MALLIAVIN STRUCTURE FOR CONDITIONALLY INDEPENDENT RANDOM VARIABLES

Laurent Decreusefond, Christophe Vuong

MALLIAVIN STRUCTURE FOR CONDITIONALLY INDEPENDENT RANDOM VARIABLES

L. DECREUSEFOND AND C. VUONG

Abstract. On any denumerable product of probability spaces, we extend the discrete Malliavin structure for conditionally independent random variables. As a consequence, we obtain the chaos decomposition for functionals of conditionally independent random variables. We also show how to derive some concentration results in that framework. The Malliavin-Stein method yields Berry-Esseen bounds for U-Statistics of such random variables. It leads to quantitative statements of conditional limit theorems: Lyapunov's central limit theorem, De Jong's limit theorem for multilinear forms. The latter is related to the fourth moment phenomenon. The final application consists of obtaining the rates of normal approximation for subhypergraph counts in random exchangeable hypergraphs including the Erdös-Rényi hypergraph model. The estimator of subhypergraph counts is an example of homogeneous sums for which we derive a new decomposition that extends the Hoeffding decomposition.

Contents

2010 *Mathematics Subject Classification.* Primary: 60H07.

Key words and phrases. Conditional independence, Dirichlet structure, Malliavin calculus, carré du champ operator, Glauber dynamics, concentration bounds, Stein's method, U-Statistics, Berry-Esseen bounds, random exchangeable hypergraphs, motif estimation, Hoeffding decomposition.

1. Introduction

Malliavin calculus is also known as the stochastic calculus of variations. At the very core of it, it considers a gradient on a measured space. The link between these the differential geometry and the measure is made through the so-called integration by parts formula. When the measured space is the Wiener space, i.e. the set of continuous functions with the Brownian measure, the gradient generalizes the usual gradient on $\mathbb{R}^N$ and the integration by parts yields an extension of the Itô integral. When the measured space is the set of point processes on the real half-line, equipped with the law of a Poisson process, the gradient becomes a difference operator and the integration by parts is nothing but an avatar of the Mecke formula. It is only very recently that, concomitantly, the situation where the measured space is a product space, i.e. if we deal with independent random variables, has been addressed (see [14, 9, 13]). By order of complexity, the next situation which can be analyzed is that of conditionally independent random variables. This is a very common structure as de Finetti's theorem says that an infinite sequence of random variables is exchangeable if and only if these random variables are conditionally independent. This is the key theorem to develop a theory on random hypergraphs as in [1].

The first definitions of gradient (denoted by D) and divergence we introduce below for conditionally independent random variables, bear strong formal similarities with those of [9]. The difference lies into the computations which rely heavily on conditional distributions given the latent variable, which is here called Z. We can then follow the classical development of the Malliavin calculus apparatus: gradient, divergence, chaos, number operator and Ornstein-Uhlenbeck semi-group (denoted by P_t). We can even describe the dynamics of the Markov process whose infinitesimal generator is the number operator. At a formal level, the computations are almost identical to those of [9] with expectations replaced by expectations given Z.

Nevertheless, for more advanced applications, namely functional identities like the covariance representation formula, we need to introduce a difference operator (see Definition 2.10) which appears more often than the gradient itself. It is in some sense a finer tool that the original gradient which is useful to define the Dirichlet structure (the Glauber process, the infinitesimal generator denoted by L, etc.) but no more. This is due to the fact that $D_a D_a = D_a$, which entails that L commutes with D, and thus we have $DP_t = P_t D$ in place of the usual formula $DP_t = e^{-t} P_t D$ which is the core formula to derive all functional inequalities in the Gaussian and Poisson cases. The difference operator Δ allows to recover the crucial e^{-t} factor (see Proposition 3.1).

The prevailing application of Malliavin calculus is nowadays, the evaluations of convergence rates via the Stein's method ([28, 8] and references therein). The question is to assess a bound of the distance between a target distribution (more often the Gaussian distribution) and the law of a deterministic transformation of a probability measure, called the initial distribution.

The Dirichlet structure is useful to construct the characterization of the target distribution and to obtain the so-called Malliavin-Stein representation formula [7]. The Malliavin gradient or the carré du champ operator on the space on which lives the initial distribution are of paramount importance to make the computations which yield the distance. In the historical version of the Stein's method, this step was achieved via exchangeable pairs or biased coupling. One of the key difference between the Gaussian case and so-called discrete situations (Poisson, Rademacher, independent random variables) is the chain rule formula: it is only in the former framework that $D\psi(F) = \psi'(F)\, DF$. For the other contexts, we need to resort to an approximate chain rule [33]. This is the role here of Lemma 4.4 and Lemma 4.7.

Motivated by the applications to random graphs statistics, we focus here on normal approximations of U-statistics as in [3, 23, 30, 35]. In passing, we extend the notion of U-statistics by allowing the coefficients to depend on the latent variable instead of being only deterministic. Following the strategy of [2], we establish a fourth moment theorem with remainder for such functionals. As an application, we apply our theorems to deduce results of asymptotic normality of subhypergraph counts in random hypergraphs.

The rest of the paper is organized as follows. The section 2 lays the foundations of the Malliavin framework. We derive some functional identities in section 3, specifically conditional versions of Poincaré inequality and McDiarmid's inequality. The section 4 presents results of normal approximation. In particular, the subsection 4.3 states a partial fourth moment theorem for U-statistics under mild assumptions. The aforementioned applications to hypergraph statistics are in Section 5.

2. Discrete Malliavin-Dirichlet structure

2.1. **Preliminaries.** Let A be an at most denumerable set equipped with the counting measure, and define:

$$\ell^2(A) := \left\{ u \,:\, A \to \mathbb{R},\ \sum_{a\in A} |u_a|^2 < \infty \right\} \text{ and } \langle u, v\rangle_{\ell^2(A)} := \sum_{a\in A} u_a v_a.$$

Let $(\Omega, \mathcal{T}, \mathbb{P})$ be a probability space, E_0 be a Polish space and $((E_a, \Upsilon_a),\, a \in A)$ be a family of Polish spaces such that

$$\begin{aligned} E_A &= \prod_{a\in A} E_a \\ \Omega &= E_0 \times E_A. \end{aligned} \tag{1}$$

The product probability space E_A is endowed with its Borel σ-algebra denoted $\Upsilon \subset \mathcal{T}$. Let Z an E_0-valued random variable. By Theorem 10.2.2 [12], all the subsequent conditional distributions in the paper admit regular versions. For any subset B of A, we denote the set $E_B := \prod_{b\in B} E_b$ and for $x \in E_A$, $x_B := (x_a, a \in B) \in E_B$ so that for $a \in B$, $x_a \in E_a$. We denote $x^B = (x_a, a \in A \setminus B)$. Let $X := (X_a)_{a\in A}$ be a sequence defined on $(\Omega, \mathcal{T}, \mathbb{P})$ of conditionally independent random variables given Z such that for all $a \in A$, X_a is an E_a-valued random variable, i.e.:

$$X_a \underset{Z}{\perp\!\!\!\perp} (X_b, b \in A \setminus \{a\}),$$

or, equivalently:

$$\mathbb{P}(X_a \in \cdot \,|\, \sigma((X_b, b \neq a), Z)) = \mathbb{P}(X_a \in \cdot \,|\, \sigma(Z)).$$

We denote by $\mathbf{P}$ the law of X and $\mathbf{P}^Z$ the law $\mathcal{L}(X|Z)$. See chapter 5 of [21] for a thorough review of conditional independence, and [32] for some limit theorems for conditionally independent random variables. We use the notation $\mathbb{E}$ for the expectation of a random variable. By the disintegration theorem, for $a \in A$, the conditional probability distribution of X_a given $\sigma(X^{\{a\}}) \vee \sigma(Z)$ admits a regular version $\mathbf{P}_a$. For $p \geq 1$, let us denote $L^p(E_A \to \mathbb{R}, \mathbf{P})$ the set of p-th-integrable functions on E_A with respect to the measure $\mathbf{P}$. It is equipped with the norm $\|\cdot\|_{L^p(E_A\to\mathbb{R},\mathbf{P})}$ defined for f a measurable function on E_A by $\|f\|_{L^p(E_A\to\mathbb{R},\mathbf{P})} := \int |f(x)|^p \mathbf{P}(\mathrm{d}x)$. For the sake of notations, $L^p(E_A)$ stands for the space of p-integrable functionals

$$L^p(E_A) := \Big\{ \omega \mapsto F(X(\omega)) : \omega \in \Omega, F \in L^p(E_A \to \mathbb{R}, \mathbf{P}) \Big\}.$$

In this respect, $L^\infty(E_A)$ is the space of bounded functionals. We shall write F in place of $F(X)$ for the sake of conciseness. We closely follow the usual construction of Malliavin calculus on that space.

Definition 2.1. A functional F is said to be cylindrical if there exists a finite subset $I \subset A$ and a functional F_I in $L^2(E_I)$ such that $\mathbb{E}[|F_I|^2] < +\infty$ and $F = F_I \circ r_I$, where r_I is the restriction operator:

$$r_I : E_A \longrightarrow E_I$$
$$(x_a, a \in A) \longmapsto (x_a, a \in I).$$

It is clear that the set of those functionals $\mathcal{S}$ is dense in $L^2(E_A)$. We set $L^2(A \times E_A)$ the Hilbert space of processes which are square-integrable with respect to the measure $\sum_{a\in A} \delta_a \otimes \mathbf{P}$, i.e.

$$L^2(A \times E_A) = \{U : \sum_{a\in A} \mathbb{E}\left[U_a(X)^2\right] < +\infty\},$$

equipped with the norm and inner product:

$$\|U\|_{L^2(A\times E_A)} := \sum_{a\in A} \mathbb{E}\left[U_a^2\right] \text{ and } \langle U, V\rangle_{L^2(A\times E_A)} := \sum_{a\in A} \mathbb{E}\left[U_a V_a\right].$$

Definition 2.2. The set of simple processes, denoted $\mathcal{S}_0(l^2(A))$ is the set of random variables defined on $A \times E_a$ of the form

$$U = \sum_{a\in A} U_a \mathbb{1}_a,$$

for $U_a \in \mathcal{S}$.

2.2. **Malliavin operators.**

Definition 2.3 (Discrete gradient)**.** For $F \in \mathcal{S}$, DF is the simple process of $L^2(A\times E_A)$ defined for all $a \in A$ by:

$$D_a F := F - \mathbb{E}\left[F \mid X^{\{a\}}, Z\right].$$

In particular, $\mathcal{S} \subset \operatorname{Dom} D$. Define the σ-field $\sigma(X^{\{a\}}) \vee \sigma(Z)$ by $\mathcal{G}^a$, so that

$$D_a F = F - \mathbb{E}\left[F \mid \mathcal{G}^a\right]. \tag{2}$$

Recall that for $K \subset A$, $X_K = (X_a, a \in K)$ and $X^K = (X_a, a \in A \setminus K)$. We shall write $\mathcal{G}^K = \sigma(X^K) \vee \sigma(Z)$ and $\mathcal{G}_K = \sigma(X_K) \vee \sigma(Z)$ for K a subset of A.

Lemma 2.1. *Let $(a, b) \in A^2$, $a \neq b$, for $F \in \operatorname{Dom} D$,*

(1) $D_a D_a F = D_a F$;
(2) $D_a D_b F = D_b D_a F$;
(3) $D_a \mathbb{E}\left[F \mid \mathcal{G}^b\right] = D_b \mathbb{E}\left[F \mid \mathcal{G}^a\right]$.

Proof of lemma 2.1. For $(a, b) \in A^2$, with $b \neq a$,

$$\begin{aligned} D_a D_b F &= D_b F - \mathbb{E}\left[D_b F \mid \mathcal{G}^a\right] \\ &= F - \mathbb{E}\left[F \mid \mathcal{G}^b\right] - \mathbb{E}\left[F \mid \mathcal{G}^a\right] + \mathbb{E}\left[\mathbb{E}\left[F \mid \mathcal{G}^b\right] \mid \mathcal{G}^a\right] \\ D_b D_a F &= D_a F - \mathbb{E}\left[D_a F \mid \mathcal{G}^b\right] + \mathbb{E}\left[\mathbb{E}\left[F \mid \mathcal{G}^a\right] \mid \mathcal{G}^b\right] \\ &= F - \mathbb{E}\left[F \mid \mathcal{G}^a\right] - \mathbb{E}\left[F \mid \mathcal{G}^b\right] + \mathbb{E}\left[\mathbb{E}\left[F \mid \mathcal{G}^a\right] \mid \mathcal{G}^b\right]. \end{aligned}$$

We note that:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[F(X) \mid \mathcal{G}^a\right] \mid \mathcal{G}^b\right] \\
&\quad = \int\int F(X_{A\setminus\{a,b\}}, x_a, x_b)\mathbf{P}_a((X_{A\setminus\{a,b\}}, Z), x_b, \mathrm{d}x_a)\mathbf{P}_b((X_{A\setminus\{a,b\}}, Z), \mathrm{d}x_b) \\
&\quad = \int\int F(X_{A\setminus\{a,b\}}, x_a, x_b)\mathbb{P}^{X_b|Z}(Z, dx_b)\mathbb{P}^{X_a|Z}(Z, \mathrm{d}x_a) \\
&\quad = \mathbb{E}\left[\mathbb{E}\left[F(X) \mid \mathcal{G}^b\right] \mid \mathcal{G}^a\right].
\end{aligned}$$

Hence, the equality follows. □

The key to the definition of the Malliavin framework is the so-called integration by parts.

Theorem 2.2 (Integration by parts I)**.** *Let $F \in \mathcal{S}$, for every simple process U,*

$$\langle DF, U\rangle_{L^2(E_A\times A)} = \mathbb{E}\left[F\sum_{a\in A} D_a U_a\right]. \tag{3}$$

Proof of theorem 2.2. We get:

$$\begin{aligned}
\langle DF, U\rangle_{L^2(E_A\times A)} &= \mathbb{E}\left[\sum_{a\in A} D_a F U_a\right] \\
&= \mathbb{E}\left[\sum_{a\in A}(F - \mathbb{E}\left[F \mid \mathcal{G}^a\right])U_a\right] \\
&= \sum_{a\in A}\mathbb{E}\left[F(U_a - \mathbb{E}\left[U_a \mid \mathcal{G}^a\right])\right] \\
&= \sum_{a\in A}\mathbb{E}\left[F D_a U_a\right],
\end{aligned}$$

by self-adjointness of the conditional expectation. □

Corollary 2.3 (Closability of the discrete gradient)**.** *The operator D is closable from $L^2(E_A)$ into $L^2(A \times E_A)$.*

Proof of corollary 2.3. The proof is analogous to the proof of closability of the gradient in [9, corollary 2.5] □

The domain of D in $L^2(E_A)$ is the closure of cylindrical functionals with respect to the norm:

$$\|F\|_{1,2} := \sqrt{\|F\|^2_{L^2(E_A)} + \|DF\|^2_{A\times L^2(E_A)}}.$$

The following lemma gives a way to define square-integrable functionals in $\operatorname{Dom} D$ that are not in $\mathcal{S}$.

Lemma 2.4. *If there exists a sequence $(F_n)_{n\in\mathbb{N}}$ of elements of* $\operatorname{Dom} D$ *such that*

(1) the sequence converges to F in $L^2(E_A)$,
(2) $\sup_{n\in\mathbb{N}} \|DF_n\|_{L^2(E_A\times A)} < +\infty$,

then F belongs to $\operatorname{Dom} D$ *and $DF = \lim_{n\to+\infty} DF_n$.*

Proof of lemma 2.4. Let $(F_n)_{n\in\mathbb{N}}$ a sequence in $L^2(E_A)$ with $\mathbf{P}$-a.s. limit F, then for $a \in A$,

$$\begin{aligned}
\mathbb{E}[|D_a F - D_a F_n|^2] &\le \mathbb{E}[|F - F_n|^2] + \mathbb{E}\left[|\mathbb{E}\left[F_n \mid \mathcal{G}^a\right] - \mathbb{E}\left[F \mid \mathcal{G}^a\right]|^2\right] \\
&\le \mathbb{E}[|F - F_n|^2] + \mathbb{E}\left[\mathbb{E}\left[|F - F_n|^2 \mid \mathcal{G}^a\right]\right] \text{ by Jensen's inequality} \\
&= 2\mathbb{E}[|F - F_n|^2] \xrightarrow{n\to+\infty} 0.
\end{aligned}$$

Let $(A_m)_{m\in\mathbb{N}}$ a family of subsets of A such that $\bigcup_{m\geq 0} A_m = A$ and $|A_m| = m$, then for all $m \in \mathbb{N}$, $(\sum_{a\in A_m} D_a F_n)_{n\in\mathbb{N}}$ converges in $L^2(E_A)$ to $\sum_{a\in A_m} D_a F$. We denote by D^m the operator on $L^2(E_A \times A)$ such that for $a \in A_m$, $D^m_a = D_a$ and otherwise D^m_a is the null operator. For $m \in \mathbb{N}$, $(D^m F_n)_{n\in\mathbb{N}}$ converges to $D^m F$ in $L^2(E_A \times A)$. Because of (2), by the uniform boundedness principle, DF is in $L^2(E_A \times A)$, and the result follows. □

Definition 2.4 (Divergence operator)**.** The domain of the divergence operator $\operatorname{Dom}\delta$ in $L^2(E_A)$ is the set of processes U in $L^2(E_A \times A)$ such that there exists δU satisfying the duality relation

$$\langle DF, U\rangle_{L^2(E_A\times A)} = \mathbb{E}[F\delta U], \text{ for all } F \in \operatorname{Dom} D. \tag{4}$$

Moreover, for any process U belonging to $\operatorname{Dom}\delta$, δU is the unique element of $L^2(E_A)$ characterized by that identity. The integration by parts formula entails that for every process $U \in \operatorname{Dom}\delta$,

$$\delta = \sum_{a\in A} D_a U_a. \tag{5}$$

Definition 2.5 (Ornstein-Uhlenbeck operator)**.** The Ornstein-Uhlenbeck operator, denoted by L is defined on its domain

$$\operatorname{Dom}\mathsf{L} = \left\{ F \in L^2(E_A) :\ \mathbb{E}\left[\left|\sum_{a\in A} D_a F\right|^2\right] < +\infty \right\} \supseteq \mathcal{S}$$

by

$$\mathsf{L}F := -\delta DF = -\sum_{a\in A} D_a F. \tag{6}$$

2.3. **Chaos decomposition.** The lemma 2.1 entails a chaos decomposition of $L^2(E_A)$ similar to the one in [13].

Theorem 2.5 (Chaos decomposition)**.** *For any $F \in L^2(E_A)$,*

$$F = \mathbb{E}\left[F \mid Z\right] + \sum_{n=1}^{+\infty} \pi_n(F), \tag{7}$$

where $(\pi_n)_{n\in\mathbb{N}}$ is a sequence of orthogonal projectors on $L^2(E_A)$.

Proof. One can notice that:

$$\mathbb{E}[D_a F(X)|\mathcal{G}^a] = D_a(\mathbb{E}[F|\mathcal{G}^a])F(X) = 0, \text{ for all } a \in A. \tag{8}$$

Let $(A_m)_{m\in\mathbb{N}}$ a family of finite subsets of A such that $|A_m| = m$ and $\bigcup_{m\in\mathbb{N}} A_m = A$. Let $m \in \mathbb{N}$, $\mathrm{Id}_{L^2(E_{A_m})} = \prod_{a\in A_m}(D_a + \mathbb{E}[\cdot|\mathcal{G}^a])$. Indeed, for all $a \in A_m$, $\mathrm{Id}_{\operatorname{Dom} D} = D_a + \mathbb{E}[\cdot|\mathcal{G}^a]$. Hence, by distributivity and by using lemma 2.1, the identity also reads off: $\mathrm{Id}_{L^2(E_{A_m})} = \sum_{n=0}^{m} \pi^m_n$, where

$$\pi^m_n := \sum_{J\subset A_m,\, |J|=n} \left(\prod_{b\in J} D_b\right)\left(\prod_{c\in A_m\setminus J} \mathbb{E}[\cdot|\mathcal{G}^c]\right) \quad \forall n \leq m. \tag{9}$$

Let $n \leq m$,

$$\begin{aligned}
\pi_n^m \pi_n^m &= \sum_{\substack{I\subset A_m\\|I|=n}} \sum_{\substack{J\subset A_m\\|J|=n}} \left(\prod_{b\in I} D_b\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\left(\prod_{d\in J} D_d\right)\left(\prod_{e\in A_m\setminus J} \mathbb{E}[\cdot|\mathcal{G}^e]\right)\\
&= \sum_{I\subset A_m,\,|I|=n} \sum_{J\subset A_m,\,|J|=n} \left(\prod_{b\in I} D_b \prod_{e\in A_m\setminus J} \mathbb{E}[\cdot|\mathcal{G}^e]\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c] \prod_{d\in J} D_d\right)\\
&= \sum_{\substack{I\subset A_m\\|I|=n}} \left(\prod_{b\in I} D_b \prod_{e\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c] \prod_{d\in I} D_d\right) \text{ by lemma 2.1}\\
&= \sum_{I\subset A_m,\,|I|=n} \left(\prod_{b\in I}\prod_{b\in I} D_b D_b\right)\left(\prod_{c\in A_m\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\mathbb{E}[\cdot|\mathcal{G}^c]\right) = \pi_n^m.
\end{aligned} \tag{10}$$

By convention $\pi_n^m(F) = 0$ for $n > m$. Analogously, for $n' \neq n$, $\pi_n^m \pi_{n'}^m = 0$. The operator π_n^m is continuous on $L^2(E_A)$. Hence, $(\pi_n^m)_{m\in\mathbb{N}}$ is a well-defined family of projectors on $L^2(E_A)$. Moreover, for all $n \in \mathbb{N}$ and $F \in L^2(E_A)$, we have $\sup_{m\in\mathbb{N}} \|\pi_n^m(F)\|_{L^2(E_A)} \leq \|F\|_{L^2(E_A)}$. Then, by the uniform boundedness principle,

$$\sup_{\substack{m\in\mathbb{N}\\ \|F\|_{L^2(E_A)}}} \|\pi_n^m(F)\|_{L^2(E_A)} < +\infty.$$

The pointwise limits of $(\pi_n^m(F))_{m\in\mathbb{N}}$ for $F \in L^2(E_A)$ define a bounded linear operator π_n on $L^2(E_A)$ for $n \in \mathbb{N}$. Thus:

$$L^2(E_A) = \bigoplus_{n=0}^{+\infty} \text{Im } \pi_n. \tag{11}$$

Given (9), for a functional $F \in \text{Dom}\,\mathsf{L}$, we have $\pi_0(F) = \mathbb{E}\,[F \mid Z]$. □

Lemma 2.6 (Spectral decomposition)**.** *Let $F \in L^2(E_A)$ of chaos decomposition*

$$F = \mathbb{E}\,[F \mid Z] + \sum_{n=1}^{+\infty} \pi_n(F).$$

(1) We say that F belongs to $\text{Dom}\,\mathsf{L}$ *whenever*

$$\sum_{n=1}^{+\infty} n^2 \|\pi_n(F)\|_{L^2(E_A)} < +\infty.$$

(2) The operator has a unit spectral gap, i.e. the spectrum of L *coincides with* $\mathbb{N}_0$.

$$L^2(E_A) = \bigoplus_{k=0}^{+\infty} \ker(\mathsf{L} + k\text{Id}). \tag{12}$$

(3) It is invertible from $L_0^2(E_A) = \{F \in L^2(E_A),\, \mathbb{E}\,[F \mid Z] = 0\}$ into itself.

Proof of lemma 2.6. Let us show that π_n is in the domain of L for all $n \in \mathbb{N}$. By summability,

$$
\begin{aligned}
|\sum_{a\in A} D_a \pi_n|^2 &= \left| \sum_{a\in A} D_a \sum_{I\subset A,\, |I|=n} \left(\prod_{b\in I} D_b\right) \left(\prod_{c\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\right|^2 \\
&= \left| \sum_{a\in A} \mathbb{1}_I(a) \sum_{I\subset A,\, |I|=n} \left(\prod_{b\in I} D_b\right) \left(\prod_{c\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\right|^2 \\
&= n^2 \left| \sum_{I\subset A,\, |I|=n} \left(\prod_{b\in I} D_b\right) \left(\prod_{c\in A\setminus I} \mathbb{E}[\cdot|\mathcal{G}^c]\right)\right|^2 \text{ since } |I| = n \\
&= n^2 |\pi_n|^2,
\end{aligned}
\tag{13}
$$

so for $F \in L^2(E_A)$, $\pi_n(F) \in \operatorname{Dom}\mathsf{L}$. Hence, because of the orthogonality of $(\operatorname{Im}\,\pi_n)_{n\in\mathbb{N}}$, $F \in \operatorname{Dom} L \iff \sum_{n=1}^{+\infty} n^2 \|\pi_n(F)\|_{L^2(E_A)} < +\infty$. With the same calculations, we get $\mathsf{L}\pi_n = -n\pi_n$. The spectrum of $-\mathsf{L}$ coincides with $\mathbb{N}$. Then, we deduce that:

$$
\mathsf{L} = \sum_{n=0}^{+\infty} -n\pi_n, \tag{14}
$$

and $\operatorname{Im}\pi_n \subset \ker(\mathsf{L} + \mathsf{n}\mathrm{Id})$. Because of the orthogonality of the kernels, we get $\operatorname{Im}\pi_n = \ker(\mathsf{L} + \mathsf{n}\mathrm{Id})$. Now let us prove the third item. The pseudoinverse L^{-1} is defined on its domain $\{F \in L^2(E_A) : \mathbb{E}[F \mid Z] = 0\}$ and reads $\sum_{n=1}^{+\infty} -\frac{\pi_n}{n}$. Then for $F \in \{G \in \operatorname{Dom}\mathsf{L} : \mathbb{E}[G \mid Z] = 0\}, \mathsf{L}^{-1}(\mathsf{L}F) = F$. □

Corollary 2.7. *For $k > 0$ and J a subset of A of cardinal k, let us denote by $\mathfrak{C}_k$ the space of functionals $\phi = \sum_{J\subset A, |J|=k} \psi_J$ such that:*

- *for every $J \subset A$, ψ_J is $\mathcal{F}_J$-measurable;*
- *for every $K \subset A$, $\mathbb{E}[\psi_J \mid \mathcal{G}_K] = 0$ unless $K \subset J$;*

then $\mathfrak{C}_k = \ker(\mathsf{L} + k\mathrm{Id})$.

Proof of corollary 2.7. From (9), for $J = (a_1, \ldots, a_n) \subset A$, the component ψ_J is $\mathcal{F}_J$-measurable. Let us compute the expression of the iterated gradient for F a $\mathcal{F}_J$-measurable function:

$$
\begin{aligned}
\prod_{a\in J} D_a F &= \sum_{k=0}^{|J|} (-1)^k \sum_{\substack{K\subseteq J \\ |K|=k}} \mathbb{E}\left[F \mid \mathcal{G}^K\right] \\
&= \sum_{L\subseteq J} (-1)^{|J|-|L|} \mathbb{E}\left[F \mid \mathcal{G}_L\right],
\end{aligned}
$$

where $\mathcal{G}^K = \sigma(X^K) \vee \sigma(Z)$ and $\mathcal{G}_L = \sigma(X_L) \vee \sigma(Z)$. In this view, we have the inclusion $\ker(\mathsf{L} + n\mathrm{Id}) = \operatorname{Im}\pi_n \subset \mathfrak{C}_n$ for $n \in \mathbb{N}$.

Conversely, let ϕ for which the properties above hold.

$$\begin{aligned}
\mathsf{L}\phi &= -\sum_{a\in A} D_a \sum_{J\subset A,|J|=n} \psi_J \\
&= -\sum_{a\in A}\sum_{J\subset A,|J|=n} (\psi_J - \mathbb{E}\left[\psi_J \mid \mathcal{G}^a\right]) \\
&= -\sum_{k\in A}\sum_{\substack{J\subset A,|J|=n\\ a\in J}} \psi_J \text{ because } \mathbb{E}\left[\psi_J \mid \mathcal{F}_{A\setminus\{a\}}\right] = 0 \text{ for } J \not\subset A\setminus\{a\} \\
&= -n \sum_{J\in A,|J|=n} \psi_J = -n\phi.
\end{aligned}$$

Therefore, $\mathfrak{C}_n = \ker(\mathsf{L} + n\mathrm{Id})$ for $n \geq 1$. □

2.4. **Dirichlet structure.** The map L can be viewed as the generator of a Glauber dynamics where the index set is a finite set of random variables indexed by A_m for $m > 1$. For practical term, we introduce a new index ∂ and $X_\partial = Z$ $\mathbb{P}$-a.s..

Definition 2.6 (Modified Glauber process)**.** Consider $(N(t))_{t\geq 0}$ a Poisson process on the half-line $[0, +\infty)$ of rate $|A_m| + 1$. Let $(X^{\circ A_m}(t))_{t\geq 0} = (X_a^{\circ A_m}(t), t \geq 0, a \in A)$ the process valued in E_A starting with $X^{\circ A_m}(0) = X$ which evolves according to the following rule. At jump time τ of the process,

- Choose randomly an index a in $A_m \sqcup \{\partial\}$ with equal probability.
- If $a \neq \partial$, replace $X_a^{\circ A_m}(\tau)$ with a conditionally independent random variable $X_a^{\backslash}$ distributed according to $\mathbf{P}_a((X^{\circ A_m}_{A\setminus\{a\}}(\tau), Z), \cdot)$, otherwise do nothing.

That Markov process has for infinitesimal generator L^{A_m}:

$$\mathsf{L}^{A_m} F = -\sum_{a\in A_m} D_a F.$$

Our aim is to show that the operator L is an infinitesimal generator, letting $m \to +\infty$. We recall the Hille-Yosida theorem [37].

Proposition 2.8 (Hille-Yosida)**.** *A linear operator L on $L^2(E_A)$ is the generator of a strongly continuous contraction semigroup on $L^2(E_A)$ if and only if*

(1) $\mathrm{Dom}\, L$ *is dense in* $L^2(E_A)$*;*

(2) L is dissipative, i.e. for any $\lambda > 0$, $F \in \mathrm{Dom}\, L$,

$$\|\lambda F - LF\|_{L^2(E_A)} \geq \lambda \|F\|_{L^2(E_A)};$$

(3) $\mathrm{Im}(\lambda Id - L)$ *is dense in* $L^2(E_A)$*.*

Theorem 2.9. L *is an infinitesimal generator on E_A of a strongly continuous contraction semigroup on $L^2(E_A)$.*

Proof of theorem 2.9. We know that $\mathcal{S}$ is dense in $L^2(E_A)$. As $\mathrm{Dom}\,\mathsf{L} \supset \mathcal{S}$, it is also dense in $L^2(E_A)$. Let A_m an increasing sequence (with respect to $\subset$) of subsets of A such that $\bigcup_{n\geq 1} A_m = A \cup \partial$ and $|A_m| = m$. Then $(\mathcal{F}_{A_m})_{m\in\mathbb{N}}$ is a filtration. For $F \in L^2(E_A)$, let $F_m = \mathbb{E}\left[F \mid \mathcal{F}_{A_m}\right]$. Since $(F_m)_{m\in\mathbb{N}}$ is a square-integrable $\mathcal{F}_A$-martingale, $(F_m)_{m\in\mathbb{N}}$ converges both almost surely and in $L^2(E_A)$ to F. For any $m \in \mathbb{N}$, F_m depends only on X_{A_m}. Because of the conditional independence of the random variables X_a given X_∂, for all $a \in A$, we get that $D_a F_m = \mathbb{E}\left[D_a F \mid \mathcal{F}_{A_m}\right]$.

Using that L_{A_m} is dissipative for all $m \in \mathbb{N}$, we have:

$$\lambda^2 \|F_m\|^2_{L^2(E_A)} \leq \|\lambda F_m - \mathsf{L}^{A_m} F_m\|^2_{L^2(E_A)} = \mathbb{E}\left[\left(\lambda F_m + \sum_{a\in A_n} D_a F_m\right)^2\right]$$
$$= \mathbb{E}\left[\left(\lambda F_m + \sum_{a\in A} D_a F_m\right)^2\right] \text{ because } D_a F_m = 0, \text{ if } a \notin A_m.$$
$$= \mathbb{E}\left[\mathbb{E}\left[\lambda F + \sum_{a\in A} D_a F \,\middle|\, \mathcal{F}_{A_m}\right]^2\right].$$

It means that the operator L is dissipative. Thus, by the Hille-Yosida theorem, L is the infinitesimal generator of a strongly continuous contraction semigroup on $L^2(E_A)$ denoted P. □

Lemma 2.10. *Let $F \in L^2(E_A)$, then*

$$\mathbb{E}\left[F(X^{\circ A_m}) \,\middle|\, X, Z\right] = P_t^{A_m} F \xrightarrow{\mathbf{P}-a.s.} P_t F$$

and

$$X^{\circ A_m} \xrightarrow{d} X^\circ.$$

Proof of lemma 2.10. The theorem 17.25 of [21, Trotter, Sova, Kurtz, Mackevičius] gives the convergence in distribution of $X^{\circ A_m}$ towards X° the Markov process associated to L, and the almost sure convergence of the semigroup. □

These are pieces of the Dirichlet structure with invariant measure $\mathbf{P}$ that we complete with the carré du champ operator. Here, we note that $\mathcal{S}$ is an algebra which is a core of $\operatorname{Dom}\mathsf{L}$.

Definition 2.7 (Carré du champ operator)**.** Let $F, G \in \mathcal{S}$. The bilinear map

$$\Gamma(F,G) := \frac{1}{2}\left\{\mathsf{L}(FG) - F\mathsf{L}G - G\mathsf{L}F\right\}$$

is well-defined, and called carré du champ operator of the Markov generator L.

By an argument of density, there exists an algebra $\mathcal{A} \supset \mathcal{S}$ maximal in the sense of inclusion such that the carré du champ operator acts on it.

Definition 2.8 (Dirichlet structure)**.** The associated Dirichlet structure defined on $(E_A, \Upsilon, \mathbf{P})$ is given by the quadruple $(X^\circ, \mathsf{L}, (P_t)_{t\geq 0}, \mathcal{E})$ where X° is a Markov process with values in E_A whose infinitesimal generator is L and its semigroup is P, i.e. for any $F \in L^\infty(E_A)$:

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t F = (\mathsf{L}P_t)F.$$

Furthermore, $\mathbf{P}^Z$ is the invariant (or stationary) distribution of X° given Z and the Dirichlet form is defined by

$$\mathcal{E}(F,G) = \mathbb{E}[\Gamma(F,G)].$$

It comes with the classical properties entailed by the spectral decomposition of L, including the Mehler's formula.

Lemma 2.11 (Mehler's formula)**.** *For any $F \in L^2(E_A)$,*

(1)

$$\begin{aligned} P_t F &= \mathbb{E}\left[F \mid Z\right] + \sum_{n=1}^{\infty} e^{-nt}\pi_n(F) \\ &= \mathbb{E}\left[F(X^\circ(t)) \mid X\right], \end{aligned} \tag{15}$$

In particular $P_t F \in \operatorname{Dom}\mathsf{L} \cap \operatorname{Dom}\mathsf{L}^{-1}$.

(2)

$$\lim_{t\to\infty} P_t F(X) = \mathbb{E}\left[F(X) \mid Z\right].$$

(3) The pseudoinverse of L *can be written:*

$$\mathsf{L}^{-1}F := -\int_0^{+\infty} P_t F \, \mathrm{d}t.$$

Proof of lemma 2.11. Since formally $P_t = e^{-t\mathsf{L}}$, we get the first line of (15) from the spectral decomposition of L. The second line is deduced from the definition of the Glauber dynamics and by passing to the limit. Then,

$$\begin{aligned} \mathbb{E}\left[F \mid Z\right] - F &= \lim_{t\to+\infty} P_t F - P_0 F \\ &= \int_0^{+\infty} \frac{\mathrm{d}}{\mathrm{d}t} P_t F \, \mathrm{d}t \\ &= \mathsf{L}\left(\int_0^{+\infty} P_t F \, \mathrm{d}t\right). \end{aligned}$$

Taking $\mathbb{E}\left[F \mid Z\right] = 0$, we get the expression of the pseudoinverse. □

Remark 2.9. By the chaos expansion, $P_t F$ can be defined as the limit in $L^2(E_A)$ of elements $(P_t F_n)_{n\in\mathbb{N}}$ for F_n in $\mathcal{S}$. Hence, it is sufficient to define the semigroup acting on a functional of some finite vector of random variables X_B, using the definition of the Glauber dynamics entailed by it.

The infinitesimal generator satisfies another integration by parts formula due to the Dirichlet structure which is the key to investigating the so-called fourth moment phenomenon.

Lemma 2.12 (Integration by parts II)**.** *For* $(F, G) \in \mathcal{A}^2$*,*

$$\mathcal{E}(F, G) = -\mathbb{E}[F\mathsf{L}G]. \tag{16}$$

We introduce to the difference operator which is associated to the Malliavin-Dirichlet structure at hand. That difference operator serves the same purpose as in [24] and [15] for computations in the proofs of the limit theorems.

Definition 2.10 (Difference operator)**.** Let $F : E_A \to \mathbb{R}$, for $a \in A$, we introduce the operator

$$\begin{aligned} \Delta^{\{a\}}F \,:\, E_A \times E_a &\longrightarrow \mathbb{R} \\ (x, x'_a) &\longmapsto f(x) - f(x^{\{a\}}, x'_a). \end{aligned}$$

For the sake of conciseness, we shall write $F^{\{a\}'} = F(X^{\{a\}}, X'_a)$.

Lemma 2.13. *For* F *a functional in* $\operatorname{Dom} D$*, the gradient also reads as:*

$$D_a F = \mathbb{E}\left[\Delta^{\{a\}}F(X, X'_a)|X, Z\right], \tag{17}$$

where X'_a *has the law of* X_a *given* Z *and is conditionally independent of* $X^{\{a\}}$ *given* Z*. Similarly,*

$$\Gamma(F, G) = \frac{1}{2}\sum_{a\in A} \mathbb{E}\left[\left(\Delta^{\{a\}}F(X, X'_a)\right)\left(\Delta^{\{a\}}G(X, X'_a)\right)|X, Z\right]. \tag{18}$$

Proof of lemma 2.13. We have

$$\mathbb{E}\left[F \mid \mathcal{G}^a\right] = \int F(X^{\{a\}}, x_a)\mathbf{P}_a(\,\mathrm{d}x_a).$$

Since $\sigma(X_a)$ is independent of $\sigma(X^{\{a\}})$ given $\sigma(Z)$, we obtain

$$\mathbb{E}\left[F \mid \mathcal{G}^a\right] = \int F(X^{\{a\}}, x_a)\mathbb{P}^{X_a|Z}(\,\mathrm{d}x_a).$$

Eqn.(18) is proved similarly. □

3. Functional identities

This section is devoted to classical functional identities obtained in the Malliavin framework. We follow the approach of [18] using a covariance identity based on difference operators to deduce concentration inequalities.

Proposition 3.1. *For $F \in L^2(E_A)$ and $a \in A$, then:*

$$D_a(P_tF) = e^{-t}\mathbb{E}\left[\Delta^{\{a\}}F(X^\circ(t), X'_a) \;\middle|\; X, Z\right] \tag{19}$$

where X' has the law of X given Z.

Proof of proposition 3.1. We consider the Glauber dynamics with index set a finite subset A_m of A, as the construction of process $(X^{\circ A_m}(t))_{t\in\mathbb{R}^+}$ is explicit in that case. Let $a \in A_m$, we denote by N_a the Poisson process of intensity 1 which represents the life duration of the a-th component in the dynamics of $X^{\circ A_m}(t)$, so:

$$X_a^{\circ A_m}(t) = \mathbb{1}_{\{\tau_a \geq t\}}X_a + \mathbb{1}_{\{\tau_a < t\}}X_a^\backprime,$$

where $\tau_a = \inf\{t \geq 0,\ N_a(t) \neq N_a(0)\}$ is the life duration of the a-the component of the original sequence, exponentially distributed with parameter 1 (independent of everything else) and $X_a^\backprime$ is conditionally independent of X given Z. Then:

$$\begin{aligned}
&D_aP_t^{A_m}F = P_t^{A_m}F - \mathbb{E}\left[P_t^{A_m}F \;\middle|\; \mathcal{G}_a\right]\\
&= P_t^{A_m}F - \mathbb{E}\left[\mathbb{E}\left[F(X^{\circ A_m}(t))|X, Z\right]\mathbb{1}_{\{t\leq\tau_a\}} \;\middle|\; \mathcal{G}_a\right] - \mathbb{E}\left[F(X^{\circ A_m}(t))\mathbb{1}_{\{t>\tau_a\}}|X, Z\right]\\
&= \mathbb{E}\left[F(X^{\circ A_m}(t))\mathbb{1}_{\{t\leq\tau_a\}}|X, Z\right] - \mathbb{E}\left[\mathbb{E}\left[F(X^{\circ A_m}(t))|X, Z\right]\mathbb{1}_{\{t\leq\tau_a\}} \;\middle|\; \mathcal{G}_a\right]\\
&= e^{-t}\mathbb{E}\left[\Delta^{\{a\}}F(X^{\circ A_m}(t), X'_a) \;\middle|\; X, Z\right]
\end{aligned}$$

because the law of $X_a^\backprime$ given X is the same as the one of X'_a given X.

On one hand,

$$D_aP_t^{A_m}F \xrightarrow{\mathbb{P}-a.s.} D_aP_tF.$$

On the other hand, by the Skorohod's representation theorem, there exist copies of $X^{\circ A_m}$ and X° on a common probability space $(\tilde{\Omega}, \tilde{\mathcal{T}}, \tilde{\mathbb{P}})$ such that the sequence $(X^{\circ A_m})_{m\in\mathbb{N}}$ converges to X° $\tilde{\mathbb{P}}$-a.s. As the whole structure is invariant by copy, we can suppose the almost sure convergence on $(\Omega, \mathcal{T}, \mathbb{P})$, and the relation passes to the limit. □

Remark 3.1. In the case, we have only one random variable (or one particle), then the commutation relation simplifies to $D_a(P_tF) = D_a$.

Corollary 3.2 (Conditional covariance identity)**.** *For any $F, G \in L^2(E_A)$, then:*

$$\mathrm{cov}(F, G|Z) = \int_0^\infty e^{-t}\sum_{a\in A}\mathbb{E}\left[(D_aF)(\Delta^{\{a\}}G(X^\circ(t), X'_a)) \;\middle|\; Z\right]\mathrm{d}t. \tag{20}$$

Proof of corollary 3.2. We use the following conditional covariance formula analogous to the covariance formula:

$$\operatorname{cov}(F, G|Z) = \mathbb{E}\left[FG \mid Z\right] = \mathbb{E}\left[F\mathsf{L}\mathsf{L}^{-1}G \,\middle|\, Z\right]. \tag{21}$$

By the integration by parts I (3) which also holds with conditional expectation given Z, we get:

$$\begin{aligned}
\mathbb{E}\left[F\mathsf{L}\mathsf{L}^{-1}G \,\middle|\, Z\right] &= -\sum_{a\in A}\mathbb{E}\left[(D_aF)(D_a\mathsf{L}^{-1}G) \,\middle|\, Z\right] \\
&= -\sum_{a\in A}\mathbb{E}\left[(D_aF)(D_a\int_0^\infty P_tG\ \mathrm{d}t) \,\middle|\, Z\right] \\
&= -\sum_{a\in A}\mathbb{E}\left[(D_aF)\left(\int_0^\infty D_aP_tG\ \mathrm{d}t\right) \,\middle|\, Z\right] \\
&= -\int_0^\infty e^{-t}\sum_{a\in A}\mathbb{E}\left[(D_aF)\mathbb{E}\left[\Delta^{\{a\}}G(X^\circ(t), X'_a)|X, Z\right] \,\middle|\, Z\right]\mathrm{d}t,
\end{aligned}$$

using (19). □

As an immediate consequence of the spectral gap, we find another proof of the Efron-Stein inequality which is of independent interest.

Proposition 3.3. *If $F \in \mathfrak{C}_p$ then*

$$\operatorname{var}[F] = \frac{1}{p}\mathcal{E}(F) = \frac{1}{p}\|DF\|_{L^2(E_A)}.$$

Moreover, if there exist $F_1, \dots, F_m \in L^2(E_A)$ such that $F = \sum_{p=1}^m F_p$ with $F_p \in \mathfrak{C}_p$ for $p \in [\![1, m]\!]$, then:

$$\operatorname{var}[F] \le \|DF\|_{L^2(E_A)}. \tag{22}$$

Proof of proposition 3.3. Let us use the previous covariance identity, we have:

$$\begin{aligned}
\operatorname{var}[F] = \operatorname{cov}(F, F) &= \mathbb{E}[\Gamma(F, -\mathsf{L}^{-1}F)] \\
&= \mathbb{E}\left[\Gamma\left(\sum_{p=1}^m F_p, \sum_{q=1}^m \frac{1}{q}F_q\right)\right] \\
&= \sum_{p=1}^m\sum_{q=1}^m \frac{1}{q}\mathbb{E}\left[\Gamma\left(F_p, F_q\right)\right] \\
&= \sum_{p=1}^m \frac{1}{p}\mathbb{E}\left[\Gamma\left(F_p, F_p\right)\right] \text{ because } \mathbb{E}[\Gamma(F_p, F_q)] = 0 \text{ for } q \neq p.
\end{aligned}$$

It yields the inequality (22) noting that $\Gamma(F_p, F_p) \ge 0$ for all $p > 0$. □

We now deduce the conditional first-order Poincaré inequality for functionals of conditionally independent random variables. The equivalent for functionals of independent random variables is rather known as the Efron-Stein inequality in the literature [16].

Theorem 3.4 (Conditional Efron-Stein inequality)**.** *For $F \in L^2(E_A)$ such that $\mathbb{E}\left[F \mid Z\right] = 0$,*

$$\operatorname{var}[F|Z] \le \mathbb{E}\left[\Gamma(F, F) \mid Z\right]. \tag{23}$$

Proof of theorem 3.4. The conditional covariance formula yields

$$\begin{aligned}\mathrm{var}[F|Z] &= \int_0^\infty e^{-u} \sum_{a\in A} \mathbb{E}\left[(D_a F)(\Delta^{\{a\}} F)(X_u^\circ, X'_a) \,\middle|\, Z\right] \mathrm{d}u \\ &\leq \int_0^\infty e^{-u} \sqrt{\sum_{a\in A} \mathbb{E}[(D_a F)^2 | Z]} \sqrt{\sum_{a\in A} \mathbb{E}[\mathbb{E}\left[(\Delta^{\{a\}} F)(X_u^\circ, X'_a)|X, Z\right]^2 |Z]}\, \mathrm{d}u.\end{aligned}$$

The invariance of $\mathbf{P}^Z$ under the Glauber dynamics entails that

$$\sum_{a\in A} \mathbb{E}\left[\mathbb{E}\left[(\Delta^{\{a\}} F)(X_u^\circ, X'_a)|X, Z\right]^2 |Z\right] = \sum_{a\in A} \mathbb{E}[(D_a F)^2|Z].$$

Hence,

$$\mathrm{var}[F|Z] \leq \mathbb{E}[\Gamma(F,F)|Z],$$

proving the theorem. □

We find a version of the McDiarmid's inequality for conditionally independent random variables.

Theorem 3.5 (Conditional McDiarmid's inequality)**.** *Let F be a square-integrable functional such that for all $a \in A$:*

$$\sup_{\substack{x^{\{a\}} \in E_{A\setminus\{a\}} \\ x'_a \in E_a}} |F(x^{\{a\}}, x'_a) - F(x)| \leq d_a.$$

For any $x > 0$, we have the inequality:

$$\mathbb{P}(F(X) - \mathbb{E}\left[F(X) \mid Z\right] \geq x|Z) \leq \exp\left(-\frac{x^2}{2\sum_{a\in A} d_a^2}\right). \tag{24}$$

Our strategy of proof is different from the original McDiarmid's original proof in [27].

Proof of theorem 3.5. We assume that $F = F(X)$ is a bounded random variable verifying $\mathbb{E}\left[F|Z\right] = 0$. Using the inequality:

$$|e^{tx} - e^{ty}| \leq \frac{t}{2}|x-y|(e^{tx} + e^{ty}) \quad \forall\, x, y \in \mathbb{R}. \tag{25}$$

We have:

$$\begin{aligned}|\Delta^{\{a\}} e^{tF}(X, X'_a)| &= |e^{tF} - e^{tF^{\{a\}'}}| \\ &\leq \frac{t}{2}|\Delta^{\{a\}} F(X, X'_a)| \left(e^{tF} + e^{tF^{\{a\}'}}\right).\end{aligned}$$

Applying the covariance identity, it yields:

$$\begin{aligned}\mathbb{E}[Fe^{tF}|Z] &= \int_0^\infty e^{-u}\sum_{a\in A}\mathbb{E}[D_a e^{tF}\Delta^{\{a\}}F(X_u^\circ,X_a')|Z]\,\mathrm{d}u\\ &\le \int_0^\infty e^{-u}\sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[|\Delta^{\{a\}}e^{tF}(X,X_a')||X,Z\right]\Delta^{\{a\}}F(X_u^\circ,X_a')|Z\right]\mathrm{d}u\\ &\le \frac{t}{2}\int_0^\infty e^{-u}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}}F(X,X_a')|e^{tF}|\Delta^{\{a\}}F(X_u^\circ,X_a')|Z\right]\mathrm{d}u\\ &+ \frac{t}{2}\int_0^\infty e^{-u}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}}F(X,X_a')|e^{tF^{\{a\}'}}|\Delta^{\{a\}}(X_u^\circ,X_a')|\,\middle|\,Z\right]\mathrm{d}u\end{aligned}$$

by using the Jensen's inequality for conditional expectation in the second inequality. Since $|\Delta^{\{a\}}F(X,X_a')|^2 \le d_a$, $|\Delta^{\{a\}}F(X_u^\circ,X_a')| \le d_a$ for all $u\in\mathbb{R}^+$ and $\mathbb{E}[e^{tF^{\{a\}'}}|Z] = \mathbb{E}[e^{tF}|Z]$, this shows that:

$$\mathbb{E}[Fe^{tF}|Z] \le \left(\sum_{a\in A} d_a^2\right) t\mathbb{E}[e^{tF}|Z] = tK^2\mathbb{E}[e^{tF}|Z],$$

where $K^2 := \sum_{a\in A} d_a^2$. Thus, in all generality for F bounded:

$$\begin{aligned}\log\mathbb{E}[e^{t(F-\mathbb{E}[F])}|Z] &= \int_0^t \frac{\mathbb{E}[(F-\mathbb{E}[F|Z])e^{s(F-\mathbb{E}[F|Z])}|Z]}{\mathbb{E}[e^{s(F-\mathbb{E}[F])}]}\,\mathrm{d}s\\ &\le K^2\int_0^t s\,\mathrm{d}s = \frac{t^2}{2}K^2,\end{aligned}$$

hence:

$$\begin{aligned}e^{tx}\mathbb{P}(F-\mathbb{E}[F|Z]>x|Z) &\le \mathbb{E}[e^{t(F-\mathbb{E}[F|Z])}|Z]\\ &= e^{t^2K^2/2},\quad t\ge 0,\end{aligned}$$

and:

$$\mathbb{P}(F-\mathbb{E}[F|Z]\ge x|Z) \le e^{\frac{t^2}{2}K^2-tx},\quad t\ge 0.$$

The minimum of the right-hand side is obtained for $t = x/K^2$. If F is not bounded, the conclusion holds for $F_n = \max(-n,\min(F,n))$, $n\ge 0$, and $(F_n)_{n\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. to F. Hence:

$$\mathbb{P}(F-\mathbb{E}[F|Z]\ge x|Z) \le \exp\left(-\frac{x^2}{2K^2}\right) = \exp\left(-\frac{x^2}{2\sum_{a\in A} d_a^2}\right).$$

The proof is thus complete. □

4. Applications to normal approximation

The goal is to bound for instance the 1-Wasserstein distance

$$d_W(\mathcal{L}(F(X)),\mathcal{L}(Y))| := \sup_{h\in\mathcal{H}} |\mathbb{E}[h(F(X))] - \mathbb{E}[h(Y)]|$$

for $\mathcal{H}$ the set of 1-Lipschitz functions and Y the random variable following the target distribution. We recall the lemma 4.2 of [4] which provides with a standard implementation of the Stein's method for this probabilistic distance with respect to the normal distribution $\mathcal{N}(0,1)$.

Lemma 4.1 (Normal approximation)**.** *Let $L^\dagger h(x) := h'(x) - xh(x)$. Then,*

$$d_W(\mathcal{L}(F(X)),\mathcal{N}(0,1)) \leq \sup_{\varphi\in\mathcal{H}_*} \left|\mathbb{E}[L^\dagger\varphi(F(X))]\right|, \tag{26}$$

where $\mathcal{H}_* := \{h \in C^2(\mathbb{R},\mathbb{R}) :\ \|h'\|_\infty \leq \sqrt{\frac{2}{\pi}}, \|h''\|_\infty \leq 2\}$.

In the following, we denote $d_W(\mathcal{L}(F(X)),\mathcal{N}(0,1))$ by $d_W(F,\mathcal{N}(0,1))$. For sake of conciseness, we denote by $\Delta^{\{a\}'}F$ the quantity $\Delta^{\{a\}}F(X,X'_a)$.

4.1. **Rates in Lyapunov's conditional central limit.**

Lemma 4.2. *For any $F \in \mathcal{S}$ such that $\mathbb{E}[F \mid Z] = 0$. Then,*

$$d_W(F,\mathcal{N}(0,1)) \leq \sup_{\psi\in\mathcal{H}_*} \left|\mathbb{E}\left[\sum_{a\in A}\psi(F(X^{\{a\}},X'_a))\Delta^{\{a\}'}FD_a(-\mathsf{L}^{-1}F) - \psi(F)\right]\right| + \sum_{a\in A}\mathbb{E}[(\Delta^{\{a\}'}F)^2|D_a\mathsf{L}^{-1}F|]. \tag{27}$$

Proof of lemma 4.2. We compute:

$$\sup_{f^\dagger\in\mathcal{H}_*} |\mathbb{E}[F(f^\dagger)(F) - (f^\dagger)'(F)]|.$$

Since F is centered,

$$\begin{aligned}
\mathbb{E}[F(f^\dagger)(F)] &= \mathbb{E}[\mathsf{L}(\mathsf{L}^{-1}F)f^\dagger(F)] \\
&= -\sum_{a\in A}\mathbb{E}[D_a\mathsf{L}^{-1}FD_af^\dagger(F)] \text{ by integration by parts} \\
&= -\sum_{a\in A}\mathbb{E}\left[D_a\mathsf{L}^{-1}F\mathbb{E}\left[(f^\dagger)'(F) - f^\dagger(F^{\{a\}'})\,\middle|\, X,Z\right]\right] \\
&= -\sum_{a\in A}\mathbb{E}[D_a\mathsf{L}^{-1}F\Delta^{\{a\}'}f^\dagger(F)].
\end{aligned}$$

Then, we use the Taylor expansion taking the reference point to be $F^{\{a\}'}$ instead of F, for all $a \in A$ yielding:

$$\begin{aligned}
\Delta^{\{a\}'}f^\dagger(F) &= f^\dagger(F) - f^\dagger(F^{\{a\}'}) \\
&= (f^\dagger)'(F^{\{a\}'})\Delta^{\{a\}'}F + R_a,
\end{aligned}$$

with $|R_a| \leq \frac{\|(f^\dagger)''\|_\infty}{2}(\Delta^{\{a\}'}F)^2 = (\Delta^{\{a\}'}F)^2$. Then,

$$\begin{aligned}
&|\mathbb{E}[Ff^\dagger(F) - (f^\dagger)'(F)]| \\
&\qquad\leq \left|\mathbb{E}\left[\sum_{a\in A}\Delta^{\{a\}'}F(D_a(-\mathsf{L}^{-1}F))\left((f^\dagger)'(F^{\{a\}'}) - (f^\dagger)'(F)\right)\right]\right| \\
&\qquad + \sum_{a\in A}\mathbb{E}[(\Delta^{\{a\}'}F)^2|D_a\mathsf{L}^{-1}F|].
\end{aligned}$$

Because $(f^\dagger)''$ has Lipschitz-constant equal to 2, we get the result. □

We prove a quantitative Lyapunov's conditional central limit theorem for random variables with moments of order 3.

Corollary 4.3 (Lyapunov's conditional central limit theorem)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of thrice integrable, conditionally independent random variables given a latent random variable Z. Let us observe that*

$$\sigma^2_{j,Z} = \mathrm{var}(X_j|Z),\ s^2_{n,Z} = \sum_{j=1}^n \sigma^2_{j,Z} \text{ and } \bar{X}_n = \frac{1}{s_{n,Z}} \sum_{j=1}^n (X_j - \mathbb{E}\left[X_j \mid Z\right]).$$

Then,

$$d_W(\bar{X}_n, \mathcal{N}(0,1)) \leq 2(\sqrt{2}+1)\mathbb{E}\left[\frac{1}{s^3_{n,Z}} \sum_{i=1}^n |X_i - \mathbb{E}\left[X_i \mid Z\right]|^3\right]. \tag{28}$$

The proof of the corollary follows the same steps as the one of [9, Corollary 5.11], using lemma 4.2.

Example 4.1 (Conditional Bernoulli random variables)**.** Let $(U_i)_{i\in\mathbb{N}}$ independent uniform random variables, and $X_i = \mathbb{1}_{\{U_i \leq Z\}}$, with Z an arbitrary random variable lying in $[0,1]$, then $(X_i)_{i\in\mathbb{N}}$ forms a sequence of conditionally independent random variables given Z. The law of $\mathcal{L}(X_i|X^{\{i\}}, Z)$ is a Bernoulli law of parameter Z. We compute the right-hand side of the Lyapunov theorem in this case.

$$s^2_{n,Z} = nZ(1-Z)$$
$$\mathbb{E}\left[\,|X_i - \mathbb{E}\left[X_i \mid Z\right]|^3 \,\middle|\, Z\right] = Z(1-Z)(1-2Z).$$

Hence,

$$d_W(\bar{X}_n, \mathcal{N}(0,1)) \leq 2(\sqrt{2}+1)\mathbb{E}\left[\frac{1-2Z+2Z^2}{\sqrt{Z(1-Z)}}\right] n^{-1/2}.$$

4.2. **Abstract bounds for U-statistics.** The chaos decomposition has a natural interpretation as a decomposition in terms of degenerate U-statistics.

Definition 4.2 (U-statistic [17])**.** Let a family of measurable functions $h_I : E_I \to \mathbb{R}$. A U-statistic of degree (or order) p is defined for any $n \geq p$ by:

$$U = \sum_{I\in(A,p)} h_I(X_I) = \sum_{I\in(A,p)} W_I.$$

Definition 4.3 (Degenerate U-statistic)**.** A degenerate U-statistic of order $p > 1$ is a U-statistic of order p such that $\mathbb{E}\left[h_I(X_I^{\{a\}}, x_a) \,\middle|\, Z\right] = 0$, for all $a \in A$ and $x_a \in E_a$.

The space of degenerate U-statistics is exactly $\mathfrak{C}_p$. Since we consider functionals given Z hereafter, h_I may be $\sigma(Z)$-measurable as well.

A convenient assumption in the proofs of quantitative limit theorems is the diffusiveness of the Markov generator at hand L, i.e. the associated carré du champ Γ_L satisfies for (F,G) in a dense algebra of $\mathrm{Dom}\, L$:

$$\Gamma_L(\phi(F), G) = \phi'(F)\Gamma_L(F,G).$$

Due to the discreteness of the Malliavin structure, the operator L is not diffusive, but it is close to. We devise the following pseudo chain rule.

Lemma 4.4 (First pseudo chain rule)**.** *Let $\psi \in C^1(\mathbb{R},\mathbb{R})$. Let $G \in \mathcal{A}$ and $F \in L^2(E_A)$ such that $\psi(F) \in \mathcal{A}$, then:*

$$\Gamma(\psi(F), G) = \frac{1}{2}\sum_{a\in A} \psi'(F)\mathbb{E}\left[(\Delta^{\{a\}'}F)(\Delta^{\{a\}'}G) \,\middle|\, X, Z\right] + R_\psi(F,G), \tag{29}$$

where:

$$|R_\psi(F,G)| \leq \frac{\|\psi''\|_\infty}{4} \sum_{a\in A} \mathbb{E}\left[|\Delta^{\{a\}'}G|(\Delta^{\{a\}'}F)^2 \,\middle|\, X, Z\right].$$

Proof of lemma 4.4. We write the Taylor expansion of ψ, and:

$$\mathbb{E}\left[(\psi(F^{\{a\}'})-\psi(F))(G^{\{a\}'}-G)\,\middle|\,X,Z\right]=\mathbb{E}\left[\psi'(F)(\Delta^{\{a\}'}F)(\Delta^{\{a\}'}G)\,\middle|\,X,Z\right]+\mathbb{E}\left[(G^{\{a\}'}-G)r_\psi(F,F^{\{a\}'}-F)\,\middle|\,X,Z\right].$$

Then,

$$2\Gamma(\psi(F),G)=\psi'(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)(\Delta^{\{a\}'}G)\,\middle|\,X,Z\right]+\sum_{a\in A}\mathbb{E}\left[(G^{\{a\}'}-G)r_\psi(F,F^{\{a\}'}-F)\,\middle|\,X,Z\right]$$

where:

$$r_\psi(x,y)=\psi(x+y)-\psi(x)-\psi'(x)y=\int_0^y(y-s)\psi''(x+s)\,\mathrm{d}s.$$

We note that r_ψ satisfies:

$$|r_\psi(x,y)|\leq\frac{\|\psi''\|_\infty}{2}y^2,$$

and we obtain the bound on the remainder. □

Theorem 4.5 (Bounds in 1-Wasserstein distance)**.** *Assume that $F\in L^3(E_A)$, such that $\mathbb{E}[F\mid Z]=0$ and $\mathbb{E}[F^2]=1$, then we get the bound:*

$$d_W(F,\mathcal{N}(0,1))\leq\sqrt{\frac{2}{\pi}}\mathbb{E}|\Gamma(F,-\mathsf{L}^{-1}F)-1|+\frac{1}{2}\sum_{a\in A}\mathbb{E}[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|(\Delta^{\{a\}'}F)^2]. \quad (30)$$

Moreover, if $F\in L^4(E_A)$, then one has the further bound:

$$d_W(F,\mathcal{N}(0,1))\leq\sqrt{\frac{2}{\pi}}\sqrt{\mathrm{var}(\Gamma(F,\mathsf{L}^{-1}F))}+\frac{\sqrt{2}}{2}\sqrt{-\mathbb{E}[F\mathsf{L}F]}\sqrt{\sum_{a\in A}\mathbb{E}[|\Delta^{\{a\}'}F|^4]}. \quad (31)$$

Proof of theorem 4.5. We have:

$$\begin{aligned}\mathbb{E}[L^\dagger f^\dagger(F)]&=\mathbb{E}[F(f^\dagger)'(F)-(f^\dagger)''(F)]\\&=\mathbb{E}[\mathsf{L}\mathsf{L}^{-1}F(f^\dagger)'(F)]-\mathbb{E}[(f^\dagger)''(F)]\\&=\mathbb{E}[\mathsf{L}^{-1}F\mathsf{L}((f^\dagger)'(F))]-\mathbb{E}[(f^\dagger)''(F)]\\&=\mathbb{E}[\Gamma(\mathsf{L}^{-1}((f^\dagger)'(F)),-\mathsf{L}^{-1}F)]-\mathbb{E}[(f^\dagger)''(F)]\end{aligned} \quad (32)$$

by integration by parts. We use lemma 4.4 and obtain that:

$$\mathbb{E}[\Gamma(\mathsf{L}((f^\dagger)'(F)),-\mathsf{L}^{-1}F)]\leq\mathbb{E}[(f^\dagger)''(F)\Gamma(F,-\mathsf{L}^{-1}F)]+\mathbb{E}[R_{(f^\dagger)^{(3)}}(F,-\mathsf{L}^{-1}F)].$$

Thus,

$$\mathbb{E}[L^\dagger f^\dagger(F)]\leq\sqrt{\frac{2}{\pi}}\mathbb{E}|\Gamma(F,-\mathsf{L}^{-1}F)-1|+\frac{1}{2}\sum_{a\in A}\mathbb{E}[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|(\Delta^{\{a\}'}F)^2].$$

By Jensen's inequality for the first term and Cauchy-Schwarz inequality (for expectation of sum of random variables) for the second one, then by integration by parts, it yields:

$$\mathbb{E}[L^\dagger f^\dagger(F)] \leq \sqrt{\frac{2}{\pi}}\sqrt{\mathrm{var}(\Gamma(F,\mathsf{L}^{-1}F))} + \frac{1}{2}\sqrt{\sum_{a\in A}\mathbb{E}[|\Delta^{\{a\}'}\mathsf{L}^{-1}F|^2]}\sqrt{\sum_{a\in A}\mathbb{E}[(\Delta^{\{a\}'}F)^4]},$$

and the proof is complete. □

Corollary 4.6. *If $F=\sum_{p=1}^m F_p$ is four times integrable functional where $F_p\in \ker(\mathsf{L}+p\mathrm{Id})$, then:*

$$d_W(F,\mathcal{N}(0,1)) \leq \sqrt{\frac{2}{\pi}}\sum_{p,q=1}^m \frac{1}{q}\sqrt{\mathrm{var}\left[\Gamma(F_p,F_q)\right]} + \sqrt{2}\sum_{p=1}^m\frac{1}{p}\sqrt{\mathbb{E}[F_p^2]}\left\{\sum_{p=1}^m p^{1/4}\left(\sum_{a\in A}\mathbb{E}\left|\Delta^{\{a\}'}F\right|^4\right)^{1/4}\right\}^2. \quad (33)$$

Proof of corollary 4.6. We use the decomposition of L^{-1} as to develop the first and second terms in (31). The final result is obtained after using Cauchy-Schwarz inequality. □

That is the starting point towards a partial fourth moment limit theorem.

4.3. **Fourth moment phenomenon.** We adapt the proof of [2], requiring a second pseudo chain rule that expresses the carré du champ operator as an approximation of a derivation operator in its two arguments.

Lemma 4.7 (Second pseudo chain rule)**.** *Let φ,ψ be twice differentiable functions such that their second derivative is bounded Lipschitz-continuous. Assume that F a four times integrable functional such that $\varphi(F)\in\mathcal{A}$, $F\in\mathcal{A}$ and $\mathbb{E}\left[F\mid Z\right]=0$, then one has:*

$$\Gamma(\varphi(F),\psi(F)) = (\varphi'\psi')(F)\Gamma(F,F) - \frac{1}{4}(\varphi''\psi'+\varphi'\psi'')(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3\,\middle|\,X,Z\right] + \sum_{a\in A}R_a, \quad (34)$$

with:

$$R_a = \frac{1}{2}\left(\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F)\,\middle|\,X,Z\right] - \varphi(F)\mathbb{E}\left[R^{(4)}_{a,\psi}(F)\,\middle|\,X,Z\right] - \psi(F)\mathbb{E}\left[R^{(4)}_{a,\varphi}(F)\,\middle|\,X,Z\right]\right)$$

and:

$$R^{(4)}_{a,\psi} \leq \frac{\|\psi^{(4)}\|_\infty}{24}\mathbb{E}\left[(\Delta^{\{a\}'}F)^4\,\middle|\,X,Z\right] \text{ for any } \psi \text{ fourth times differentiable.}$$

Proof of lemma 4.7. We have:

$$\begin{aligned}2\Gamma(\varphi(F),\psi(F)) &= 2\varphi'(F)\psi'(F)\Gamma(F,F) - \frac{3}{6}(\varphi''\psi'+\varphi'\psi'')(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3\,\middle|\,X,Z\right]\\ &+ \sum_{a\in A}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - \varphi(F)R^{(4)}_{a,\psi}(F) - \psi(F)R^{(4)}_{a,\varphi}(F)\,\middle|\,X,Z\right],\end{aligned} \quad (35)$$

with:

$$R^{(4)}_{a,\phi} = \frac{1}{6}\mathbb{E}\left[\int_F^{F^{\{a\}'}} \phi^{(4)}(x)(x-F)^4 \,\mathrm{d}x \,\middle|\, X, Z\right],$$

for ϕ a four times differentiable function. □

We focus on functionals in the p-th chaos for $p > 0$, as to obtain such kind of bound:

$$\mathrm{var}[\Gamma(F,F)] \leq C(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) + \text{remainder}.$$

Lemma 4.8. *Let $G \in \oplus_{k=0}^q \mathfrak{C}_k$. Then for any $\eta \geq q$,*

$$\mathbb{E}[G(\mathsf{L}+\eta Id)^2 G] \leq \eta\mathbb{E}[G(\mathsf{L}+\eta Id)G] \leq c\mathbb{E}[G(\mathsf{L}+\eta Id)^2 G], \tag{36}$$

where

$$c = \frac{1}{\eta - q} \wedge 1.$$

Proof of lemma 4.8. Since $G \in \oplus_{k=0}^q \mathfrak{C}_k$, we write

$$G = \sum_{k=0}^q \pi_k(G) \text{ and } \mathsf{L}G = -\sum_{k=0}^q k\pi_k(G). \tag{37}$$

It follows that

$$\begin{aligned}\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})^2 G] &= \mathbb{E}[G\mathsf{L}(\mathsf{L}+\eta\mathrm{Id})G] + \eta\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})G] \\ &= \mathbb{E}[G\sum_{k=0}^q k(k-\eta)\pi_k(G)] + \eta\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})G].\end{aligned}$$

By orthogonality of the chaos,

$$\mathbb{E}[G\sum_{k=0}^q k(k-\eta)\pi_k(G)] = -\mathbb{E}[\sum_{k=0}^q k(\eta-k)\pi_k(G)^2] \leq 0,$$

and the inequality holds in view on the assumption on η. In the same vein,

$$\begin{aligned}\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})G] &= \sum_{k=0}^q (\eta-k)\mathbb{E}[\pi_k(G)^2] \\ &\leq c\sum_{k=0}^q (\eta-k)^2\mathbb{E}[\pi_k(G)^2] \\ &= c\mathbb{E}[G(\mathsf{L}+\eta\mathrm{Id})^2].\end{aligned}$$

Thus, it yields the result. □

Lemma 4.9. *For $F \in \mathfrak{C}_p \cap L^4(E_A)$ and Q a polynomial of degree two and $a > 0$,*

$$\mathbb{E}[Q(F)(\mathsf{L}+ap\mathrm{Id})Q(F)] = p\mathbb{E}\left[aQ^2(F) - \frac{Q'(F)F}{3Q''(F)}\right] - \mathbb{E}[R_Q(F)], \tag{38}$$

where R_Q is a remainder term that depends on Q. For $Q = H_2 = X^2 - 1$ the second Hermite polynomial, the remainder reads off:

$$\mathbb{E}[R_Q] = \mathbb{E}[R_{H_2}] = \frac{1}{6}\mathbb{E}\left[\sum_{a\in A} |\Delta^{\{a\}'}F|^4\right]. \tag{39}$$

Proof of lemma 4.9. We first integrate by parts, then use the pseudo chain rule of lemma 4.7:

$$\begin{aligned}\mathbb{E}[Q(F)\mathsf{L}Q(F)] &= -\mathbb{E}[\Gamma(Q(F),Q(F))]\\ &= -\mathbb{E}[Q'(F)^2\Gamma(F,F)]\\ &+ \frac{1}{6}(Q^2)^{(3)}(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3\;\middle|\;X,Z\right]\\ &- \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[R^{(4)}_{a,Q^2}(F)\;\middle|\;X,Z\right] - 2Q(F)\mathbb{E}\left[R^{(4)}_{a,Q}(F)\;\middle|\;X,Z\right]\right].\end{aligned} \tag{40}$$

Since $Q^{(3)}=0$, we have:

$$\begin{aligned}\mathbb{E}[Q(F)\mathsf{L}Q(F)] &= -\mathbb{E}\left[[Q'(F)^2\Gamma(F,F)\right]\\ &+ \frac{1}{6}\mathbb{E}\left[(Q^2)^{(3)}(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3\;\middle|\;X,Z\right]\right]\\ &- \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[R^{(4)}_{a,Q^2}(F)\;\middle|\;X,Z\right]\right].\end{aligned} \tag{41}$$

Moreover,

$$\left(\frac{Q'(F)^3}{3Q''(F)}\right)' = \frac{3Q'(F)Q''(F)^2}{3Q''(F)^2} = Q'(F)^2. \tag{42}$$

Subsequently, we use the pseudo chain rule of lemma 4.7 taking $\psi=\mathrm{Id}$ and $\varphi=\frac{Q'(\cdot)^3}{3Q''(\cdot)}$:

$$\begin{aligned}\mathbb{E}[Q'(F)^2\Gamma(F,F)] &= \mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)},F\right)\right]\\ &+ \frac{1}{4}\mathbb{E}\left[(\varphi''\psi'+\varphi'\psi'')(F)\sum_{a\in A}\mathbb{E}\left[(\Delta^{\{a\}'}F)^3\;\middle|\;X,Z\right]\right]\\ &- \sum_{a\in A}\mathbb{E}\left[\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F)\;\middle|\;X,Z\right] - \varphi(F)\mathbb{E}\left[R^{(4)}_{a,\psi}(F)\;\middle|\;X,Z\right]\right]\\ &- \mathbb{E}\left[F\mathbb{E}\left[R^{(4)}_{a,\varphi}(F)\;\middle|\;X,Z\right]\right]\\ &= \mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)},F\right)\right] + \frac{1}{4}\mathbb{E}\left[(Q'(\cdot)^2)'(F)\sum_{a\in A}(\Delta^{\{a\}'}F)^3\right]\\ &- \sum_{a\in A}\frac{1}{2}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - FR^{(4)}_{a,\varphi}(F)\right].\end{aligned} \tag{43}$$

Finally,

$$\begin{aligned}\mathbb{E}[Q(F)\mathsf{L}Q(F)] &= -\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] \\ &+ \mathbb{E}\left[\left(\frac{1}{4}(Q'(\cdot)^2)'(F) - \frac{1}{12}(Q^2)^{(3)}(F)\right)\sum_{a\in A}(\Delta^{\{a\}'}F)^3\right] \\ &+ \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - R^{(4)}_{a,Q^2}(F) - FR^{(4)}_{a,\varphi}(F)\right] \\ &= -\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] \\ &+ \mathbb{E}\left[\left(\frac{1}{4}(Q'(\cdot)^2)'(F) - \frac{1}{12}(Q^2)^{(3)}(F)\right)\sum_{a\in A}(\Delta^{\{a\}'}F)^3\right] \\ &+ \frac{1}{2}\sum_{a\in A}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - R^{(4)}_{a,Q^2}(F)\right].\end{aligned} \tag{44}$$

Because $F \in \mathfrak{C}_p$, we have: $-\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] = \mathbb{E}[\frac{Q'(F)^3}{3Q''(F)}\mathsf{L}F] = -p\mathbb{E}\left[\frac{Q'(F)^3}{3Q''(F)}F\right]$. For $Q = H_2 = X^2 - 1$ the second Hermite polynomial,

$$\frac{Q'(F)^3}{3Q''(F)} = \frac{4}{3}X^3,$$

so $\left(\frac{Q'(\cdot)^3}{3Q''(\cdot)}\cdot\right)^{(4)} = 32$ and $(Q^2)^{(4)} = 24$. Thus,

$$\sum_{a\in A}\mathbb{E}\left[R^{(4)}_{a,\varphi\psi}(F) - R^{(4)}_{a,Q^2}(F)\right] = \frac{(32-24)}{24}\sum_{a\in A}\mathbb{E}\left[|\Delta^{\{a\}'}F|^4\right]. \tag{45}$$

Since $(Q'(\cdot)^2)'(F) = 8F$, and $(Q^2)^{(3)}(F) = 24F$, the result follows. □

The assumption under which a fourth moment theorem holds, is that $F \in \mathfrak{C}_p$ is a chaos eigenfunction with respect to the Markov generator L i.e.:

$$F^2 \in \oplus_{k=0}^{2p}\mathfrak{C}_k. \tag{EGF}$$

It is analog to the one in [25, 2]. We show that it holds for an important class of U-statistics, homogeneous sums. We shall use the notation (A, p) that stands for the set of p-tuples of distinct elements of A.

Example 4.4 (**Conditionally independent homogeneous sums**)**.** Let $p > 0$. If there exists $(a_I)_{I\subset A} \in \mathbb{R}^{\mathcal{P}(A)}$ such that

$$W = \sum_{k=1}^{p}\sum_{I\in(A,k)} a_I \prod_{i\in I} X_i, \tag{46}$$

then

(1) W is square-integrable homogeneous sum of order p if X_i are $2p$-integrable. In that case, $W \in \mathcal{S}$.

(2)
$$\mathbb{E}\left[W \mid Z\right] = \sum_{k=1}^{p}\sum_{I\in(A,k)} a_I \prod_{i\in I}\mathbb{E}\left[X_i \mid Z\right]$$
is a homogeneous sum of random variables $\hat{X}_i = \mathbb{E}\left[X_i \mid Z\right]$ for $i \in I$ with $I \in (A, k)$ for $k \le p$.

Remark that $(a_I)_{I\subset A}$ may be a sequence of random variables, in which case there exists a family of functions $(g_I)_{I\subset A}$ such that $a_I = g_I(Z)$.

Lemma 4.10. *Let W a homogeneous sums of conditionally independent random variables given Z. Then* (EGF) *holds.*

Proof of lemma 4.10. Let us denote by W_I the component of F in (46) proportional to $\prod_{\alpha\in I} X_\alpha$. We want to prove that there exist $G_1,\dots,G_{2p}$ with $G_i \in \mathfrak{C}_i \cup \{0\}$ such that $W_I W_J = \sum_{i=1}^{2p} G_i$. Note that if $I\cap J=\emptyset$, and $a\in I$, then a is not in J and vice versa. Therefore, $W_I W_J \in \mathfrak{C}_{|I|+|J|}$. In general,

$$\begin{aligned} W_I W_J &\propto \prod_{\alpha\in I} Y_\alpha \prod_{\beta\in J} Y_\beta \\ &= \prod_{\gamma\in (I\setminus J)\cup(J\setminus I)} Y_\gamma \prod_{\delta\in I\cap J} Y_\delta^2 \\ &= \prod_{\gamma\in (I\setminus J)\cup(J\setminus I)} Y_\gamma \prod_{\delta\in I\cap J} (Y_\delta^2 - \mathbb{E}\left[Y_\delta^2 \mid Z\right] + \mathbb{E}\left[Y_\delta^2 \mid Z\right]) \\ &= \sum_{K\subset I\cap J} \prod_{\gamma\in (I\setminus J)\cup(J\setminus I)} Y_\gamma \prod_{\delta\in K} (Y_\delta^2 - \mathbb{E}\left[Y_\delta^2 \mid Z\right]) \prod_{\delta\in (I\cap J)\setminus K} \mathbb{E}\left[Y_\delta^2 \mid Z\right]. \end{aligned}$$

For $a\in A$:

$$\begin{aligned} &\mathbb{E}\left[\prod_{\gamma\in (I\setminus J)\cup(J\setminus I)} Y_\gamma \prod_{\delta\in K} (Y_\delta^2 - \mathbb{E}\left[Y_\delta^2 \mid Z\right]) \prod_{\delta\in (I\cap J)\setminus K} \mathbb{E}\left[Y_\delta^2 \mid Z\right] \,\middle|\, \mathcal{G}_a^Z \right] \\ &= \begin{cases} 0 \text{ if } a\in K\cup((I\setminus J)\cup(J\setminus I)) \\ \displaystyle\prod_{\gamma\in (I\setminus J)\cup(J\setminus I)} Y_\gamma \prod_{\delta\in K} (Y_\delta^2 - \mathbb{E}\left[Y_\delta^2 \mid Z\right]) \prod_{\delta\in (I\cap J)\setminus K} \mathbb{E}\left[Y_\delta^2 \mid Z\right] \text{ otherwise.} \end{cases} \end{aligned}$$

Hence, we get

$$\prod_{\gamma\in (I\setminus J)\cup(J\setminus I)} Y_\gamma \prod_{\delta\in K} (Y_\delta^2 - \mathbb{E}\left[Y_\delta^2 \mid Z\right]) \prod_{\delta\in (I\cap J)\setminus K} \mathbb{E}\left[Y_\delta^2 \mid Z\right] \in \mathfrak{C}_{|K\cup(I\setminus J)\cup(J\setminus I)|}$$

with $|K\cup((I\setminus J)\cap(J\setminus I))| \le |I\cup J| \le 2p$. Thus, (EGF) holds. □

Proposition 4.11. *For $F\in \mathfrak{C}_p \cap L^2(E_A)$ such that $\mathbb{E}[F^2]=1$ and* (EGF) *holds, one has:*

$$\mathbb{E}[(\Gamma(F,F)-p)^2] \le \frac{p^2}{3}|\mathbb{E}[F^4]-3| + \frac{p}{12}\mathbb{E}\left[\sum_{a\in A} |\Delta^{\{a\}'}F|^4\right]. \tag{47}$$

Proof of proposition 4.11. By the very definition of Γ, one has:

$$\begin{aligned} \Gamma(F,F)-p &= \frac{1}{2}\mathsf{L}(F^2) - F\mathsf{L}F - p = \frac{1}{2}\mathsf{L}(F^2) + pF^2 - p \text{ for } F\in\mathfrak{C}_p \\ &= \frac{1}{2}(\mathsf{L}+2p\mathrm{Id})(F^2-1). \end{aligned}$$

It follows that:

$$\mathbb{E}[(\Gamma(F,F)-p)^2] = \frac{1}{4}\mathbb{E}[\big((\mathsf{L}+2p\mathrm{Id})(F^2-1)\big)^2].$$

Since L is a self-adjoint operator, this yields:

$$\mathbb{E}[(\Gamma(F,F)-p)^2] = \frac{1}{4}\mathbb{E}[H_2(F)(\mathsf{L}+2p\mathrm{Id})^2 H_2(F)].$$

As (EGF) holds, we are in position to apply lemma 4.8 with $q=2p$ and $\eta=2p$:

$$\mathbb{E}[(\Gamma(F,F)-p)^2] \le \frac{p}{2}\mathbb{E}[H_2(F)(\mathsf{L}+2p\mathrm{Id})H_2(F)]. \tag{48}$$

According to lemma 4.9, with $a = 2$,

$$\begin{aligned}\frac{p}{2}\mathbb{E}[H_2(F)(\mathsf{L}+2p\mathrm{Id})H_2(F)] &= \frac{p^2}{2}\mathbb{E}\left[2(F^2-1)^2 - \frac{4}{3}F^4\right] + \frac{p}{2}\mathbb{E}[R_{H_2}(F)] \\ &= \frac{p^2}{6}\mathbb{E}\left[6(F^2-1)^2 - 4F^4\right] + \frac{p}{2}\mathbb{E}[R_{H_2}] \\ &= \frac{p^2}{3}\mathbb{E}[F^4 - 6F^2 + 3] + \frac{p}{2}\mathbb{E}[R_{H_2}].\end{aligned}$$

Thus, it yields

$$\mathbb{E}[(\Gamma(F,F)-p)^2] \leq \frac{p^2}{3}|\mathbb{E}[F^4 - 6F^2 + 3]| + \frac{p}{2}|\mathbb{E}[R_{H_2}]|, \tag{49}$$

and the proof is complete, using again lemma 4.9. □

4.4. **Quantitative De Jong's theorems.** Many papers are devoted to find the optimal conditions for the asymptotic normality of U-statistics. The criterion established in [5] is related to the fourth moment phenomenon. The extra assumption is a negligibility condition also known as the Lindeberg-Feller condition. Fix A_m a finite subset of cardinal m such that $F = F(X_{A_m})$ and $\mathbb{E}[F^2] = 1$, that means:

$$\rho^2_{A_m} = \max_{i\in A_m} \sum_{I\ni i,\, I\subseteq A_m,\, |I|=p} \mathbb{E}[W_I^2] \xrightarrow{m\to+\infty} 0. \tag{50}$$

In some papers [10], the term ρ_{A_m} is called maximal influence of the random variables on the total variance of the degenerate U-statistics F. In the following, we shall denote it by ρ. The condition (50) is not necessary for asymptotic normality to hold, but there exist counterexamples for which the sequence of fourth cumulants of functionals of independent Rademacher random variables converges to 0 while (50) does not hold (see [11]). We show that the quantity is related to the remainder above.

Definition 4.5 (Connectedness of subsets)**.** The r-tuple $(I_1,\dots,I_r)$ subsets of A is connected if the intersection graph of $\{I_1,\dots,I_r\}$ is connected, i.e. the graph G with vertex set $\{I_1,\dots,I_r\}$ and edge set $E(G) = \{\{I_i, I_j\} |\, i\neq j, I_i\cap I_j \neq \emptyset\}$ is connected.

Lemma 4.12. *If $F \in \mathfrak{C}_p \cap L^4(E_A)$, then:*

$$\sum_{a\in A}\mathbb{E}[|\Delta^{\{a\}'}F|^4] \leq 16p \sum_{(I,J,K,L)\text{ connected}} |\mathbb{E}[W_IW_JW_KW_L]|. \tag{51}$$

Moreover, assuming the hypercontractivity condition, i.e.

$$\sup_{J\in(A,p)} \frac{\mathbb{E}[W_J^4]}{\mathbb{E}[W_J^2]^2} < +\infty, \tag{HC}$$

there exists a constant c_p that depends only on p such that:

$$\sum_{a\in A}\mathbb{E}[|\Delta^{\{a\}'}F|^4] \leq c_p\rho^2. \tag{52}$$

Proof of lemma 4.12. Because $(a+b)^4 \leq 8(a^4+b^4)$, one has:

$$\begin{aligned}
\sum_{a\in A}\mathbb{E}\left|\Delta^{\{a\}'}F\right|^4 &\leq 8\sum_{a\in A}\mathbb{E}\left[\left(\sum_{I\ni a,|I|\leq p}W_I^{\{a\}'}\right)^4+\left(\sum_{I\ni a,|I|\leq p}W_I\right)^4\right]\\
&= 16\sum_{a\in A}\mathbb{E}\left[\left(\sum_{I\ni a,|I|\leq p}W_I\right)^4\right]\\
&\leq 16\sum_{I\cap J\cap K\cap L\neq\emptyset}|I\cap J\cap K\cap L|\mathbb{E}[W_IW_JW_KW_L]\\
&\leq 16p\sum_{I\cap J\cap K\cap L\neq\emptyset}|\mathbb{E}[W_IW_JW_KW_L]|\\
&\leq 16p\sum_{I,J,K,L\text{ connected}}|\mathbb{E}[W_IW_JW_KW_L]|.
\end{aligned}$$

Then, we bound it by the maximal influence, using the generalized Hölder inequality:

$$\begin{aligned}
|\mathbb{E}[W_IW_JW_KW_L]| &\leq \left(\mathbb{E}[W_I^4]\mathbb{E}[W_J^4]\mathbb{E}[W_K^4]\mathbb{E}[W_L^4]\right)^{1/4}\\
&\leq \max_{J\in A,|J|=p}\frac{\mathbb{E}[W_J^4]}{\mathbb{E}[W_J^2]^2}\left(\mathbb{E}[W_I^2]^2\mathbb{E}[W_J^2]^2\mathbb{E}[W_K^2]^2\mathbb{E}[W_L^2]^2\right)^{1/4}
\end{aligned}$$

with $\sigma_I^2=\mathbb{E}[W_I^2]$. Then the proposition 2.9 of [10] can be extended for functionals of conditionally independent random variables and implies that:

$$\sum_{I\cap J\cap K\cap L\neq\emptyset}\sigma_I\sigma_J\sigma_K\sigma_L\leq C_p\rho^2,$$

where the finite constant C_p only depends on p. It yields the existence of $c_p>0$ such that the inequality (52) holds true. □

We are now in position to state a partial fourth moment limit theorem.

Theorem 4.13 (Quantitative De Jong's limit theorem I)**.** *Let $F\in L^4(E_A)$ a degenerate U-statistics of order p of conditionally independent random variables such that $\mathbb{E}[F\mid Z]=0$ and $\mathbb{E}[F^2]=1$. If we suppose the hypercontractivity condition* (HC) *and the assumption* (EGF)*, then one has the bound:*

$$d_W(F,\mathcal{N}(0,1))\leq\sqrt{\frac{2}{3\pi}}\sqrt{|\mathbb{E}[F^4]-3|}+\tilde{C}_p\rho, \tag{53}$$

with $\tilde{C}_p$ a positive constant that only depends on p.

Proof. By corollary 4.6,

$$d_W(F,\mathcal{N}(0,1))\leq\sqrt{\frac{2}{\pi}}\frac{1}{p}\sqrt{\mathrm{var}\left[\Gamma(F,F)\right]}+\sqrt{2}\sqrt{\mathbb{E}[F^2]}\left(\sum_{a\in A}\mathbb{E}\left[\left|\Delta^{\{a\}'}F\right|^4\right]\right)^{1/2}.$$

The combination of (47) and lemma 4.12 yields the final upper bound. □

The upper bound of the remainder expressed in terms of maximal influence is not used in the subsequent applications, so we drop the (HC) condition.

A related result to the fourth moment phenomenon appears in [6]. We prove the associated quantitative statement for functionals of conditionally independent random variables. We prepare the proof with the following proposition.

Proposition 4.14. *If $F = \sum_{p=1}^{m} F_p$ where $F_p = \sum_{|I|=p} W_I \in \mathfrak{C}_p$, assuming there exists $C \in \mathbb{R}^+$ such that for all $I, J \subset A$, and $a \in A$, that*

$$\frac{\mathbb{E}\left[W_I W_J \mid \mathcal{G}^a\right]}{W_{I\setminus\{a\}} W_{J\setminus\{a\}}} < C \ \mathbb{P}\text{-a.s.}, \tag{H1}$$

then for $p \neq q$:

$$\sqrt{\mathrm{var}\left[\Gamma(F_p, F_q)\right]} \lesssim \sqrt{\sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|}, \tag{54}$$

for I, J, K, L sets of size less than $\max(p, q)$.

Proof of proposition 4.14. The carré du champ reads for $p \neq q$:

$$\begin{aligned}\Gamma(F_p, F_q) &= \Gamma\Big(\sum_{|I|=p} W_I, \sum_{|J|=q} W_J\Big) \\ &= \sum_{|I|,|J|=p,q} \Gamma(W_I, W_J).\end{aligned}$$

Hence,

$$\begin{aligned}2\Gamma(F_p, F_q) &= \sum_{|I|,|J|=p,q} \left(\mathsf{L}(W_I W_J) + (p+q) W_I W_J\right) \\ &= \sum_{|I|,|J|=p,q} \left((p+q) W_I W_J - \sum_{a\in A} D_a(W_I W_J)\right) \\ &= \sum_{|I|,|J|=p,q} \left((p+q) W_I W_J - \sum_{a\in I\cup J} D_a(W_I W_J)\right) \\ &= (p+q) \sum_{\substack{|I|,|J|=p,q \\ I\cap J=\emptyset}} W_I W_J + \sum_{\substack{|I|,|J|=p,q \\ I\cap J\neq\emptyset}} (|I| + |J| - |I\cup J|) W_I W_J \\ &+ \sum_{a\in I\cup J} \mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right].\end{aligned}$$

Because of the spectral decomposition, $\mathbb{E}\left[W_I \mid \mathcal{G}_a\right] = 0$ for $a \in I$. Let J such that $a \notin J$, then $\mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right] = W_J \mathbb{E}\left[W_I \mid \mathcal{G}_a\right] = 0$.

$$2\Gamma(F_p, F_q) = (p+q) \sum_{\substack{|I|,|J|=p,q \\ I\cap J=\emptyset}} W_I W_J + \sum_{\substack{|I|,|J|=p,q \\ I\cap J\neq\emptyset}} \sum_{a\in I\cap J} \left(W_I W_J + \mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right]\right).$$

Then for $p \neq q$, using the convexity of $x \longmapsto x^2$,

$$\begin{aligned}\mathrm{var}(\Gamma(F_p, F_q)) \leq \frac{1}{2}\,\mathrm{var}&\left[(p+q) \sum_{\substack{|I|,|J|=p,q \\ I\cap J=\emptyset}} W_I W_J\right] \\ &+ \frac{1}{2}\,\mathrm{var}\left[\sum_{\substack{|I|,|J|=p,q \\ I\cap J\neq\emptyset}} \sum_{a\in I\cap J} \left(W_I W_J + \mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right]\right)\right]\end{aligned}$$

$$\operatorname{var}(\Gamma(F_p,F_q)) \leq \frac{1}{2}\mathbb{E}\left[\left((p+q)\sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}} W_I W_J\right)^2\right] + \frac{1}{2}\operatorname{var}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{a\in I\cap J}(W_I W_J + \mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right])\right]$$

$$\begin{aligned}2\operatorname{var}(\Gamma(F_p,F_q)) \leq{}& \sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L=\emptyset}}\mathbb{E}[W_I W_J W_K W_L]\\ &+\mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} W_I W_J W_K W_L\right]\\ &+\mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} W_I W_J \mathbb{E}\left[W_K W_L \mid \mathcal{G}_b\right]\right]\\ &+\mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} \mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right] W_K W_L\right]\\ &+\mathbb{E}\left[\sum_{\substack{|I|,|J|=p,q\\ I\cap J\neq\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L\neq\emptyset}}\sum_{a\in I\cap J}\sum_{b\in K\cap L} \mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right]\mathbb{E}\left[W_K W_L \mid \mathcal{G}_b\right]\right].\end{aligned}$$

We shall write

$$|C_{I,J,a}| = \left|\frac{\mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right]}{W_{I\setminus\{a\}} W_{J\setminus\{a\}}}\right| \text{ for all } I,J,a$$

with the convention $W_\emptyset = 1$.

Let us deal with each term one by one:

- If $I\cap J=\emptyset$, $K\cap L=\emptyset$, and if there is more than 2 other pairs with null intersection, the contribution of the term is 0, hence the first term is non-zero if (I,J,K,L) is connected, then:
$$\sum_{\substack{|I|,|J|=p,q\\ I\cap J=\emptyset}}\sum_{\substack{|K|,|L|=p,q\\ K\cap L=\emptyset}}\mathbb{E}[W_I W_J W_K W_L] \leq \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|.$$
- The second term consists of the sums of product of factors indexed by connected sets since there are at least two pairs that have non-null intersection. Since $p\neq q$, $\mathbb{E}\left[W_I W_J \mid Z\right]=0$ for $|I|=p$ and $|J|=q$, so if the terms are non-zero, $W_I W_J$ and $W_K W_L$ are not conditionally independent.
- For the third term, using self-adjointness, the terms are non-zero if $b\in I\cap J$, hence it is equivalent to:
$$|C_{I,J,a}\mathbb{E}[W_{I\setminus\{b\}} W_{J\setminus\{a\}} W_K W_L]| = |C_{I,J,a}||\mathbb{E}[W_{I\setminus\{b\}} W_{J\setminus\{a\}} W_K W_L]|.$$

If b is the unique element that lies in the intersection, the contribution is 0, otherwise I, J, K, L are connected or the contribution is

$$\mathbb{E}\left[W_I W_J \mid Z\right]\mathbb{E}\left[W_K W_L \mid Z\right] = 0$$

because $|I| \neq |J|$.

- For the last term, it is the same argument.

Then, there exists a constant C independent of others such that

$$\mathrm{var}(\Gamma(F_p, F_q)) \leq (1 + m^2 + 2Cm^2 + C^2m^2) \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|.$$

□

In [31], Privault and Serafin proves a partial fourth moment theorem for F a functional of independent random variables sum of element in the first and second chaos of their own Malliavin structure. To that end, we devise another strategy which is to reexpress the remainder in the partial fourth moment theorem as a fourth order term.

Theorem 4.15 (Quantitative De Jong's theorem II)**.** *If $F = \sum_{p=1}^m F_p$ where $F_p \in \mathfrak{C}_p$ and let us assume:*

- *F_p are chaos eigenfunctions* (EGF)*;*
- *the condition* (H1)*;*
- $$\kappa = \sup_{I,J \subset A} \frac{\mathbb{E}[W_I^2]\mathbb{E}[W_J^2]}{\mathbb{E}[W_I^2 W_J^2]} < \infty \tag{H2}$$
  is independent of A.

Then:

$$d_W(F, \mathcal{N}(0,1)) \leq C_m \sqrt{\sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|}, \tag{55}$$

where the constant C_m grows quadratically with m, independent of all others.

Proof of theorem 4.15. Let us prove the upper bound of $\mathrm{var}\left[\Gamma(F_p, F_p)\right]$ by bounding the fourth cumulant:

$$\begin{aligned}
\mathbb{E}[F_p^4] &= 3 \sum_{\substack{I,J,K,L \in (A,p) \\ (I \cup J) \cap (K \cup L) = \emptyset}} \mathbb{E}[W_I W_J]\mathbb{E}[W_K W_L] + \sum_{\substack{I,J,K,L \in (A,p) \\ I,J,K,L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] \\
&= 3 \sum_{I,J \in (A,p)} \mathbb{E}[W_I^2]\mathbb{E}[W_J^2] - 3 \sum_{I \cap J \neq \emptyset \neq} \mathbb{E}[W_I^2]\mathbb{E}[W_J^2] \\
&+ \sum_{\substack{I,J,K,L \in (A,p) \\ I,J,K,L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] \\
&= 3\mathbb{E}[F_p^2]^2 + \sum_{\substack{I,J,K,L \in (A,p) \\ I,J,K,L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] - 3 \sum_{\substack{I \cap J \neq \emptyset \\ I \neq J}} \mathbb{E}[W_I^2]\mathbb{E}[W_J^2].
\end{aligned}$$

Then, one has:

$$|\mathbb{E}[F_p^4] - 3|\mathbb{E}[F_p^2]^2| \leq (1 + 3\kappa) \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|. \tag{56}$$

□

The assumptions may seem cumbersome, but as shown in lemma 4.10 concerning (EGF), they are valid for homogeneous sums.

5. Application to motif estimation

We interest in the applications of the bounds of probability distances to asymptotic normality of subhypergraph counts in exchangeable random hypergraphs.

5.1. **Basic hypergraph definitions.** The hypergraph model is a generalization of graph notion that aims at model more complex model in network analysis.

Definition 5.1. A hypergraph denoted by $G = (V, E = (e_i)_{i\in\mathcal{P}(V)})$ on a finite set $V = V(G)$ is a family of subsets of V called hyperedges. Vertices in a hypergraph are adjacent if there is a hyperedge which contains them. The vertices not in any edge are the isolated vertices of G. A hypergraph is connected if it contains no isolated vertices and if the intersection graph of E is connected.

We denote by $[e]$ the set of vertices of the hyperedge e.

Definition 5.2. A k-uniform hypergraph $G = (V, E)$ is a hypergraph where each hyperedge has cardinality k. In particular, such hypergraph has hyperedge set in $\binom{V}{k}$, the collection of k-tuples of the set of vertices V.

Definition 5.3. For $k > 3$, a subhypergraph (or simply subgraph) of a hypergraph $G = (V, E)$ is a hypergraph $H = (V', E')$ such that $V' \subset V$ and $E' \subset E \cap \binom{V}{k}$.

We denote by v_H and e_H the number of vertices and number of hyperedges of a hypergraph H respectively.

A 2-uniform hypergraph is a graph. A 3-uniform hypergraph is a hypergraph whose hyperedges are triangles only. We also denote by $G^{(j)}$ the hypergraph induced by the hypedges of cardinality $j \leq k$ included in the hyperedges of the k-uniform hypergraph G.

5.2. **Exchangeable random hypergraphs.** The random hypergraphs are natural extensions of random graphs. A vast majority of the literature deals with the Erdös-Rényi model and its generalization. It is an example of exchangeable random hypergraphs.

Definition 5.4. A k-uniform exchangeable random hypergraph $\mathbf{G}$ of vertex set $V = [n]$ is defined by the set of $\{0,1\}$-valued random variables $(X_\alpha, \alpha \subset \binom{[n]}{k})$ such that:

- one associates each realization of the random variables a hypergraph $([n], E)$ with $\alpha \in E$ if and only if $X_\alpha = 1$;
- (X_α) form an exchangeable array, i.e. $X_{(\sigma(u))_{u\in\alpha}} \overset{d}{=} X_\alpha$.

One can formulate a recipe for exchangeable random hypergraphs as done in [1, definition 2.8]. Fix a sequence of ingredients which consist of a sequence of sample spaces and probability kernels that determine the presence of k-hyperedges in the hypergraph based on the indicators X_β for $(k-1)$-hyperedges β:

$$(\{*\}), (V, P_1), (\{0,1\}, P_2), (\{0,1\}, P_3), \ldots, (\{0,1\}, P_{k-1}), (\{0,1\}, P_k)$$

where we write $\{*\}$ for a one-point space, $(P_k)_{k\in\mathbb{N}}$ is a family of probability kernels such that for all $k \in \mathbb{N}$, P_k is a probability kernel from $\prod_{j=0}^{k-1}\{0,1\}^{\binom{V}{j}}$ to $\{0,1\}^{\binom{V}{k}}$.

- Color each vertex $s \in V$ by some $x_s \in \{0,1\}$ chosen independently according to $P_1(*,\cdot)$;
- Color each edge $a = \{s,t\} \in \binom{V}{2}$ by some $x_a \in \{0,1\}$ chosen independently according to $P_2(*, x_s, x_t, \cdot)$;

$$\vdots$$

- Color each $(k-1)$-hyperedge $u \in \binom{V}{k-1}$ by some $x_u \in \{0,1\}$ chosen independently according to $P_{k-1}(*, (x_s)_{s\in\binom{[u]}{1}}, *, \dots, *, (x_v)_{v\in\binom{[u]}{k-2}}, \cdot)$;
- Color each k-hyperedge $e \in \binom{V}{k}$ by some color $x_e \in \{0,1\}$ chosen independently according to $P_k(*, (x_s)_{s\in\binom{[e]}{2}}, *, \dots, *, (x_u)_{u\in\binom{[e]}{k-1}}, \cdot)$.

Example 5.5 (Erdös-Rényi random model)**.** The randomness intervenes at the level of edges. $P_1(*,\cdot)$ is the uniform distribution on V. We color each edge $a = \{s,t\} \in \binom{V}{2}$ by some $z_a \in \{0,1\}$ chosen independently according to $P_2(*, x_s, x_t, \cdot) \overset{d}{=} \mathcal{B}(p)$ the Bernoulli distribution with parameter p for some $p \in [0,1]$ which is called the *edge density*.

Example 5.6 (Stochastic block model)**.** A stochastic block model corresponds to a model where there are communities, and each edge has a probability of belonging to the model according to the community of the vertices that the edge links. Likewise, the randomness intervenes at the level of the edges. Let a partition $V = C_1 \sqcup \ldots \sqcup C_q$. Let $(p_{i,j})_{i,j\in[\![1,q]\!]^2}$ a sequence of reals in $[0,1]$. We can assign a community to each vertex s, let call it $c(s)$. Then:

- $P_1(*,\cdot)$ is the uniform distribution;
- $P_2(*, z_s, z_t, \cdot) \overset{d}{=} \mathcal{B}(p_{c(s),c(t)})$.

The natural extension of the Erdös-Rényi model denoted $\mathbb{G}^{(3)}(n,p_n)$ consists of having

$$P_3(*, x_{st}, x_{tu}, x_{us}) \overset{d}{=} \mathcal{B}(p_n),$$

i.e. we draw every triangle of the hypergraph with probability p_n. We also consider another random model based on the recipe. Let $(\mathbb{T}^{(3)}(n,q_n,p_n))_{n\in\mathbb{N}}$ the sequence of 3-uniform hypergraphs such that for $(s,t,u) \in V^3$:

- $$P_2(*, x_s, x_t) \overset{d}{=} \mathcal{B}(q_n);$$
- $$P_3(*, x_{st}, x_{tu}, x_{us}) \overset{d}{=} \mathcal{B}(p_n).$$

It differs from $\mathbb{G}^{(3)}(n,p_n)$ in many ways as pointed out by [26, Example 23.11], but we note that $\mathbb{G}^{(3)}(n,p_n)$ and $\mathbb{T}^{(3)}(n,1,p_n)$ have the same law. The case $q_n < 1$ has not been much studied in the literature. The functional identities in Section 3 can be applied to random hypergraphs in the same way as for random graphs [19, corollary 2.27]. In that section, we consider once for all A to be the set of hyperedges. We use the notation A for other purposes.

5.3. **Motif estimation in random hypergraphs.** One of the oldest problem of motif estimation is subgraph counting in random graphs. Small subgraph counts can be used as summary statistics for large random graphs. The asymptotic normality of subgraph count in Erdös-Rényi model is well-known, as well as the convergence rate [20]. There are many extensions that revolve around the definition of a random graph as a sequence of independent random variables, for example a clique complex of Bernoulli random graphs. In this work, we study subgraph counting in 3-uniform random hypergraphs. To the best of our knowledge, this is the first paper about asymptotic normality of subgraph counting of such models.

The number of subhypergraphs of $\mathbb{G}^{(3)}(n,p_n)$ isomorphic to G is

$$M_G = \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \prod_{\alpha\in H} \hat{X}_\alpha. \tag{57}$$

For $\sigma \in \mathrm{Aut}(G)$, $(x,y,z) \in E(G)$ if and only if $(\sigma(x),\sigma(y),\sigma(z)) \in E(G)$. The random variable M_G has a finite Hoeffding decomposition [6, p.11(115)]. Since $\hat{X}_\alpha = p_n + (\hat{X}_\alpha - p_n)$, M_G admits the decomposition:

$$M_G = \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{J \subseteq H} p_n^{|H|-|J|} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n), \tag{58}$$

where the summation extends over all subsets J of I, in virtue of the inclusion-exclusion principle. By interchanging the sums, we find the chaotic decomposition of $M_G - \mathbb{E}[M_G]$ that is:

$$\begin{aligned}
M_G - \mathbb{E}[M_G] &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\substack{J \subseteq H \\ J \neq \emptyset}} p_n^{|I|-|J|} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n), \\
&= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{j=1}^{e_G} p_n^{e_G - j} \sum_{\substack{J \subset H \\ |J|=j}} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n) \\
&= \sum_{j=1}^{e_G} p_n^{e_G - j} \sum_{|J|=j} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n) \left(\sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \\
&= \sum_{j=1}^{e_G} \pi_j(M_G),
\end{aligned}$$

where

$$\pi_k(M_G) = p_n^{e_G - j} \sum_{|J|=j} \left(\sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \prod_{\alpha \in J} \hat{Y}_\alpha \tag{59}$$

with $\hat{Y}_\alpha$ is the centered version of $\hat{X}_\alpha$ for all α hyperedges of K_n. We note that the decomposition above corresponds to the Hoeffding decomposition of the U-statistics with

$$W_J \propto \left(\sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \prod_{\alpha \in J} \hat{Y}_\alpha. \tag{60}$$

We proceed in the same manner in $\mathbb{T}^{(3)}(n, q_n, p_n)$. Define N_G the number of sub-hypergraphs isomorphic to G

$$N_G = \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \prod_{\alpha \in H} X_\alpha. \tag{61}$$

Here, $(X_\alpha)_{\alpha \in \binom{[n]}{3}}$ is a sequence of conditionally independent Bernoulli random variables given $Z = \mathbb{G}(n, q_n)$. The chaos decomposition yields:

$$\begin{aligned}
N_G &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{J \subseteq H} \prod_{\beta \in H \setminus J} \mathbb{E}\left[X_\beta \mid \mathbb{G}(n, q_n) \right] \prod_{\alpha \in J} \left(X_\alpha - \mathbb{E}\left[X_\alpha \mid \mathbb{G}(n, q_n) \right] \right) \\
&= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{J \subseteq H} p_n^{|H|-|J|} \mathbb{1}_{\{(H \setminus J)^{(2)} \subset \mathbb{G}(n, q_n)\}} \prod_{\alpha \in J} \left(X_\alpha - \mathbb{E}\left[X_\alpha \mid \mathbb{G}(n, q_n) \right] \right).
\end{aligned} \tag{62}$$

Hence, $N_G - \mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right]$ reads off:

$$\sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \sum_{\emptyset\neq J\subseteq I} p_n^{|H|-|J|} \mathbb{1}_{\{(H\setminus J)^{(2)}\subset\mathbb{G}(n,q_n)\}} \prod_{\alpha\in J} (X_\alpha - \mathbb{E}\left[X_\alpha \mid \mathbb{G}(n,q_n)\right]). \tag{63}$$

The corresponding degenerate U-statistics in the decomposition are given for $J \subset \binom{[n]}{3}$ by

$$W_J \propto \left(\sum_{\substack{I\in\binom{[n]}{3}\\ I\simeq G, I\supseteq J}} 1 \right) \prod_{\alpha\in J} Y_\alpha, \tag{64}$$

where Y_α is the centered version of X_α given $\mathbb{G}(n,q_n)$ and: and:

$$w_J = \left(\sum_{\substack{I\in\binom{[n]}{3}\\ H\simeq G, I\supseteq J}} p_n^{|H|-|J|} \mathbb{1}_{\{(H\setminus J)^{(2)}\subset\mathbb{G}(n,q_n)\}} \right).$$

Historically, normal approximation for subgraph counting had been dealt with the method of moments [34] which requires tedious computations, but is quite adapted to this application. In particular, thresholds of asymptotic normality for the density of edges are obtained in function of n the number of vertices. In [3], the authors used Stein's method to derive convergence rates of the number of subgraph counting in random graphs in the 1-Wasserstein distance. The combination with Malliavin calculus has brought another feature to the usual coupling constructions in Stein's method, leveraging chaos representation property for independent identically distributed (see [29]). M_G is a Rademacher functional, so it has its Walsh chaotic decomposition. It has led to applications to subgraph counting in random graphs [30] and percolation problems [23]. By applying theorem 4.15 to (60), we obtain a quantitative version of the main theorem in [6] as well as its counterpart for $\mathbb{T}^{(3)}(n,q_n,p_n)$. To the best of our knowledge, there is no study of normal approximation of motif estimation in $\mathbb{T}^{(3)}(n,q_n,p_n)$. Let us denote $\bar{M}_G$ and $\bar{N}_G$ the respective rescaled statistic of the number of isomorphic copies of G with respect to their expectation, and let $\tilde{N}_G$ the rescaled statistic with respect to its conditional mean.

Theorem 5.1. *Let G a hypergraph without isolated vertices. Then,*

$$d_W(\tilde{N}_G, \mathcal{N}(0,1)) \leq C_{e_G} \sqrt{\sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]| / \operatorname{var}[\tilde{N}_G]}. \tag{65}$$

Proof of Theorem 5.1. We check whether the conditions of theorem 4.15 hold. For both statistics, the (EGF) assumption holds. By conditional independence of $(X_\alpha)_{\alpha\in\binom{[n]}{3}}$, we have:

$$\begin{aligned}\frac{\mathbb{E}[W_I^2]\mathbb{E}[W_J^2]}{\mathbb{E}[W_I^2 W_J^2]} &\propto \frac{\mathbb{E}\left[\prod_{\alpha\in I}\mathbb{E}[Y_\alpha^2|Z]\prod_{\alpha\in J}\mathbb{E}[Y_\alpha^2|Z]\right]}{\prod_{I\setminus J}\mathbb{E}[Y_\alpha^2]\prod_{J\setminus I}\mathbb{E}[Y_\alpha^2]\prod_{I\cap J}\mathbb{E}[Y_\alpha^2]} \\ &= (p_n(1-p_n))^{|I|+|J|-|I\cup J|}\, q_n^{|I^{(2)}|+|J^{(2)}|-|I^{(2)}\cup J^{(2)}|} \leq 1.\end{aligned}$$

Let us note that for all a, $W_{I\setminus\{a\}}$ is non-zero with the definition of $N_G - \mathbb{E}\left[N_G \mid Z\right]$. Let $W_I = w_I \prod_{i\in I} X_i$, then for $a \in I \cap J$:

$$\begin{aligned}\mathbb{E}\left[W_I W_J \mid \mathcal{G}_a\right] &= w_I w_J \prod_{i\in I\setminus\{a\}} Y_i \prod_{j\in J\setminus\{a\}} Y_j \mathbb{E}\left[Y_a^2 \;\middle|\; Z\right] \\ &= \frac{w_I w_J}{w_{I\setminus\{a\}} w_{J\setminus\{a\}}} \mathbb{E}\left[Y_a^2 \;\middle|\; Z\right] W_{I\setminus\{a\}} W_{J\setminus\{a\}} \\ &= C_{I,J,a} W_{I\setminus\{a\}} W_{J\setminus\{a\}},\end{aligned}$$

with

$$C_{I,J,a} = \frac{w_I w_J}{w_{I\setminus\{a\}} w_{J\setminus\{a\}}} \mathbb{E}\left[Y_a^2 \;\middle|\; Z\right] < +\infty \quad \mathbb{P}\text{-almost surely.}$$

□

We deduce those convergence rates for $p_n < c < 1$ for some c.

Theorem 5.2. *Let G a hypergraph without isolated vertices. Then, we have*

$$d_W(\tilde{N}_G, \mathcal{N}(0,1)) \lesssim \left(\min_{\substack{H\subset G \\ e_H>1}} \{n^{v_H} p_n^{e_H}\}\right)^{-1/2} \tag{66}$$

and

$$d_W(\tilde{N}_G, \mathcal{N}(0,1)) \lesssim \left(\min_{\substack{H\subset G \\ e_H>1}} \{n^{v_H} p_n^{e_H} q_n^{e_H^{(2)}}\}\right)^{-1/2}, \tag{67}$$

where $e_H^{(2)}$ is the number of edges included in the hyperedges of H.

Proof of theorem 5.2. We are left to upper bound the quantity:

$$\sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|$$

$$\propto_Z \sum_{(I,J,K,L) \text{ connected}} \left|\mathbb{E}\left[\mathbb{E}\left[\prod_{\alpha\in I} Y_\alpha \prod_{\alpha\in J} Y_\alpha \prod_{\alpha\in K} Y_\alpha \prod_{\alpha\in L} Y_\alpha \;\middle|\; Z\right]\right]\right|$$

where the notation $\propto_Z$ accounts for an equality up to a factor depending only on Z. The terms are non-zero if and only if α lies in at least two elements of the quadruple, i.e. if α does not lie in $I \setminus (J \cup K \cup L)$, etc. Then, the number of non-zero terms is $I \cup J \cup K \cup L$. We recall that:

$$\mathbb{E}[Y_\alpha|Z] = 0$$

$$\mathbb{E}[Y_\alpha^2|Z] = p_n(1-p_n)\mathbb{1}_{\{\alpha^{(1)}\in Z\}} \prod_{i=1}^{3} \mathbb{1}_{\{\alpha^{(i)}\in Z\}}$$

$$\mathbb{E}[Y_\alpha^3|Z] = p_n(1-p_n)(1-2p_n) \prod_{i=1}^{3} \mathbb{1}_{\{\alpha^{(i)}\in Z\}} \lesssim p_n(1-p_n) \prod_{i=1}^{3} \mathbb{1}_{\{\alpha^{(i)}\in Z\}}$$

$$\mathbb{E}[Y_\alpha^4|Z] = p_n(1-p_n)(1-3p_n(1-p_n)) \prod_{i=1}^{3} \mathbb{1}_{\{\alpha^{(i)}\in Z\}} \lesssim p_n(1-p_n) \prod_{i=1}^{3} \mathbb{1}_{\{\alpha^{(i)}\in Z\}}.$$

Now, we remark I, J, K, L are respectively isomorphic to A, B, C, D subhypergraphs of G. Hence, we can sum first over (A, B, C, D), and then over all the quadruples

(I,J,K,L) whose components are respectively isomorphic to the ones of the fixed quadruple (A,B,C,D). We shall write:

$$\sum_{I,J,K,L} \cdot = \sum_{A,B,C,D \subset G} \sum_{\substack{I \simeq A, J \simeq B \\ K \simeq C, L \simeq D}} \cdot := \sum_{A,B,C,D} \sum_{I,J,K,L}^{*A,B,C,D} \cdot$$

$v(A)$ denotes the number of vertices in A. We have that $|\{I,J,K,L \in \binom{[n]}{r} : I \simeq A, J \simeq B, K \simeq C, L \simeq D\}|$ is bounded by the number of collection of vertices of cardinal $v(A \cup B \cup C \cup D)$. By a counting argument, we see that is of order $n^{v(A \cup B \cup C \cup D)}$. Because (I,J,K,L) is connected and copies of subhypergraphs of G, we also have that $|I \cup J \cup K \cup L| = |A \cup B \cup C \cup D|$ and $|I^{(2)} \cup J^{(2)} \cup K^{(2)} \cup L^{(2)}| = |A^{(2)} \cup B^{(2)} \cup C^{(2)} \cup D^{(2)}|$. Hence, for a fixed connected quadruple (I,J,K,L) associated to (A,B,C,D),

$$|\mathbb{E}[W_I W_J W_K W_L]| \lesssim w_I w_J w_K w_L \, n^{v(A \cup B \cup C \cup D)} p_n^{|A \cup B \cup C \cup D|} q_n^{|A^{(2)} \cup B^{(2)} \cup C^{(2)} \cup D^{(2)}|}.$$

Let us bound the variance of $N_G - \mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right]$:

$$\begin{aligned} \mathrm{var}^2[N_G - \mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right]] &= \left(\sum_{I \cap J \neq \emptyset} \mathbb{E}[W_I W_J] \right)^2 = \sum_{I \cap J \neq \emptyset} \left(\mathbb{E}[W_I^2] + \mathbb{E}[W_J^2] \right)^2 \\ &= \frac{1}{2^2} \sum_{\substack{A,B \subset G \\ A \cap B \neq \emptyset}} \left(\sum_I^{*A} \mathbb{E}[W_I^2] + \sum_J^{*B} \mathbb{E}[W_J^2] \right)^2. \end{aligned}$$

For a fixed connected quadruple (A,B,C,D), by applying repeatedly the inequality $a^2 + b^2 \geq 2ab$, we get:

$$\begin{aligned} &\mathrm{var}^2[N_G - \mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right]] \\ &\qquad \geq \frac{1}{16} \left(\sum_I^{*A} \mathbb{E}[W_I^2] + \sum_J^{*B} \mathbb{E}[W_J^2] + \sum_K^{*C} \mathbb{E}[W_K^2] + \sum_L^{*D} \mathbb{E}[W_L^2] \right)^2 \\ &\qquad \geq \frac{1}{16} \left(\sum_I^{*A} \mathbb{E}[W_I^2] \times \sum_J^{*B} \mathbb{E}[W_J^2] \times \sum_K^{*C} \mathbb{E}[W_K^2] \times \sum_L^{*D} \mathbb{E}[W_L^2] \right)^{1/2}. \end{aligned}$$

Then using that $\mathbb{E}[W_I^2] = w_I^2 q_n^{|I^{(2)}|} (1-p_n)^{|I|} p_n^{|I|} = q_n^{|A^{(2)}|} (1-p_n)^{|A|} p_n^{|A|}$, so

$$\sum_I^{*A} \mathbb{E}[W_I^2] = \sum_I^{*A} w_I^2 (1-p_n)^{|A|} p_n^{|A|} q_n^{|A^{(2)}|} = n^{v(A)} (1-p_n)^{|A|} p_n^{|A|} q_n^{|A^{(2)}|}.$$

In particular, one has for a fixed quadruple (I,J,K,L) and associated (A,B,C,D):

$$\mathrm{var}^2[N_G - \mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right]] \geq \frac{1}{16} w_I w_J w_K w_L \left(n^{v^*(A,B,C,D)} (p_n(1-p_n))^{e^*(A,B,C,D)} q_n^{e^{(2)*}(A,B,C,D)} \right)^{1/2} \quad (68)$$

where $v^*(A,B,C,D) = v(A) + v(B) + v(C) + v(D)$ and $e^*(A,B,C,D) = |A| + |B| + |C| + |D|$ and $e^{(2)*}(A,B,C,D) = |A^{(2)}| + |B^{(2)}| + |C^{(2)}| + |D^{(2)}|$. It yields the result. Using the Lemma 9 of [6], for any quadruple (I,J,K,L) of collections I,J,K,L isomorphic to (sub)collections of H, there exists two (not both empty)

subcollections of G, say M and M', which may contain a nonzero number of isolated vertices, say i_M and $i_{M'}$, such that

$$v(M) + v(M') + i_M + i_{M'} = v(I) + v(J) + v(K) + v(L) - 2v(I \cup J \cup K \cup L), \quad (69)$$

$$|M^{(2)}| + |M'^{(2)}| = |I^{(2)}| + |J^{(2)}| + |K^{(2)}| + |L^{(2)}| - 2|I^{(2)} \cup J^{(2)} \cup K^{(2)} \cup L^{(2)}| \quad (70)$$

and by extension:

$$|M| + |M'| = |I| + |J| + |K| + |L| - 2|I \cup J \cup K \cup L|. \quad (71)$$

As G does not have isolated vertices, so do M and M'. As M and M' are subcollections of G, their average degree does not exceed $m(G)$. Hence, by theorem 5.1,

$$d_W(\tilde{N}_G, \mathcal{N}(0,1)) \lesssim (n^{v(M)+v(M')} p_n^{|M|+|M'|} q_n^{|M^{(2)}|+|M'^{(2)}|})^{1/2}.$$

Thus, (67) follows. The first result for M_G is obtained with $q_n = 1$. □

Remark 5.7. This bound is relevant only for the regime $p_n \xrightarrow{n\to\infty} 0$.

The Malliavin structure for conditionally independent random variables yields a chaos decomposition and rates of normal convergence of the conditionally centered statistic given Z.

5.4. **A modified Hoeffding decomposition.** In that section, we readopt the notations of the previous chapter by denoting A the index set of the random variables. Let another set $\hat{A}$ that index auxiliary random variables in addition to $(\hat{X}_\beta)_{\beta\in\hat{A}}$. We shall write the sequence of conditionally independent random variables given Z, $\mathsf{X} = (X_\alpha, \ldots, \hat{X}_\beta, \ldots)_{\alpha\in A, \beta\in\hat{A}}$ where the subsequence $(\hat{X}_\beta)_{\beta\in\hat{A}}$ is a sequence of independent random variables, and $\sigma(Z) = \sigma(\hat{X}_a, a \in A)$. This setting is new to the best of our knowledge, and is specifically tailored for the application in $\mathbb{T}^{(3)}(n, q_n, p_n)$. We assume that A is the set of 3-hyperedges, $\hat{A}$ is the set of edges included in the hyperedges of A and

$$X_\alpha = g(U_\alpha) \prod_{b\subset\alpha} \hat{X}_b \quad (72)$$

where $(U_\alpha)_{\alpha\in A}$ forms a sequence of conditionally independent random variables given $\hat{X}$, following the uniform distribution.

Lemma 5.3. *The sequence* X *is a sequence of conditionally independent random variables.*

Proof. Since, by assumption, for f bounded and $(\alpha, \beta) \in \hat{A}^2$ such that $\alpha \neq \beta$:

$$\mathbb{E}\left[f(\hat{X}_\beta) \,\middle|\, \hat{X}_\alpha, Z \right] = \mathbb{E}\left[f(\hat{X}_\beta) \,\middle|\, Z \right],$$

the subsequence $(\hat{X}_\beta)_{\beta\in\hat{A}}$ is a sequence of conditionally independent random variables given Z. Let $\alpha \in A$ and $\beta \in \hat{A}$, by definition

$$\mathbb{E}\left[f(X_\alpha) \mid \hat{X}_\beta, Z \right] = \mathbb{E}\left[f(X_\alpha) \mid Z \right],$$

and:

$$\mathbb{E}\left[f(\hat{X}_\beta) \,\middle|\, X_\alpha, Z \right] = \mathbb{E}\left[f(\hat{X}_\beta) \,\middle|\, Z \right]$$

because $\hat{X}_\beta$ is a function of Z. □

Remark 5.8. That type of sequence is a degenerate case of sequence of conditionally independent random variables since the $\hat{X}_\alpha$ are constant given Z.

For our purpose, the following lemma shows the commutation relation.

Lemma 5.4. *For $F \in L^2(E_A)$ a homogeneous sum of conditionally independent random variables X and $\alpha \in A$ and $\beta \in \hat{A}$ such that*

$$\mathbb{E}\left[\mathbb{E}\left[F \mid \hat{X}^{\{\beta\}}, X\right] \,\middle|\, X^{\{\alpha\}}, \hat{X}\right] = \mathbb{E}\left[\mathbb{E}\left[F \mid X^{\{\alpha\}}, \hat{X}\right] \,\middle|\, X, \hat{X}^{\{\beta\}}\right]. \tag{73}$$

Proof of lemma 5.4. It suffices to consider functionals of the type:

$$X_\alpha X_{\alpha_1} \dots X_{\alpha_n}$$

for $n \geq 1$. If β is not included in the edges of α, we have the property by independence of the associated random variables.

Let consider the case where β is one of the edge of α.

$$\mathbb{E}\left[F \mid X^{\{\alpha\}}, \hat{X}\right] = \mathbb{E}[g(U_\alpha)] \prod_{b \subset \alpha} \hat{X}_b \prod_{i=1}^{n} X_{\alpha_i}$$

and

$$\mathbb{E}\left[F \mid \hat{X}^{\{\beta\}}, X\right] = g(U_\alpha)\mathbb{E}\left[\hat{X}_\beta^{1+\sum_{i=1}^n \mathbb{1}_{\{\beta \in \alpha_i\}}}\right] \prod_{i=1}^{n} \prod_{b \subset \alpha_i} \hat{X}_b.$$

Then,

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[F \mid \hat{X}^{\{\beta\}}, X\right] \,\middle|\, X^{\{\alpha\}}, \hat{X}\right] &= \mathbb{E}[g(U_\alpha)]\mathbb{E}\left[\hat{X}_\beta^{1+\sum_{i=1}^n \mathbb{1}_{\{\beta \in \alpha_i\}}}\right] \prod_{i=1}^{n} \prod_{b \subset \alpha_i} \hat{X}_b. \\ &= \mathbb{E}\left[\mathbb{E}\left[F \mid X^{\{\alpha\}}, \hat{X}\right] \,\middle|\, X, \hat{X}^{\{\beta\}}\right].\end{aligned}$$

□

Those commutation relations of lemma 5.4 entail a modified Hoeffding decomposition of functionals of Bernoulli random variables.

Lemma 5.5. *Given X, the modified chaos decomposition is given by:*

$$F = \mathbb{E}[F] + \sum_{n=1}^{+\infty} \pi_n(F)$$

with

$$\pi_n(F) = \sum_{\substack{I \subset A \cup \hat{A} \\ |I|=n}} \left(\prod_{b \in I} D_b\right) \left(\prod_{c \in (A \cup \hat{A}) \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c]\right) \tag{74}$$

with $\mathcal{G}^c = \sigma(\mathsf{X}^{\{c\}})$.

Proof of lemma 5.5. We redefine a gradient D and Ornstein-Uhlenbeck operator L in the same fashion as in subsection 2.2 such that for $a \in A \cup \hat{A}$:

$$D_a F = F - \mathbb{E}\left[F \mid \mathsf{X}^{\{a\}}\right].$$

Then, we follow the same scheme of proof as lemma 2.6 with that modified gradient. Thus, we obtain that $\ker \mathsf{L} = \{F \in \operatorname{Dom} \mathsf{L} \,:\, \mathbb{E}[F] = 0\}$ and (74).

□

The resulting Malliavin framework is analogous to the Malliavin-Dirichlet structure in [9], whose underlying Markov process is the usual Glauber dynamics starting from X. It extends the scope to a particular type of sequences of conditionally independent random variables. That applies to N_G. We recall that in the application to motif estimation, Z is the underlying Erdös-Rényi random graph $\mathbb{G}(n, q_n)$. The

decomposition is similar to (62) except that this time the decomposition involves the random variables $(\hat{X}_b)_{b\in\hat{A}}$. Using the inclusion-exclusion principle,

$$\begin{aligned}\mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right] &= \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \prod_{\alpha\in H} \mathbb{E}\left[X_\alpha \mid \mathbb{G}(n,q_n)\right] = \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \prod_{\alpha\in H}\prod_{\beta\subset\alpha} \hat{X}_\beta \\ &= \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} p_n^{|H|} \prod_{\beta\in H^{(2)}} ((\hat{X}_\beta - \mathbb{E}[\hat{X}_\beta]) + \mathbb{E}[\hat{X}_\beta])\end{aligned}$$

Hence,

$$\mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right] - \mathbb{E}[N_G] = \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} p_n^{|H|} \sum_{\emptyset\neq J\subseteq H} q_n^{|H^{(2)}|-|J^{(2)}|} \prod_{\beta\in J^{(2)}} (\hat{X}_\alpha - q_n).$$

It entails that:

$$\begin{aligned}N_G - \mathbb{E}[N_G] &= (N_G - \mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right]) + (\mathbb{E}\left[N_G \mid \mathbb{G}(n,q_n)\right] - \mathbb{E}[N_G]) \\ &= \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \sum_{\emptyset\neq J\subseteq I} p_n^{|H|-|J|} \mathbb{1}_{\{(H\setminus J)^{(2)}\subset\mathbb{G}(n,q_n)\}} \prod_{\alpha\in J} (X_\alpha - \mathbb{E}\left[X_\alpha \mid \mathbb{G}(n,q_n)\right]) \\ &+ \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} p_n^{|H|} \sum_{\emptyset\neq J\subseteq H} q_n^{|H^{(2)}|-|J^{(2)}|} \prod_{\beta\in J^{(2)}} (\hat{X}_\alpha - q_n) \\ &= \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \sum_{\emptyset\neq J\subseteq I} p_n^{|H|-|J|} \prod_{\beta\in(H\setminus J)^{(2)}} \hat{X}_\beta \prod_{\alpha\in J} (X_\alpha - \mathbb{E}\left[X_\alpha \mid \mathbb{G}(n,q_n)\right]) \\ &+ \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} p_n^{|H|} \sum_{\emptyset\neq J\subseteq H} q_n^{|H^{(2)}|-|J^{(2)}|} \prod_{\beta\in J^{(2)}} (\hat{X}_\alpha - q_n) \\ &= \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \sum_{\emptyset\neq J\subseteq H} p_n^{|H|-|J|} \prod_{\beta\in(H\setminus J)^{(2)}} \hat{Y}_\beta \prod_{\alpha\in J} Y_\alpha \\ &+ \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \sum_{\emptyset\neq J\subseteq H} p_n^{|H|-|J|} q_n^{(H\setminus J)^{(2)}} \prod_{\alpha\in J} Y_\alpha \\ &+ \sum_{\substack{H\in\binom{[n]}{3}\\ H\simeq G}} \sum_{\emptyset\neq J\subseteq H} p_n^{|H|} q_n^{|H^{(2)}|-|J^{(2)}|} \prod_{\beta\in J^{(2)}} \hat{Y}_\beta\end{aligned}$$

where $Y_\alpha = X_\alpha - \mathbb{E}\left[X_\alpha \mid \mathbb{G}(n,q_n)\right]$ and $\hat{Y}_\beta = \hat{X}_\beta - \mathbb{E}[\hat{X}_\beta]$. It can be rewritten as $N_G - \mathbb{E}[N_G] = \sum_J W_J^{(1)} + W_J^{(2)} + W_J^{(3)}$ where

$$W_J^{[1]} = p_n^{e_G - |J|} q_n^{e_G^{(2)} - |J^{(2)}|} \left(\sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \prod_{\alpha \in J} Y_\alpha; \quad W_J^{[2]} = p_n^{e_G - |J|} q_n^{e_G^{(2)} - |J^{(2)}|} p_n^{|J|} \prod_{\beta \in J^{(2)}} \hat{Y}_\beta;$$

$$W_J^{[3]} = p_n^{e_G - |J|} \left(\sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \prod_{\beta \in (H \setminus J)^{(2)}} \hat{Y}_\beta \right) \prod_{\alpha \in J} Y_\alpha. \tag{75}$$

We consider the Malliavin structure associated to $\mathsf{Y} = (Y_\alpha, \dots, \hat{Y}_\beta, \dots)_{\alpha \in \binom{[n]}{3}, \beta \in \binom{[n]}{3}}$. Then, for each J, there exists $m \in \mathbb{N}$ such that $W_J^{(i)} \in \mathfrak{C}_m$ for $i \in \{1,2,3\}$.

Theorem 5.6. *Let G a hypergraph without isolated vertices. Then, let $p_n \xrightarrow{n \to +\infty} 0$ and $q_n \xrightarrow{n \to +\infty} 0$:*

$$d_W(\bar{N}_G, \mathcal{N}(0,1)) \lesssim \left(\min_{\substack{H \subset G \\ e_H > 1}} \{ n^{v_H} p_n^{e_H} q_n^{e_H^{(2)}} \} \right)^{-1/2}. \tag{76}$$

Proof of theorem 5.6. We follow the same lines as the proof of theorem 5.2, with the difference that $\pi_0(N_G) = \mathbb{E}[F]$. The (EGF) assumption holds. We recall the bound in our context

$$d_W(\bar{N}_G, \mathcal{N}(0,1)) \le C_{e_G} \sqrt{\sum_{(I,J,K,L) \text{ connected}} \sum_{i_i, i_j, i_k, i_l = 1}^{3} |\mathbb{E}[W_I^{[i_i]} W_J^{[i_j]} W_K^{[i_k]} W_L^{[i_l]}]| / \operatorname{var}[N_G]}. \tag{77}$$

As each connected quadruple (I,J,K,L) is associated to (H_1,H_2,H_3,H_4) subhypergraphs of G such that $I \simeq H_1$, $J \simeq H_2$, $K \simeq H_3$ and $L \simeq H_4$, from theorem 5.2, we have:

$$\begin{aligned} &|\mathbb{E}[W_I^{[1]} W_J^{[i_j]} W_K^{[i_K]} W_L^{[i_l]}]| \\ &\quad \le w_I w_J w_K w_L \, n^{v(H_1 \cup H_2 \cup H_3 \cup H_4)} p_n^{|H_1 \cup H_2 \cup H_3 \cup H_4|} q_n^{|H_1^{(2)} \cup H_2^{(2)} \cup H_3^{(2)} \cup H_4^{(2)}|} \end{aligned} \tag{78}$$

for $i_j, i_k, i_l \in \{1,2\}$ as $\prod_{\alpha \in J} \mathbb{E}\left[Y_\alpha \mid Z\right] \propto_Z \prod_{\beta \in J^{(2)}} \hat{Y}_\beta$ and $\mathbb{E}[\hat{Y}_\beta^k] \propto p_n$ for $k \ge 2$. As $\hat{X}_\beta \le 1$ a.s., we also have (78) for all $i_i, i_j, i_k, i_l \in \{1,2,3,4\}$. Likewise, the variance reads off in function of the quadruples:

$$\operatorname{var}[N_G] = \frac{1}{4} \sum_{\substack{H_1, H_2 \subset G \\ H_1 \cap H_2 \neq \emptyset}} \left(\sum_I^{*H_1} \mathbb{E}[W_I^2] + \sum_J^{*H_2} \mathbb{E}[W_J^2] \right)$$

where $\sum_I^{*H_1} \cdots$ stands for a sum over I such that I is isomorphic to H_1.

We follow the same lines of computations as those leading to (68). Then, for fixed quadruples (H_1,H_2,H_3,H_4),

$$\operatorname{var}[N_G] \ge \frac{1}{16} \left(\sum_I^{*H_1} \mathbb{E}[W_I^2] \times \sum_J^{*H_2} \mathbb{E}[W_J^2] \times \sum_K^{*H_3} \mathbb{E}[W_K^2] \times \sum_L^{*H_4} \mathbb{E}[W_L^2] \right)^{1/2}. \tag{79}$$

As $\mathbb{E}[W_I^2] = p_n^{|H| - |H_1|} q_n^{|H^{(2)}| - |H_1^{(2)}|} q_n^{|H_1^{(2)}|} (1-q_n)^{|H_1^{(2)}|} (1-p_n)^{|H_1|} p_n^{|H_1|}$, we have:

$$\sum_{(I,J,K,L) \text{ connected}} \sum_{i_i,i_j,i_k,i_l=1}^{3} |\mathbb{E}[W_I^{[i_i]} W_J^{[i_j]} W_K^{[i_k]} W_L^{[i_l]}]| / \operatorname{var}[N_G]$$

$$\leq \frac{n^{v(H_1\cup H_2\cup H_3\cup H_4)} p_n^{|H_1\cup H_2\cup H_3\cup H_4|} q_n^{|H_1^{(2)}\cup H_2^{(2)}\cup H_3^{(2)}\cup H_4^{(2)}|}}{\left(n^{v(H_1)+v(H_2)+v(H_3)+v(H_4)} p_n^{|H_1|+|H_2|+|H_3|+|H_4|} q_n^{|H_1^{(2)}|+|H_2^{(2)}|+|H_3^{(2)}|+|H_4^{(2)}|} \right)^{1/2}}.$$

At that point, we arrive at the same upper bound as in the proof of theorem 5.2. □

While in [22, 36], the probability of keeping a hyperedge does not depend on the number of vertices, we let p_n tend to 0. As a consequence, we can state thresholds for subhypergraph containment that complement the ones in [19, p.61]. As done in [22, 38] for random graphs, it should be possible to derive with our method the convergence rates considering an arbitrary exchangeable random hypergraph generated by a hypergraphon, the analog of graphon in graph limit theory.

References

[1] Tim Austin. On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probability Surveys*, 5(none):80–145, January 2008.

[2] Ehsan Azmoodeh, Simon Campese, and Guillaume Poly. Fourth moment theorems for Markov diffusion generators. *Journal of Functional analysis*, 266(4):2341–2359, 2014.

[3] Andrew D Barbour, Michał Karoński, and Andrzej Ruciński. A central limit theorem for decomposable random variables with applications to random graphs. *Journal of Combinatorial Theory, Series B*, 47(2):125–145, 1989.

[4] Sourav Chatterjee. A new method of normal approximation. *The Annals of Probability*, 36(4):1584–1610, 2008.

[5] Peter De Jong. A Central Limit Theorem for Generalized Multilinear Forms. *Journal of Multivariate Analysis*, 34(2):275–289, 1990.

[6] Peter De Jong. A Central Limit Theorem with Applications to Random Hypergraphs. *Random Structures and Algorithms*, 8(2):105–120, 1996.

[7] Laurent Decreusefond. The Stein-Dirichlet-Malliavin method. *ESAIM: Proceedings and Surveys*, 51:49–59, 2015.

[8] Laurent Decreusefond. *Selected Topics in Malliavin Calculus: Chaos, Divergence and So Much More*. Springer Nature, 2022.

[9] Laurent Decreusefond and Hélène Halconruy. Malliavin and Dirichlet structures for independent random variables. *Stochastic Processes and their Applications*, 129(8):2611–2653, August 2019.

[10] Christian Döbler and Giovanni Peccati. Quantitative de Jong theorems in any dimension. *Electronic Journal of Probability*, 22, 2017.

[11] Christian Döbler and Giovanni Peccati. Quantitative CLTs for symmetric U-statistics using contractions. *Electronic Journal of Probability*, 24, 2019.

[12] Richard M. Dudley. *Real Analysis and Probability*. Cambridge studies in advanced mathematics. Cambridge University Press, 2002.

[13] Mitia Duerinckx. On the size of chaos via Glauber calculus in the classical mean-field dynamics. *Communications in Mathematical Physics*, pages 1–41, 2021.

[14] Nguyen Tien Dung. Poisson and normal approximations for the measurable functions of independent random variables. *arXiv preprint arXiv:1807.10925*, 2018.

[15] Nguyen Tien Dung. Rates of convergence in the central limit theorem for nonlinear statistics under relaxed moment conditions. *Acta Mathematica Vietnamica*, pages 1–26, 2021.

[16] Bradley Efron and Charles Stein. The jackknife estimate of variance. *The Annals of Statistics*, pages 586–596, 1981.

[17] Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, pages 293–325, 1948.

[18] Christian Houdré and Nicolas Privault. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, 8(6):697–720, 2002.

[19] Svante Janson, Tomasz Luczak, and Andrzej Ruciński. *Random Graphs*. John Wiley & Sons, Inc., 2000.

[20] Svante Janson and Krzysztof Nowicki. The asymptotic distributions of generalized U-statistics with applications to random graphs. *Probability theory and related fields*, 90(3):341–375, 1991.
[21] Olav Kallenberg. *Foundations of Modern Probability*. Springer, 1997.
[22] Gursharn Kaur and Adrian Röllin. Higher-order fluctuations in dense random graph models. *Electronic Journal of Probability*, 26:1–36, 2021.
[23] Kai Krokowski, Anselm Reichenbachs, and Christoph Thäle. Discrete Malliavin–Stein method: Berry–Esseen bounds for random graphs and percolation. *The Annals of Probability*, 45(2):1071–1109, 2017.
[24] Raphaël Lachièze-Rey and Giovanni Peccati. New Berry–Esseen bounds for functionals of binomial point processes. *The Annals of Applied Probability*, 27(4):1992–2031, 2017.
[25] Michel Ledoux. Chaos of a Markov operator and the fourth moment condition. *Annals of Probability*, 40(6):2439–2459, 2012.
[26] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.
[27] Colin McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.
[28] Ivan Nourdin and Giovanni Peccati. *Normal approximations with Malliavin calculus: from Stein's method to universality*. Number 192 in Cambridge Tracts in Mathematics. Cambridge University Press, 2012.
[29] Nicolas Privault. Stochastic analysis of Bernoulli processes. *Probability Surveys*, 5:435–483, 2008.
[30] Nicolas Privault and Grzegorz Serafin. Normal approximation for sums of discrete U-statistics – Application to Kolmogorov bounds in random subgraph counting. *Bernoulli*, 26(1):587–615, 2020.
[31] Nicolas Privault and Grzegorz Serafin. Berry-Esseen bounds for functionals of independent random variables. *Electronic Journal of Probability*, 27:1–37, 2022.
[32] B.L.S. Prakasa Rao. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61(2):441–460, 2009.
[33] Gesine Reinert, Ivan Nourdin, and Giovanni Peccati. Stein's method and stochastic analysis of Rademacher functionals. *Electronic Journal of Probability*, 15:1703–1742, 2010.
[34] Andrzej Ruciński. When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields*, 78(1):1–10, 1988.
[35] Adrian Röllin. Kolmogorov bounds for the normal approximation of the number of triangles in the Erdos-Rényi random graph. *Probability in the Engineering and Informational Sciences*, 36(3):747–773, 2022.
[36] Tadas Temčinas, Vidit Nanda, and Gesine Reinert. Multivariate Central Limit Theorems for Random Clique Complexes, June 2022. arXiv:2112.08922 [math] type: article.
[37] Kösaku Yosida. *Functional Analysis*. Classics in Mathematics. Springer Berlin Heidelberg, 1995.
[38] Zhuo-Song Zhang. Berry–Esseen bounds for generalized U-statistics. *Electronic Journal of Probability*, 27:1–36, 2022.