}^{d-1} f_K \right\|^2 = \sum_{i,j} \left[\sum_{a_1,\ldots,a_{d-1}} f_K\left(i, a_1, \ldots, a_{d-1}\right) f_K\left(j, a_1, \ldots, a_{d-1}\right) \right]^2 \geq \tau_K.$$

The conclusion follows once again from the previous criterion for CLT on a Wiener chaos.

[1] Peccati, Solé, Taqqu and Utzet (2010). Stein's method and normal approximation of Poisson functionals. *Ann. Probab.*

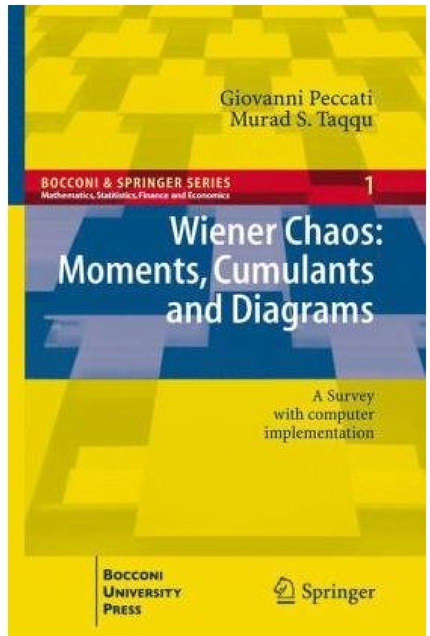
[2] Peccati and Zheng (2010). Multi-dimensional Gaussian fluctuations on the Poisson space. *Electronic J. Probab.*

[2] Nourdin, Peccati and Reinert (2010). Invariance principles for homogeneous sums: universality of Wiener chaos. *Ann. Probab.*

[3] Nourdin and Peccati (2009). Stein's method on Wiener chaos. *Probab. Theory Rel. Fields*.

[4] Nourdin and Peccati (2010). Stein's method and exact Berry-Esséen asymptotics for functionals of Gaussian fields. *Ann. Probab*.

[6] de Blasi, Peccati and Prünster (2009). Posteriors for random hazards. *Ann. Stat*.



Giovanni Peccati
Murad S. Taqqu
BOCCONI & SPRINGER SERIES
Mathematics, Statistics, Finance and Economics
1
Wiener Chaos:
Moments, Cumulants
and Diagrams
A Survey
with computer
implementation
BOCCONI
UNIVERSITY
PRESS
Springer