Electronic Journal of Probability

Electron. J. Probab. **25** (2020), article no. 144, 1–28.
ISSN: 1083-6489 https://doi.org/10.1214/20-EJP538

On almost sure convergence of random variables with finite chaos decomposition*

Radosław Adamczak†

Abstract

Under mild conditions on a family of independent random variables (X_n) we prove that almost sure convergence of a sequence of tetrahedral polynomial chaoses of uniformly bounded degrees in the variables (X_n) implies the almost sure convergence of their homogeneous parts. This generalizes a recent result due to Poly and Zheng obtained under stronger integrability conditions. In particular for i.i.d. sequences we provide a simple necessary and sufficient condition for this property to hold.

We also discuss similar phenomena for sums of multiple Wiener-Itô integrals with respect to Poisson processes, answering a question by Poly and Zheng.

Keywords: multiple stochastic (Wiener-Itô) integrals; polynomial chaos; random multi-linear forms; Poisson process.
MSC2020 subject classifications: 60F99; 60H05; 60B11.
Submitted to EJP on November 4, 2019, final version accepted on October 22, 2020.
Supersedes arXiv:1909.09576.

1 Introduction

Investigation of real and vector valued multi-linear forms in independent random variables is a classical topic in probability theory closely related to multiple Wiener-Itô integration. Such random variables have been thoroughly studied, e.g., in the context of harmonic analysis on the discrete cube, analysis of Boolean functions, geometric theory of Banach spaces, random graphs, concentration of measure, Malliavin calculus or more recently the Malliavin-Stein method. We refer the reader to the monographs [11, 3, 14, 25, 7, 18, 16, 6, 19] for extensive exposition of various aspects of the theory.

Recently Poly and Zheng [22] have observed that for a large class of sequences $\mathbb{X} = (X_n)_{n\in\mathbb{N}}$ of independent random variables the almost sure convergence of a sequence of sums of tetrahedral (i.e., affine in each variable) multi-linear forms of bounded degrees in the sequence $\mathbb{X}$ can be decomposed into the almost sure convergence of their homogeneous parts. They also proved a counterpart of this result for sums of multiple

*Research partially supported by the National Science Centre, Poland, grant no. 2015/18/E/ST1/00214.
†Institute of Mathematics, University of Warsaw, Poland. E-mail: r.adamczak@mimuw.edu.pl

Wiener-Itô integrals with respect to a Gaussian process and posed certain questions concerning similar phenomena for sequences of variables with less regularity than those covered by their theorems, as well as for sums of multiple Wiener-Itô integrals with respect to a Poisson process.

The goal of this article is to provide answers to the questions raised by Poly and Zheng and to further study the almost sure convergence of sums of tetrahedral multi-linear forms, also in the vector valued setting. In order to formulate our results in a precise way and to put them in the right perspective let us start with the formulation of the main theorems by Poly and Zheng.

1.1 Results by Poly and Zheng

Denote by $\ell_0(\mathbb{N})^{\odot d}$ the set of all d-tensors (d-indexed matrices) of the form $a = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d=0}^{\infty}$, symmetric in their arguments (i.e., $a_{i_1,\ldots,i_d} = a_{i_{\sigma(1)},\ldots,i_{\sigma(d)}}$ for any permutation σ of the set $[d] = \{1,\ldots,d\}$), with vanishing diagonals (i.e., such that $a_{i_1,\ldots,i_d} = 0$ whenever $i_k = i_l$ for some $k \neq l$). For $d = 0$ we interpret $a \in \ell(\mathbb{N})^{\odot d}$ as a single real number $a_\emptyset$ (corresponding to the empty multi-index).

Let $X_0, X_1, X_2, \ldots$ be a family of independent random variables. Assume that $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = 1$ and for some $\delta > 0$, $\sup_i \mathbb{E}|X_i|^{2+\delta} < \infty$.

Assume now that $(Z_n)_{1\le n\le\infty}$ is a sequence of random variables of the form[1]

$$Z_n = \sum_{k=0}^{d} Z_{n,k},$$

where

$$Z_{n,k} = \sum_{i_1,\ldots,i_k=0}^{\infty} a^{(n,k)}_{i_1,\ldots,i_k} X_{i_1}\cdots X_{i_k}$$

for some $a^{(n,k)} \in \ell_0(\mathbb{N})^{\odot k}$ such that $\sum_{i_1,\ldots,i_k=1}^{\infty} |a^{(n,k)}_{i_1,\ldots,i_k}|^2 < \infty$. Here the infinite sums defining $Z_{n,k}$ are understood as a.s. (or L_2) limits of sums over $i_1,\ldots,i_k \in \{0,\ldots,n\}$ (their existence follows easily from the martingale convergence theorem). Note that $Z_{n,0}$ are just constants (products over empty index set are interpreted as one).

One of the results proved by Poly and Zheng is

Theorem 1.1 (Theorem 1.3. in [22])**.** *In the above setting, if Z_n converges to Z_∞ a.s. as $n \to \infty$, then for all $k \le d$, $Z_{n,k} \to Z_{\infty,k}$ a.s.*

In other words the almost sure convergence of sums of tetrahedral multilinear forms of uniformly bounded degrees in the variables X_i decomposes into almost sure convergence of their homogeneous components.

While we postpone the rigorous formulation of our results to subsequent sections, let us announce that we provide a weaker sufficient conditions for this property to hold (see Theorem 2.8), which in particular allows to replace the finiteness of higher moments in Theorem 1.1 by uniform square integrability (Corollary 2.9). We also completely characterize i.i.d. sequences with the above property (Theorem 2.13) and extend this phenomenon to the case of multi-linear forms with coefficients from a Banach space (Proposition 2.4).

Another result from [22] is a counterpart of Theorem 1.1 for sums of Gaussian multiple Wiener-Itô integrals. Since we are not going to use it (we state it only for comparison with the Poissonian case which we will consider in Section 3) we refer, e.g., to the monograph [7] for the necessary definitions. We remark that the original formulation of the theorem involved rather isonormal Gaussian processes over a separable Hilbert

[1] Note that we include here $n = \infty$.

space. To be able to draw analogy with the Poissonian setting, we state it in an equivalent form in terms of Gaussian stochastic measures.

Theorem 1.2 (Theorem 1.1. in [22]). *Let G be a Gaussian stochastic measure on a measure space $(\mathcal{X}, \mathcal{F}, \mu)$ and let I_n denote the corresponding n-fold Gaussian stochastic integral on $L_{2,s}(\mathcal{X}^n, \mathcal{F}^{\otimes n}, \mu^{\otimes n})$ (the space of square integrable functions, symmetric in their arguments). Let $d \in \mathbb{N}$ and consider a sequence $(F_n)_{n=0}^{\infty}$ of random variables of the form*

$$F_n = \mathbb{E}F_n + \sum_{k=1}^{d} I_k(f_{n,k}),$$

where $f_{n,k} \in L_{2,s}(\mathcal{X}^k, \mathcal{F}^{\otimes k}, \mu^{\otimes k})$ and $d \in \mathbb{N}$. If the sequence F_n converges almost surely to a random variable F, then $\mathbb{E}F_n \to \mathbb{E}F$ and there exist functions $f_{\infty,k} \in L_{2,s}(\mathcal{X}^k, \mathcal{F}^{\otimes k}, \mu^{\otimes k})$, such that for all $k \le d$, $I_k(f_{n,k})$ converges almost surely as $n \to \infty$ to $I_k(f_{\infty,k})$.

Poly and Zheng ask if an analogous result holds for Poisson multiple Wiener-Itô integrals. While we show (see Example 3.1 below) that this is not the case (even for $d = 1$), we will also prove that under an additional assumption that the converging sequence is majorized by an integrable random variable, one can indeed deduce the almost sure convergence of individual summands from the convergence of the sum (Theorem 3.2).

Let us mention in passing that there are many common aspects of the analysis on the Gauss and Poisson space, e.g., they both have the chaos representation property, however the Poisson space lacks many regularity aspects of the Gauss space (e.g., hypercontractivity and related strong concentration properties). Searching for counterparts of Gaussian results in the context of Malliavin calculus, concentration of measure or hypercontractivity is an active area of research (see, e.g., the recent articles [12, 24, 17]). Our result is another example showing that the behaviour of multiple Wiener-Itô integrals with respect to the Poisson process resembles to some extent the Gaussian case, however at the cost of introducing some additional assumptions.

2 Results for independent random variables

We will now present new results for independent random variables, deferring the proofs to further sections. We will start by discussing certain general properties, then we will state the main theorems concerning extensions of Theorem 1.1.

2.1 Preliminaries

In order to make the presentation more transparent we need to introduce some additional terminology. Below $\mathbb{X} = (X_i)_{i\in\mathbb{N}}$ is a sequence of independent random variables.

Definition 2.1. For a nonnegative integer d define $Q_d(\mathbb{X})$ – the *homogeneous tetrahedral chaos of degree* d, as the space of all random variables Z, which are limits in probability of a sequence of random variables of the form

$$\sum_{i_1,\ldots,i_d=0}^{\infty} a_{i_1,\ldots,i_d} X_{i_1} \cdots X_{i_d},$$

where $a \in \ell_0(\mathbb{N})^{\odot d}$ is a d-tensor with only finitely many non-zero coefficients.

Remark 2.2. If the sequence $\mathbb{X}$ consists of i.i.d. Rademacher variables, then $Q_d(\mathbb{X})$ coincides with the Walsh-Rademacher chaos of order d, however if the variables X_i are i.i.d. standard Gaussian, Q_d is distinct from the d-th Wiener-Itô chaos corresponding to the Gaussian Hilbert space spanned by $\mathbb{X}$ (it is just a proper subspace). Since in this section we discuss only sequences of independent random variables, we believe that this should not lead to misunderstanding. Let us also remark that in general the spaces Q_d may have a non-trivial intersection (see however Proposition 2.5 below) and (even if all variables X_i are square integrable) they need not span $L_2(\mathbb{X})$. Indeed, if the variables are i.i.d. standard Gaussian, then, e.g., $X_0^2 - 1$ is orthogonal to all the spaces $Q_d(\mathbb{X})$, whereas it is an element of the second chaos understood in the Wiener-Itô sense (i.e., the orthogonal complement of the subspace of $L_2(\mathbb{X})$ spanned by polynomials of degree one in the subspace spanned by polynomials of degree two). If one works in the setting of Gaussian stochastic measures, as in Theorem 1.2, and one assumes that μ is non-atomic, then thanks to infinite-divisibility one can approximate in L_2 any element of the d-th Wiener-Itô chaos by a tetrahedral polynomial in Gaussian variables (we will use a similar idea in the proofs of results for the Poisson space).

Note also that $Q_0(\mathbb{X})$ is just the space of almost surely constant random variables.

Definition 2.3. We will say that $\mathbb{X}$ has the *convergence decomposition property* (abbrev. CDP) if for every nonnegative integer d and every sequence $(Z_n)_{1\leq n\leq\infty}$ of random variables of the form

$$Z_n = \sum_{k=0}^{d} Z_{n,k}, \tag{2.1}$$

where $Z_{n,k} \in Q_k(\mathbb{X})$, such that $Z_n \to Z_\infty$ almost surely as $n \to \infty$, we have

$$Z_{n,k} \to Z_{\infty,k} \text{ a.s.} \tag{2.2}$$

for all $k = 0, \ldots, d$.

The results by Poly and Zheng have been formulated for real valued chaos variables, however it turns out that the CDP automatically extends to an analogous property for polynomial chaoses with coefficients in an arbitrary separable Banach space $(E, \|\cdot\|)$. More precisely, if we define $Q_d(\mathbb{X}, E)$ as the sets of limits in probability of homogeneous tetrahedral polynomials of degree d in $\mathbb{X}$, with coefficients from E, then the following result holds.

Proposition 2.4. The sequence $\mathbb{X}$ satisfies the CDP iff for every separable Banach space $(E, \|\cdot\|)$, every nonnegative integer d and every sequence $(Z_n)_{1\leq n\leq\infty}$ of random variables of the form

$$Z_n = \sum_{k=0}^{d} Z_{n,k},$$

where $Z_{n,k} \in Q_k(\mathbb{X}, E)$, such that $Z_n \to Z_\infty$ almost surely as $n \to \infty$, we have

$$Z_{n,k} \to Z_{\infty,k}$$

almost surely for all $k = 0, \ldots, d$.

An obvious necessary condition for the sequence $\mathbb{X}$ to satisfy the CDP is linear independence of the spaces $Q_k(\mathbb{X})$, i.e., uniqueness of representations of random variables Z as sums of variables from a finite number of spaces $Q_k(\mathbb{X})$ (if such uniqueness does not hold then the sequence $Z_n = Z$ together with two distinct representations provides a counterexample for the CDP). The following proposition asserts that this minimal condition of uniqueness of the chaos decomposition is in fact also sufficient for the CDP.

Proposition 2.5. A sequence $\mathbb{X}$ satisfies the CDP if and only if for every $d \in \mathbb{N}$ and every $Y_0, Y_0' \in Q_0(\mathbb{X}), \ldots, Y_d, Y_d' \in Q_d(\mathbb{X})$, if

$$Y_0 + Y_1 + \ldots + Y_d = Y_0' + Y_1' + \ldots + Y_d' \text{ a.s.},$$

then for all $k \leq d$, $Y_k = Y_k'$ a.s.

Remark 2.6. In fact, as follows from our main technical tool, Lemma 5.1 in Section 5, if the CDP does not hold, then we can find finite sums $Z_n = b_n + \sum_{k=0}^{k_n} a_k^{(n)} X_k$, where $b_n, a_k^{(n)} \in \mathbb{R}$, such that $Z_n \to 0$ almost surely while $b_n \to -1$, $Z_n - b_n \to 1$ a.s. In particular the uniqueness of the decomposition is lost already for $d = 1$.

Let us conclude this section with a comment on the assumed structure of the limiting random variable Z_∞. In the formulation of Theorem 1.1 and Proposition 2.4 as well as in Definition 2.3 it is assumed that Z_∞ can be also represented as a finite sum of variables from $Q_k(\mathbb{X})$. The next proposition states that if $\mathbb{X}$ satisfies the CDP, then it is in fact enough to assume just the existence of the limit.

Proposition 2.7. Assume that the sequence $\mathbb{X}$ satisfies the CDP and let $(E, \|\cdot\|)$ be a separable Banach space. Consider a sequence of random variables $(Z_n)_{1 \leq n < \infty}$ as in (2.1), with $Z_{n,k} \in Q_k(\mathbb{X}, E)$. If the sequence Z_n converges in probability to some random variable Z_∞, then there exist unique random variables $Z_{\infty,k}$, $k = 0, \ldots, d$ such that $Z_\infty = \sum_{k=0}^{d} Z_{\infty,k}$ and $Z_k \in Q_k(\mathbb{X}, E)$.

2.2 Main results

We will now present the main results for sequences of independent random variables. We will start with a mild sufficient condition for the CDP to hold. Before we formulate it let us note that since the spaces $Q_k(\mathbb{X})$ do not change when one scales the variables X_n by nonzero factors, there is no loss of generality in assuming that $\mathbb{X}$ is a tight sequence.

Theorem 2.8. *Let $\mathbb{X}$ be a tight sequence of independent random variables. Assume that for some $\delta > 0$ and all $n \in \mathbb{N}$,*

$$\sup_{x \in \mathbb{R}} \mathbb{P}(X_n \in (x - \delta, x + \delta)) \leq 1 - \delta. \tag{2.3}$$

Assume moreover that there exist $C \geq 0$, $t_0 > 0$ such that for any $t \in (0, t_0]$ and any $n \in \mathbb{N}$,

$$\left| \mathbb{E} X_n \mathbb{1}_{\{|X_n| \leq \frac{1}{t}\}} \right| \leq C\Big(\frac{1}{t} \mathbb{P}\Big(|X_n| > \frac{1}{t} \Big) + t \operatorname{Var}\Big(X_n \mathbb{1}_{\{|X_n| \leq \frac{1}{t}\}} \Big) \Big). \tag{2.4}$$

Then the sequence $\mathbb{X}$ satisfies the CDP.

The condition (2.4) may seem quite technical, therefore let us now state a corollary to the above theorem, which is a strengthening of Theorem 1.1.

Corollary 2.9. If the variables X_n are centered of variance one and the family $\{X_n^2\}_{n \in \mathbb{N}}$ is uniformly integrable, then $\mathbb{X}$ satisfies the CDP.

Remark 2.10. The assumption (2.3) is an anti-concentration type condition, preventing the random variables X_n from being *too deterministic*. It is not difficult to see that if the variables become too strongly concentrated in points or small intervals away from zero, then the CDP cannot hold, as illustrated by the following example, which answers a question posed by Poly and Zheng [22].

Example 2.11. Assume that X_n is a centered two point variable of variance one (not necessarily symmetric), i.e., for some $p_n \in (0, 1)$,

$$\mathbb{P}\Big(X_n = \frac{1 - p_n}{\sqrt{p_n(1 - p_n)}} \Big) = p_n, \; \mathbb{P}\Big(X_n = \frac{-p_n}{\sqrt{p_n(1 - p_n)}} \Big) = 1 - p_n.$$

Assume that $\limsup_{n\to\infty} p_n = 1$ (the situation when $\liminf_{n\to\infty} p_n = 0$ is completely analogous). In particular there exists an increasing sequence k_n such that

$$\sum_{n=0}^{\infty}(1-p_{k_n}) < \infty.$$

Set now $Z_{n,0} = 1$, $Z_{n,1} = -\frac{\sqrt{p_{k_n}(1-p_{k_n})}}{1-p_{k_n}} X_{k_n}$, $Z_n = Z_{n,0} + Z_{n,1}$. By the Borel-Cantelli lemma we obtain that with probability one, for sufficiently large n, $X_{k_n} = \frac{1-p_{k_n}}{\sqrt{p_{k_n}(1-p_{k_n})}}$ and as a consequence $Z_{n,1} = -1$, $Z_n = 0$. Setting $Z_{\infty,0} = Z_{\infty,1} = 0$ one can see that there is no convergence of $Z_{n,i}$ to $Z_{\infty,i}$. One can also see that the decomposition of Z_∞ into a sum of variables from $Q_0(\mathbb{X})$ (constants) and $Q_1(\mathbb{X})$ is not unique, since $-1 \in Q_1(\mathbb{X})$ (cf. Proposition 2.5 and Remark 2.6). Of course if one insists on representing Z_∞ in the form $c + \sum_{n=0}^{\infty} a_n X_n$, where $c, a_n \in \mathbb{R}$ then one must have $c = 0$, $a_n \equiv 0$.

On the other hand, if the numbers p_n are separated from zero and one, then it follows from Theorem 1.1 that the sequence $\mathbb{X}$ satisfies the CDP.

Remark 2.12 (A note added in revision)**.** When the present article was under review Pratelli and Rigo [23] provided an example of a sequence p_n and a sequence of polynomials Z_n of degree two in the variables X_n such that $Z_n \to 0$ a.s. *and* in L_2, while $Z_{n,1}$ does not converge almost surely to zero.

If the variables X_n are i.i.d., one can show that the condition of Theorem 2.8 is in fact necessary for the CDP to hold, i.e., we have the following theorem.

Theorem 2.13. *Assume that the variables X_n, $n \in \mathbb{N}$ are i.i.d. Then the sequence $\mathbb{X}$ satisfies the CDP if and only if there exists $C \geq 0$ and $t_0 > 0$ such that for all $t \in (0, t_0)$,*

$$\Big|\mathbb{E}X_0 \mathbb{1}_{\{|X_0| \leq \frac{1}{t}\}}\Big| \leq C\Big(\frac{1}{t}\mathbb{P}\Big(|X_0| > \frac{1}{t}\Big) + t\,\mathrm{Var}\Big(X_0 \mathbb{1}_{\{|X_0|\leq \frac{1}{t}\}}\Big)\Big). \tag{2.5}$$

Remark 2.14. The proof of the above theorem is presented (along with the proofs of other results concerning independent random variables) in Section 5, let us however already now explain briefly the reasons behind the simplification of the conditions for i.i.d. sequences and the fact that in this case one can provide a full characterization of the CDP. Tightness (which as indicated before Theorem 2.8, is just a technical condition one assumes without loss of generality to simplify the statements for the results) follows immediately from the i.i.d. assumption. As for the condition (2.3), it is easy to see that in the i.i.d. case it can fail only in the degenerate case, i.e., when X_0 is deterministic. This situation can be easily treated separately from the truly random case. Since the condition (2.4) in the i.i.d. situation reduces to (2.5), sufficiency of the latter in the non-degenerate case follows directly from Theorem 2.8. As for the necessity, if (2.5) fails, then the i.i.d. assumption and the Law of Large Numbers allow for a construction of an appropriate sequence of averages of the variables X_i, converging almost surely to a non-zero constant, yielding a counterexample to the CDP.

Remark 2.15. Using the fact that $\lim_{t\searrow 0} t|\mathbb{E}X_0\mathbb{1}_{\{|X_0|\leq\frac{1}{t}\}}| = 0$, it is easy to see that (2.5) is satisfied for some $C \geq 0$, $t_0 > 0$ and all $t \in (0, t_0)$ if and only if for some $C_1 \geq 0$, $t_1 > 0$ and all $t \in (0, t_1)$,

$$\Big|\mathbb{E}X_0 \mathbb{1}_{\{|X_0| \leq \frac{1}{t}\}}\Big| \leq C_1\Big(\frac{1}{t}\mathbb{P}\Big(|X_0| > \frac{1}{t}\Big) + t\mathbb{E}X_0^2 \mathbb{1}_{\{|X_0|\leq \frac{1}{t}\}}\Big). \tag{2.6}$$

Furthermore, this is equivalent to the existence of $C_2 \geq 0$, such that

$$\Big|\mathbb{E}X_0 \mathbb{1}_{\{|X_0| \leq \frac{1}{t}\}}\Big| \leq C_2\Big(\frac{1}{t}\mathbb{P}\Big(|X_0| > \frac{1}{t}\Big) + t\mathbb{E}X_0^2 \mathbb{1}_{\{|X_0|\leq \frac{1}{t}\}}\Big) \tag{2.7}$$

for all $t > 0$.

Indeed, if X_0 is not equal identically to zero, then for t large enough (say $t > t_2$) and C_3 large enough, $\left|\mathbb{E}X_0\mathbb{1}_{\{|X_0|\le \frac{1}{t}\}}\right| \le \frac{1}{t} \le C_3\frac{1}{t}\mathbb{P}(|X_0| > \frac{1}{t})$, while (2.7) for $t \in [t_1, t_2]$ can be easily obtained from (2.6) for $t < t_1$ (with C_2 depending only on C_1, t_1, t_2).

Using the Fubini theorem and (2.7) one can easily prove that if Y is any random variable independent of X_0 and X_0 satisfies (2.5), then so does X_0Y (possibly with different t_0, C). This clearly follows from Theorem 2.13 but is perhaps somewhat hidden at the level of inequality (2.5).

Example 2.16. It is clear from the law of large numbers that if $\mathbb{X}$ is an i.i.d. sequence with X_0 integrable but not centered, then $\mathbb{X}$ cannot satisfy the CDP. Let us present a sequence violating the CDP with $\mathbb{E}X_0 = 0$. To this end consider X_0 satisfying

$$\mathbb{P}\Big(X_0 = \frac{2^n}{n^2}\Big) = \frac{1}{2^{n+1}} \text{ for } n = 1, 2, \ldots,$$

and

$$\mathbb{P}\Big(X_0 = -\frac{\pi^2}{6}\Big) = \frac{1}{2}.$$

Then $\mathbb{E}X_0 = 0$ and for $t = n^2/2^n$ and n large, we obtain

$$\mathbb{E}X_0\mathbb{1}_{\{|X_0|\le\frac{1}{t}\}} = \mathbb{E}X_0\mathbb{1}_{\{|X_0|>\frac{1}{t}\}} = \sum_{k=n+1}^{\infty}\frac{1}{2k^2} \ge \frac{1}{2(n+1)}.$$

On the other hand

$$\frac{1}{t}\mathbb{P}\Big(|X_0| > \frac{1}{t}\Big) = \frac{2^n}{n^2}\cdot\frac{1}{2^{n+1}} = \frac{1}{2n^2}$$

and

$$t\,\mathrm{Var}\Big(X_0\mathbb{1}_{\{|X_0|\le\frac{1}{t}\}}\Big) \le \frac{n^2}{2^n}\mathbb{E}|X_0|^2\mathbb{1}_{\{|X_0|\le 2^n/n^2\}} = \frac{n^2}{2^n}\Big(\frac{\pi^4}{72} + \sum_{k=1}^{n}\frac{2^{k-1}}{k^4}\Big) \le \frac{K}{n^2}$$

for some numerical constant K.

This shows that the condition (2.5) is not satisfied and as a consequence $\mathbb{X}$ consisting of i.i.d. copies of X_0 does not satisfy the CDP.

Our last result concerning independent random variables is the following corollary on reversing the triangle inequality in L_0, which should be compared with Lemma A.2 from the Appendix, dealing with L_p spaces for $p \ge 1$. It turns out that in contrast to the L_p case, reversing the triangle inequality at the level of tails requires additional regularity of the distribution of the underlying random variables.

Corollary 2.17. Assume that the variables X_n, $n \in \mathbb{N}$ are i.i.d. and X_0 satisfies (2.5). Then for any $d \in \mathbb{N}$ there exists a constant C_d such that for any separable Banach space $(E, \|\cdot\|)$, any sequence of random variables $Z_i \in Q_i(\mathbb{X}, E)$, $i = 0, \ldots, d$ and any $t > 0$,

$$\sum_{i=0}^{d}\mathbb{P}(\|Z_i\| \ge t) \le C_d\mathbb{P}(\|Z_0 + Z_1 + \ldots + Z_d\| \ge t/C_d). \tag{2.8}$$

Moreover, if (2.8) holds for $E = \mathbb{R}$ and $d = 1$, then X_0 satisfies (2.5).

3 Multiple Poisson integrals

Let us now pass to the Poissonian setting and discuss a counterpart of Theorem 1.2. Since a formal introduction of all the underlying notions is quite lengthy here we will only present the counterexample and the formulation of our theorem, using standard notation from the theory of Poisson processes and Poisson multiple integrals (see, e.g., [13]), postponing the precise definitions to Section 6, which will be devoted solely to the Poissonian case.

Example 3.1. Consider a Poisson process η with uniform intensity on the interval $[0,1]$. Let $f_n = n\mathbb{1}_{[0,\frac{1}{n}]}$ and let $F_n = I_1(f_n) = \int_0^1 f_n d\eta - \int_0^1 f_n dx$ be the compensated Poisson stochastic integral of f_n (in particular F_n is an element of the first Wiener-Poisson chaos, see Section 6 for the definition). Then F_n converges almost surely to -1. Since -1 is an element of the Poisson chaos of order 0, we see that the counterpart of Theorem 1.2 does not hold even for $d = 1$.

On the other hand we have the following result.

Theorem 3.2. *Let η be a Poisson point process on a measurable space $(\mathcal{X},\mathcal{F})$ with a σ-finite intensity measure λ. For $n \geq 1$ let I_n be the corresponding n-fold (compensated) stochastic integral on $L_{2,s}(\mathcal{X}^n,\mathcal{F}^{\otimes n},\lambda^{\otimes n})$ (the space of square integrable functions, symmetric in their arguments). Consider $d \in \mathbb{N}$ and a sequence $(F_n)_{1\leq n<\infty}$ of random variables of the form*

$$F_n = \mathbb{E}F_n + \sum_{k=1}^{d} I_k(f_{n,k}),$$

where $f_{n,k} \in L_{2,s}(\mathcal{X}^k,\mathcal{F}^{\otimes k},\lambda^{\otimes k})$. If the sequence F_n converges almost surely to some random variable F_∞, and there exists an integrable random variable X such that for all n, $|F_n| \leq X$ a.s., then $\mathbb{E}F_n \to \mathbb{E}F_\infty$ and there exist random variables $F_{\infty,k}$, $k = 1,\ldots,d$ such that as $n \to \infty$, $I_k(f_{n,k})$ converges almost surely and in L_1 to $F_{\infty,k}$. If moreover $(F_n)_{1\leq n<\infty}$ is bounded in L_2, then there exist functions $f_{\infty,k} \in L_{2,s}(\mathcal{X}^k,\mathcal{F}^{\otimes k},\lambda^{\otimes k})$, such that for all $1 \leq k \leq d$, $F_{\infty,k} = I_k(f_{n,k})$.

Example 3.3. The assumption that $(F_n)_{n=1}^\infty$ is bounded in L_2 cannot be dropped, i.e., if the other assumptions of the theorem are satisfied, but this one is not, it is possible that the sequence F_n converges almost surely to a random variable which is not in L_2. To see this it is enough to consider $d = 1$, a function $f_\infty \colon \mathcal{X} \to [0,\infty)$, which is integrable but not square integrable and a sequence of functions $f_n \in L_2(\mathcal{X},\mu) \cap L_1(\mathcal{X},\mu)$ converging pointwise to f from below. Setting $F_n = I_1(f_n) = \int_{\mathcal{X}} f_n d\eta - \int_{\mathcal{X}} f_n d\lambda$ one can easily see that F_n converges almost surely to $I_1(f) \notin L_2(\eta)$, moreover $|F_n| \leq \int f_\infty d\eta + \int f_\infty d\lambda \in L_1(\eta)$, so $|F_n|$ is indeed dominated by an integrable random variable.

Remark 3.4. (A note added in revision) In the initial submission we remarked that it was not clear to us whether under the assumption of L_2 boundedness or even under a stronger assumption that F_n converge to F_∞ in L_2, one could drop the assumption of majorization by an integrable random variable. When the article was under review Pratelli and Rigo [23] provided a negative answer to this question. They constructed a sequence of random variables of the form $F_n = I_1(f_{n,1}) + I_2(f_{n,2})$ such that $F_n \to 0$ a.s. and moreover for some $\delta > 0$, $F_n \to 0$ in $L_{2+\delta}$ and $\mathbb{E}\sup_n |F_n|^\delta < \infty$, while $I_1(f_{n,1})$ does not converge almost surely to 0.

4 Further comments

4.1 Overview of the proof

The original proofs of Theorem 1.1 and Theorem 1.2 due to Poly and Zheng are based on the notion of hypercontractivity, which over the years has proved very useful in analysis of polynomials in random variables. Our approach is based on decoupling inequalities, introduced by McConnell and Taqqu in the 1980s [15] and subsequently developed by many authors, in particular by Kwapień [9] in the case of multilinear forms, with the most general result dealing with U-statistics and U-processes obtained by de la Peña and Montgomery-Smith [4] (see Theorem A.1 in Appendix A). This technique reduces the analysis of homogeneous tetrahedral polynomials in a single sequence of independent random variables, to polynomials in multiple copies of this sequence, which are linear in each of the copies (see [5], where general decoupling inequalities for U-statistics have been used for polynomials in a similar way as in the proof of Lemma 5.1 below). This often allows for conditioning and inductive arguments based on the analysis of sequences of independent random variables, which are well understood. The downside of the decoupling approach, when compared with hypercontractivity methods is a typically much worse dependence of constants in the inequalities on the degree of the polynomial. On the other hand decoupling inequalities are more general as they work for all sequences of independent random variables and also in arbitrary Banach spaces, while hypercontractivity depends heavily on the distribution of the underlying sequence and on the Banach space considered. One can note that from a conditional application of the results by Poly and Zheng it follows that all sequences of symmetric random variables satisfy the CDP, while not all of them satisfy hypercontractive estimates (see [10] for a characterization in terms of the distribution). This, and the fact that the CDP is a qualitative and not quantitative statement, suggests that in this case decoupling may work more efficiently than hypercontraction. On the other hand we should mention that Corollary 2.9 may probably follow by hypercontractive estimates that can be recovered from the proofs in [10]. Also, Theorem 3.2 can be proved by means of the Mehler formula for the Poisson process, mimicking the approach Poly and Zheng used in the Gaussian case. We present this alternate argument in Section 6. In terms of notation it is in fact simpler than our main approach based on decoupling, which on the other hand seems to be more easily generalizable to other settings (in particular to more general random measures or U-statistics).

4.2 Organisation of the article

Section 5 is devoted to proofs of results for independent random variables. It is split into Subsection 5.1, where we formulate the main technical lemma and use it to prove Propositions 2.4, 2.5, 2.7, and Subsection 5.2, where we present proofs of Theorem 2.8, Corollary 2.9, Theorem 2.13 and Corollary 2.17. Section 6 contains two proofs of Theorem 3.2 on multiple Poisson integrals. In Appendix A we formulate the main decoupling results that all our proofs are based on, in Appendix B an elementary proof of Proposition 5.5 used in Section 5.2, and in Appendix C we formulate a technical lemma concerning density of simple symmetric functions, used in the proof of Theorem 3.2.

5 Proofs of results for independent random variables

In this section we provide proofs of results concerning sequences of independent random variables, formulated in Section 2. First we will state a technical lemma and demonstrate abstract propositions, then prove the main theorems.

5.1 A technical lemma and proofs of results from Section 2.1

In what follows by $\stackrel{\mathbb{P}}{\to}$ we denote convergence in probability.

Lemma 5.1. *Let $\mathbb{X} = (X_n)_{n=0}^{\infty}$ be a sequence of independent random variables, satisfying the following implication. For all sequences k_n, $n \in \mathbb{N}$, of nonnegative integers and all sequences $a^{(n)} = (a_0^{(n)}, \ldots, a_{k_n}^{(n)}) \in \mathbb{R}^{k_n+1}$, $b_n \in \mathbb{R}$, $n \in \mathbb{N}$,*

$$\Big(b_n + \sum_{k=0}^{k_n} a_k^{(n)} X_k \stackrel{\mathbb{P}, n\to\infty}{\to} 0\Big) \implies \Big(b_n \stackrel{n\to\infty}{\to} 0\Big). \tag{5.1}$$

Then for every separable Banach space $(E, \|\cdot\|)$, every nonnegative integer d and every sequence $(Z_n)_{1\le n\le\infty}$ of random variables of the form

$$Z_n = \sum_{k=0}^{d} Z_{n,k},$$

where $Z_{n,k} \in Q_k(\mathbb{X}, E)$, such that $Z_n \to Z_\infty$ almost surely as $n \to \infty$, we have

$$Z_{n,k} \to Z_{\infty,k} \text{ a.s.}$$

for all $k = 0, \ldots, d$.

Before we present the proof of the above lemma, let us make a few comments and describe its basic consequences. In particular we will prove Propositions 2.4, 2.5 and 2.7.

Remark 5.2. Since k_n in the above lemma may be an arbitrary sequence of nonnegative integers, it is easy to see that its assumption can be equivalently stated as

$$\text{There does not exist a sequence } Z_n \in Q_1(\mathbb{X}) \text{ such that } Z_n \overset{\mathbb{P}, n\to\infty}{\to} 1.$$

Let us also make the following remark, which will be used in the proof of Lemma 5.1.

Remark 5.3. Consider the implication (5.1) with convergence in probability replaced by almost sure convergence. It is easy to see that this formally weaker property of the sequence $\mathbb{X}$ is in fact equivalent to (5.1). Indeed, assume that such a weaker version holds and consider any $a^{(n)}, b_n$ such that

$$b_n + \sum_{k=0}^{k_n} a_k^{(n)} X_k \overset{\mathbb{P}}{\to} 0.$$

For every increasing sequence n_m of nonnegative integers we can find a subsequence n_{m_l} such that

$$b_{n_{m_l}} + \sum_{k=0}^{k_{n_{m_l}}} a_k^{(n)} X_k \overset{a.s., l\to\infty}{\to} 0,$$

which implies that $b_{n_{m_l}} \to 0$ as $l \to \infty$. Therefore, $b_n \to 0$ as $n \to \infty$, which proves (5.1).

The above remark and Lemma 5.1 immediately implies the following corollary.

Corollary 5.4. A sequence $\mathbb{X}$ of independent random variables satisfies the CDP if and only if it satisfies the implication (5.1).

Having the above corollary we can easily prove Propositions 2.4 and 2.5.

Proof of Proposition 2.4. If $\mathbb{X}$ satisfies the CDP, then by Corollary 5.4 it satisfies the implication (5.1). The assertion of the proposition follows thus by Lemma 5.1. □

Proof of Proposition 2.5. The necessity of the uniqueness of the decomposition is obvious. To show that it is also sufficient for the CDP, note that by Lemma 5.1 if this property does not hold, then we can find a sequence of linear forms in the variables X_i converging in probability to 1. Thus $1 \in Q_1(\mathbb{X})$. Since obviously $1 \in Q_0(\mathbb{X})$, this shows that there is no uniqueness of the decomposition for $d = 1$. □

Finally let us demonstrate Proposition 2.7.

Proof of Proposition 2.7. It is enough to prove the existence of the variables $Z_{\infty,k}$. Consider thus any strictly increasing sequences l_n, m_n of integers. Since Z_n converges in probability, the difference $S_n := Z_{l_n} - Z_{m_n}$ converges in probability to zero. Thus from an arbitrary subsequence of S_n we can select a further subsequence along which the almost sure convergence holds. Using the CDP, we obtain that along this subsequence also the homogeneous parts of S_n tend almost surely to zero. Thus from every subsequence of $S_{n,k} := Z_{l_n,k} - Z_{m_n,k}$ one can select a further subsequence converging almost surely to zero, which implies that $S_{n,k}$ converges to zero in probability. But, as the sequences l_n, m_n were arbitrary, this implies that the Cauchy condition for convergence in probability is satisfied, and so by the completeness of $L_0(E)$, $Z_{n,k}$ converges in probability to some random variable $Z_{\infty,k}$. Clearly we have then $Z_\infty = \sum_{k=0}^{d} Z_{\infty,k}$. □

We will now pass to the proof of Lemma 5.1.

Proof of Lemma 5.1. We will use the notation as in Definition 2.3. Clearly it is enough to consider the case $Z_{\infty,k} = 0$ for all $k \le d$. Also, we can assume that $Z_{n,k}$ are multilinear tetrahedral forms in a finite number of variables X_i, i.e.,

$$Z_{n,k} = \sum_{i_1,\ldots,i_k=0}^{\infty} a^{(n,k)}_{i_1,\ldots,i_k} X_{i_1} \cdots X_{i_k} \tag{5.2}$$

where $a^{(n,k)} \in \ell_0^{\odot n}(\mathbb{N})$ and there exist $m_{n,k} < \infty$ such that $a^{(n,k)}_{i_1,\ldots,i_k} = 0$ if $\max(i_1,\ldots,i_k) > m_{n,k}$.

Indeed, by the Borel-Cantelli lemma we can find $\widetilde{Z}_{n,k}$, $0 \le k \le d, n \ge 0$, being such tetrahedral forms, such that with probability one for all $k \le d$, $\widetilde{Z}_{n,k} - Z_{n,k} \to 0$ as $n \to \infty$. In particular $\sum_{k=0}^{d} \widetilde{Z}_{n,k} \xrightarrow{a.s.} 0$ and for all k, $Z_{n,k} \xrightarrow{a.s.} 0$ iff $\widetilde{Z}_{n,k} \xrightarrow{a.s.} 0$. For the purpose of the proof we can thus assume that $\widetilde{Z}_{n,k} = Z_{n,k}$.

We will now prove by induction on $d \ge 1$ that for *any sequence* $\mathbb{X}$, satisfying (5.1), if $Z_{n,k}$, $k \le d$, are as in (5.2) and $Z_n = \sum_{k=0}^{d} Z_{n,k} \xrightarrow{a.s.} 0$, then for all $k \le d$, $Z_{n,k} \xrightarrow{a.s.} 0$.

The base of induction: $d = 1$ In this case we have $Z_n = Z_{n,0} + Z_{n,1}$, where $Z_{n,0} \in E$ is deterministic and $Z_{n,1} = \sum_{k=0}^{k_n} a_k^{(n,1)} X_k$ for some $a_k^{(n,1)} \in E$. By the Hahn-Banach theorem there exist norm one linear functionals φ_n on E such that $\varphi_n(Z_{n,0}) = \|Z_{n,0}\|$. If $Z_n \xrightarrow{a.s.} 0$, then $\varphi_n(Z_n) = \|Z_{n,0}\| + \sum_{k=0}^{k_n} \varphi_n(a_k^{(n,1)}) X_k \xrightarrow{a.s.} 0$. By assumption (5.1) this implies that $\|Z_{n,0}\| \to 0$, which clearly yields $Z_{n,1} \xrightarrow{a.s.} 0$.

The induction step Let us assume that the induction hypothesis holds for all numbers smaller than d. Note that by the convergence $Z_n \xrightarrow{a.s.} 0$, we have for arbitrary $\varepsilon > 0$,

$$\lim_{n\to\infty} \sup_{m \ge n} \mathbb{P}(\max_{n \le l \le m} \|Z_l\| > \varepsilon) = 0. \tag{5.3}$$

For $l \in \mathbb{N}$ define the functions $h^{(l)}_{i_1,\ldots,i_d} : \mathbb{R}^d \to E$ by the formula

$$h^{(l)}_{i_1,\ldots,i_d}(x_1,\ldots,x_d) = \sum_{k=0}^{d} \frac{(d-k)!}{d!} \frac{(N-d)!}{(N-k)!} \sum_{1 \le r_1 \ne \ldots \ne r_k \le d} a^{(l,k)}_{i_{r_1},\ldots,i_{r_k}} x_{r_1} \cdots x_{r_k}$$

and note that the random vector $(Z_n,\ldots,Z_m)$ can be written as

$$\sum_{1 \le i_1 \ne i_2 \ne \ldots \ne i_d \le N} \left(h^{(l)}_{i_1,\ldots,i_d}(X_{i_1} \ldots, X_{i_d}) \right)_{l=n}^{m}$$

where $N = \max_{n \le l \le m} \max_{0 \le k \le d} m_{l,k}$ (here and in what follows the notation $i_1 \ne \ldots \ne i_d$ denotes the condition that the indices i_j are *pairwise* distinct).

Let now $\mathbb{X}^{(i)} = (X_n^{(i)})_{n=0}^{\infty}$, $i \in [d]$ be i.i.d. copies of the sequence $\mathbb{X}$ and define Z_l^{dec}, the decoupled version of Z_l as

$$
\begin{aligned}
Z_l^{dec} &:= \sum_{1 \le i_1 \neq i_2 \neq \ldots \neq i_d \le N} h^{(l)}_{i_1,\ldots,i_d}(X^{(1)}_{i_1} \ldots, X^{(d)}_{i_d}) \\
&= \sum_{k=0}^{d} \frac{1}{\binom{d}{k}} \sum_{1 \le r_1 < \ldots < r_k \le d} \sum_{i_1,\ldots,i_k=1}^{\infty} a^{(l,k)}_{i_1,\ldots,i_k} X^{(r_1)}_{i_1} \cdots X^{(r_k)}_{i_k},
\end{aligned}
$$

where the equality follows from the symmetry of the tensors $a^{(l,k)}$, and the fact that they have vanishing diagonals and finite support.

We have $h^{(l)}_{i_{\pi(1)},\ldots,i_{\pi(d)}}(x_{\pi(1)},\ldots,x_{\pi(d)}) = h^{(l)}_{i_1,\ldots,i_d}(x_{i_1},\ldots,x_{i_d})$ for every permutation π of the set $[d]$, so by Theorem A.1 from the Appendix, applied to the space $F = E^{\{n,\ldots,m\}}$ equipped with the norm $\|(x_n,\ldots,x_m)\| = \max_{n \le l \le m} \|x_i\|$, we have for every $\varepsilon > 0$,

$$
\frac{1}{C_d}\mathbb{P}(\max_{n \le l \le m} \|Z_l\| \ge C_d \varepsilon) \le \mathbb{P}(\max_{n \le l \le m} \|Z_l^{dec}\| \ge \varepsilon) \le C_d \mathbb{P}(\max_{n \le l \le m} \|Z_l\| \ge \varepsilon / C_d)
$$

Combining this with (5.3) we obtain that $Z_n^{(dec)} \overset{a.s.}{\to} 0$ as $n \to \infty$.

Consider a sequence $\mathbb{Y} = (Y_n)_{n=0}^{\infty}$ defined as

$$
Y_{kd+r} = X_k^{(r+1)}
$$

for $k \in \mathbb{N}$, $r \in \{0,\ldots,d-1\}$, and any sequences $a^{(n)} = (a_0^{(n)},\ldots,a_{k_n}^{(n)}) \in \mathbb{R}^{k_n+1}$, $b_n \in \mathbb{R}$, $n \in \mathbb{N}$, such that $b_n + \sum_{k=0}^{k_n} a_k^{(n)} Y_k \to 0$ almost surely. Using the Fubini theorem and applying successively (5.1) to $X^{(r)}$ ($r \in [d]$) conditionally on $X^{(l)}$, $l \in [d] \setminus \{r\}$, we easily obtain that $b_n \to 0$ as $n \to \infty$. Taking into account Remark 5.3 we can infer that $\mathbb{Y}$ satisfies the implication (5.1).

Now, for fixed $r \in [d]$, applying this implication to $(X_n^{(r)})_{n \in \mathbb{N}}$ and Z_n^{dec}, conditionally on $\{X_i^{(r)}\}_{n \in \mathbb{N}, r \in [d] \setminus \{r\}}$ we obtain via the Fubini theorem, that

$$
\sum_{k=0}^{d-1} \frac{1}{\binom{d}{k}} \sum_{\substack{1 \le r_1 < \ldots < r_k \le d \\ \forall_i r_i \neq r}} \sum_{i_1,\ldots,i_k=1}^{\infty} a^{(n,k)}_{i_1,\ldots,i_k} X^{(r_1)}_{i_1} \cdots X^{(r_k)}_{i_k} \overset{a.s.}{\to} 0.
$$

The induction hypothesis applied to $\mathbb{Y}$ implies now that for each $r \in [d]$ and each $k \le d-1$

$$
\sum_{\substack{1 \le r_1 < \ldots < r_k \le d \\ \forall_i r_i \neq r}} \sum_{i_1,\ldots,i_k=1}^{\infty} a^{(n,k)}_{i_1,\ldots,i_k} X^{(r_1)}_{i_1} \cdots X^{(r_k)}_{i_k} \overset{a.s.}{\to} 0
$$

(we use here that every tetrahedral homogeneous polynomial can be represented in the form (5.2)).

Now we get

$$
\begin{aligned}
&\sum_{1 \le r_1 < \ldots < r_k \le d} \sum_{i_1,\ldots,i_k=1}^{\infty} a^{(n,k)}_{i_1,\ldots,i_k} X^{(r_1)}_{i_1} \cdots X^{(r_k)}_{i_k} \\
&\qquad = \frac{1}{d-k} \sum_{r=1}^{d} \sum_{\substack{1 \le r_1 < \ldots < r_k \le d \\ \forall_i r_i \neq r}} \sum_{i_1,\ldots,i_k=1}^{\infty} a^{(n,k)}_{i_1,\ldots,i_k} X^{(r_1)}_{i_1} \cdots X^{(r_k)}_{i_k} \overset{a.s.}{\to} 0,
\end{aligned}
$$

for all $k \le d-1$ and as a consequence

$$
Z_{n,d}^{dec} := \sum_{i_1,\ldots,i_d=1}^{\infty} a^{(n,d)}_{i_1,\ldots,i_d} X^{(1)}_{i_1} \cdots X^{(d)}_{i_d} \overset{a.s.}{\to} 0 \,.
$$

Applying now again the decoupling inequality, we conclude that for all $\varepsilon > 0$,

$$\limsup_{n\to\infty} \sup_{m\geq n} \mathbb{P}(\max_{n\leq l\leq m} \|Z_{l,d}\| > \varepsilon) \leq C \lim_{n\to\infty} \sup_{m\geq n} \mathbb{P}(\max_{n\leq l\leq m} \|Z_{l,d}^{dec}\| > \varepsilon/C) = 0,$$

i.e., $Z_{n,d} \xrightarrow{a.s.} 0$. From this we obtain $\sum_{k=0}^{d-1} Z_{n,k} \xrightarrow{a.s.} 0$, which by another application of the induction hypothesis implies that $Z_{n,k} \xrightarrow{a.s.} 0$ also for all $k < d$, and ends the induction step. □

5.2 Proofs of results from Section 2.2

We will use the following proposition, characterizing convergence in probability to a constant for row sums of a triangular array of independent random variables. Let us remark that with some not difficult but technical calculations it can be obtained from a much deeper result, namely [21, Chapter IV, Theorem 3], characterizing weak convergence of such sums to an arbitrary infinitely divisible distribution. However, to make the presentation more self contained and elementary, we provide a direct proof in Appendix B.

Proposition 5.5. Let $X_{n,k}$, $n \in \mathbb{N}, k \in \{0,\ldots,k_n\}$ be a triangular array of random variables such that for each n, $X_{n,0},\ldots,X_{n,k_n}$ are independent. Assume that for all $\varepsilon > 0$,

$$\max_{0\leq k\leq k_n} \mathbb{P}(|X_{n,k}| \geq \varepsilon) \to 0 \tag{5.4}$$

as $n \to \infty$. Let τ be an arbitrary positive number. Then $\sum_{k=0}^{k_n} X_{n,k}$ converges in probability to 1 if and only if

(i)

$$\sum_{k=0}^{k_n} \mathbb{E} X_{n,k} \mathbb{1}_{\{|X_{n,k}|\leq\tau\}} \to 1 \tag{5.5}$$

and

(ii)

$$\sum_{k=0}^{k_n} \Big(\mathbb{P}(|X_{n,k}| > \tau) + \mathrm{Var}\Big(X_{n,k}\mathbb{1}_{\{|X_{n,k}|\leq\tau\}}\Big)\Big) \to 0 \tag{5.6}$$

as $n \to \infty$.

We are now ready for the proof of Theorem 2.8.

Proof of Theorem 2.8. By Lemma 5.1 it is enough to verify that under the assumptions of Theorem 2.8 the implication (5.1) holds. We will proceed by contradiction. Assume thus that there are sequences $a^{(n)} = (a_0^{(n)},\ldots,a_{k_n}^{(n)}) \in \mathbb{R}^{k_n+1}$, $b_n \in \mathbb{R}$, $n \in \mathbb{N}$, such that $b_n + \sum_{k=0}^{k_n} a_k^{(n)} X_k \xrightarrow{\mathbb{P}, n\to\infty} 0$ but b_n does not converge to 0. By passing to a subsequence we can further assume that b_n's are separated from zero. Dividing by b_n and setting $t_{n,k} = -a_{n,k}/b_n$ we thus obtain a sequence $t_n = (t_{n,0},\ldots,t_{n,k_n})$ such that

$$\sum_{k=0}^{k_n} t_{n,k} X_k \xrightarrow{\mathbb{P}} 1.$$

Let $\mathbb{X}' = (X_n')_{n=0}^{\infty}$ be an independent copy of $\mathbb{X}$. We have

$$\sum_{k=0}^{k_n} t_{n,k}(X_k - X_k') \xrightarrow{\mathbb{P}} 0.$$

By assumption (2.3) and the Fubini Theorem we obtain for all k,

$$\mathbb{P}(|X_k - X_k'| \geq \delta) \geq \delta.$$

On the other hand, by symmetry of $X_k - X_k'$, for any $\varepsilon > 0$,

$$\mathbb{P}(|t_{n,k}| \cdot |X_k - X_k'| \geq \varepsilon) \leq 2\mathbb{P}\Big(\Big|\sum_{k=0}^{k_n} t_{n,k}(X_k - X_k')\Big| \geq \varepsilon\Big) \to 0,$$

which shows that $t_{n,k}$ converge with n to zero, uniformly in $k \in \mathbb{N}$. Together with tightness this implies that the triangular array given by $X_{n,k} = t_{n,k}X_k$ satisfies (5.4). As a consequence, by Proposition 5.5 we obtain

$$A_n := \sum_{\substack{1\leq k\leq k_n\\ t_{n,k}\neq 0}} t_{n,k}\mathbb{E}X_k\mathbb{1}_{\{|X_k|\leq \frac{1}{|t_{n,k}|}\}} \to 1$$

and

$$B_n := \sum_{\substack{1\leq k\leq k_n\\ t_{n,k}\neq 0}} \Big(\mathbb{P}\Big(|X_k| > \frac{1}{|t_{n,k}|}\Big) + t_{n,k}^2 \operatorname{Var}\Big(X_k\mathbb{1}_{\{|X_{n,k}|\leq \frac{1}{|t_{n,k}|}\}}\Big)\Big) \to 0,$$

which is however impossible, since by (2.4), for n large enough, we have $|A_n| \leq CB_n$. This ends the proof of the theorem. □

Let us now prove Corollary 2.9.

Proof of Corollary 2.9. Let us first prove the condition (2.4). Let t_0 be such that for all $n \in \mathbb{N}$, $\mathbb{E}X_n^2\mathbb{1}_{\{|X_n|\leq \frac{1}{t_0}\}} \geq 1/2$. Using the mean zero assumption, for $t \leq t_0$ we can estimate

$$\begin{aligned}|\mathbb{E}X_n\mathbb{1}_{\{|X_n|\leq\frac{1}{t}\}}| &= |\mathbb{E}X_n\mathbb{1}_{\{|X_n|>\frac{1}{t}\}}| \leq t\mathbb{E}X_n^2 = t\\ &\leq 2t\mathbb{E}X_n^2\mathbb{1}_{\{|X_n|\leq\frac{1}{t}\}} = 2t\operatorname{Var}\Big(X_n\mathbb{1}_{\{|X_n|\leq\frac{1}{t}\}}\Big) + 2t\Big(\mathbb{E}X_n\mathbb{1}_{\{|X_n|>\frac{1}{t}\}}\Big)^2.\end{aligned}$$

Now, by the Schwarz inequality, the second term on the right-hand side above is bounded by

$$2t\mathbb{E}X_n^2\mathbb{P}\Big(|X_n| > \frac{1}{t}\Big) \leq 2t_0^2\frac{1}{t}\mathbb{P}\Big(|X_n| > \frac{1}{t}\Big),$$

which shows that (2.4) holds with $C = \max(2, 2t_0^2)$.

Tightness of the sequence $\mathbb{X}$ follows from uniform integrability, so to finish the proof it remains to demonstrate the condition (2.3). This will follow by uniform integrability and a Paley-Zygmund type argument.

By the de la Vallée Poussin theorem, there exists a convex, nondecreasing function $\varphi\colon [0,\infty) \to [0,\infty)$ such that $\varphi(0) = 0$, $\lim_{x\to\infty}\varphi(x)/x = \infty$ and $M := \sup_{n\in\mathbb{N}}\mathbb{E}\varphi(|X_n|^2) < \infty$. Consider any $x \in [-2,2]$. By convexity

$$\mathbb{E}\varphi\Big(\frac{(X_n - x)^2}{4}\Big) \leq \frac{1}{2}(\mathbb{E}\varphi(|X_n|^2) + \varphi(4)) \leq \frac{1}{2}(M + \varphi(4))$$

Therefore, again by convexity, there exists K such that for all $n \in \mathbb{N}$ and $x \in [-2,2]$,

$$\mathbb{E}\varphi\Big(\frac{(X_n - x)^2}{K}\Big) \leq \frac{1}{4}.$$

On the other hand $\mathbb{E}(X_n - x)^2 = \mathbb{E}X_n^2 + x^2 \geq 1$. Denoting by φ^* the Legendre transform of φ, given by the formula $\varphi^*(x) = \sup_{y\geq 0}(xy - \varphi(y))$, we can estimate

$$\frac{3}{4} \leq \mathbb{E}(X_n - x)^2\mathbb{1}_{\{|X_n - x|>\frac{1}{2}\}} \leq \mathbb{E}\varphi\Big(\frac{(X_n - x)^2}{K}\Big) + \varphi^*(K)\mathbb{P}\Big(|X_n - x| \geq \frac{1}{2}\Big),$$

which together with the definition of K yields

$$\mathbb{P}\Big(|X_n - x| \geq \frac{1}{2}\Big) \geq \frac{1}{2\varphi^*(K)}$$

for all $x \in [-2, 2]$. For $|x| > 2$, by Chebyshev's inequality we have $\mathbb{P}(|X - x| \leq 1/2) \leq \mathbb{P}(|X| \geq 3/2) \leq \frac{4}{9}$.

Combining the last two estimates we obtain (2.3) with $\delta = 2^{-1}\min(1, 1/\varphi^*(K))$, which ends the proof of the corollary. $\square$

We will conclude this section by proving the characterizations of the CDP (Theorem 2.13) and the corresponding reverse triangle inequality (Corollary 2.17) in the i.i.d. case.

Proof of Theorem 2.13. One can easily check that the equivalence between the CDP and condition (2.5) holds in the case of almost surely constant variable X_0 (both conditions are satisfied if and only if X_0 vanishes almost surely), therefore from now on we will assume that X_0 is not deterministic. We will first prove that (2.5) implies the CDP. To this end we will use Theorem 2.8. The condition (2.4) in the i.i.d. case clearly reduces to (2.5), tightness of $\mathbb{X}$ is obvious, and the condition (2.3) follows easily from the assumption that X_0 is not deterministic. Indeed, for any pair of sequences $x_n \in \mathbb{R}$ and $\delta_n \to 0$, such that $\mathbb{P}(X_0 \in (x_n - \delta_n, x_n + \delta_n)) \geq 1 - \delta_n$, the sequence x_n must be bounded, and thus passing to a convergent subsequence we would obtain that X_0 is deterministic. Thus, as all the assumptions of Theorem 2.8 hold, we can conclude that $\mathbb{X}$ satisfies the CDP.

Let us now prove the converse implication. Assume that (2.5) is not satisfied. Thus there exists a sequence of positive numbers t_n, such that $t_n \to 0$ and

$$t_n\Big|\mathbb{E}X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big| > n\Big(\mathbb{P}\Big(|X_0| > \frac{1}{t_n}\Big) + t_n^2 \operatorname{Var}\Big(X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big)\Big).$$

Set $a_n = t_n$ if $\mathbb{E}X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}} > 0$ and $a_n = -t_n$ otherwise. Define moreover

$$k_n = \Big\lfloor \Big(t_n\Big|\mathbb{E}X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big|\Big)^{-1}\Big\rfloor - 1.$$

Note that by the Lebesgue dominated convergence theorem

$$t_n\Big|\mathbb{E}X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big| \to 0 \tag{5.7}$$

and so $k_n \to \infty$.

Now consider the sequence $Z_n = \sum_{k=0}^{k_n} a_n X_k = \sum_{k=0}^{k_n} X_{n,k}$, where $X_{n,k} = a_n X_k$. Since $a_n \to 0$ and X_n have the same distribution, the condition (5.4) is satisfied. Using the definition of k_n and (5.7) we get

$$\sum_{k=0}^{k_n} \mathbb{E}X_{n,k}\mathbb{1}_{\{|X_{n,k}|\leq 1\}} = (k_n + 1)t_n\Big|\mathbb{E}X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big| \to 1$$

as $n \to \infty$, which yields (5.5) of Proposition 5.5.

Moreover,

$$\begin{aligned}\sum_{k=0}^{k_n} &\Big(\mathbb{P}(|X_{n,k}| > 1) + \operatorname{Var}(X_{n,k}\mathbb{1}_{\{|X_{n,k}|\leq 1\}})\Big)\\ &= (k_n + 1)\Big(\mathbb{P}\Big(|X_0| > \frac{1}{t_n}\Big) + t_n^2 \operatorname{Var}\Big(X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big)\Big)\\ &\leq \frac{\mathbb{P}\Big(|X_0| > \frac{1}{t_n}\Big) + t_n^2 \operatorname{Var}\Big(X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big)}{t_n\Big|\mathbb{E}X_0\mathbb{1}_{\{|X_0|\leq \frac{1}{t_n}\}}\Big|} < \frac{1}{n}\end{aligned}$$

and so (5.6) is also satisfied.

Thus, by Proposition 5.5 we obtain that $\sum_{k=0}^{k_n} a_n X_n \to 1$ in probability. Passing to a subsequence we can upgrade this to the almost sure convergence, which implies that the CDP cannot hold. □

Proof of Corollary 2.17. To prove the first part of the corollary we will proceed by contradiction, constructing a sequence of polynomials with coefficients in c_0, which violate the assertion of Proposition 2.4. Let us thus assume that (2.5) holds but (2.8) is violated. Then there exist d, $k \le d$, a sequence of Banach spaces E_n, $t_n > 0$ and $Z_{n,i} \in Q_i(\mathbb{X}, E_n)$ ($i \le d$), such that

$$\mathbb{P}(\|Z_{n,k}\| \ge 2t_n) > 4n^2 \mathbb{P}\Big(\|Z_{n,0} + \ldots + Z_{n,d}\| \ge \frac{t_n}{2n^2}\Big),$$

(for simplicity we will denote all the norms appearing in the proof by $\|\cdot\|$). Scaling $Z_{n,k}$ if necessary, we can assume that $t_n = 1$. By approximation we obtain homogeneous tetrahedral forms (in particular depending on a finite number of variables) $Z'_{n,i}$ of degree i ($i \le d$), such that

$$\mathbb{P}(\|Z'_{n,k}\| \ge 3/2) > 4n^2 \mathbb{P}\Big(\|Z'_{n,0} + \ldots + Z'_{n,d}\| \ge \frac{2}{3n^2}\Big).$$

Passing to subspaces spanned by coefficients of $Z'_{n,i}$ we may further assume that all spaces E_n are finite dimensional, which by a standard embedding gives a sequence N_n of positive integers, and tetrahedral forms $Z''_{n,k}$ with values in $\ell_\infty^{N_n}$ such that

$$\mathbb{P}(\|Z''_{n,k}\| \ge 1) > n^2 \mathbb{P}\Big(\|Z''_{n,0} + \ldots + Z''_{n,d}\| \ge \frac{1}{n^2}\Big).$$

Let now $m_n = \lceil 1/\mathbb{P}(\|Z''_{n,k}\| \ge 1)\rceil$. Since $Z''_{n,i}$ depend only on finitely many variables X_n, using the sequence $\mathbb{X}$ we can construct i.i.d. copies $(Z''_{n,1}(j), \ldots, Z''_{n,d}(j))$, $j = 1, \ldots, m_n$ of the vectors $(Z''_{n,1}, \ldots, Z''_{n,d})$. Then $\widehat{Z}_{n,i} := (Z''_{n,i}(j))_{j=1}^{m_n}$ may be considered a tetrahedral homogeneous polynomial of degree i with coefficients in $\ell_\infty^{N_n m_n}$ embedded in c_0 in a natural way. Recall also the following elementary inequality for independent random variables ξ_i:

$$\frac{1}{2}\min\Big(\sum_j \mathbb{P}(\xi_j > t), 1\Big) \le \mathbb{P}(\max_j \xi_j > t) \le \sum_j \mathbb{P}(\xi_j > t).$$

Using this inequality together with independence over $j = 1, \ldots, m_n$ we obtain

$$\mathbb{P}(\|\widehat{Z}_{n,k}\| \ge 1) = \mathbb{P}(\max_{j\le m_n} \|Z''_{n,k}(j)\| \ge 1) \ge \frac{1}{2}\min\Big(m_n \mathbb{P}(\|Z''_{n,k}\| \ge 1), 1\Big) = 1/2 \tag{5.8}$$

and

$$\begin{aligned}\mathbb{P}\Big(\|\widehat{Z}_{n,0} + \ldots + \widehat{Z}_{n,d}\| \ge \frac{1}{n^2}\Big) &= \mathbb{P}\Big(\max_{j\le m_n} \|Z''_{n,0}(j) + \ldots + Z''_{n,d}(j)\| \ge \frac{1}{n^2}\Big)\\ &\le m_n \mathbb{P}\Big(\|Z''_{n,0} + \ldots + Z''_{n,d}\| \ge \frac{1}{n^2}\Big)\\ &\le m_n \frac{1}{n^2}\mathbb{P}(\|Z''_{n,k}\| \ge 1) \le \frac{2}{n^2}.\end{aligned}$$

Thus, by the Borel-Cantelli lemma, the sequence $\widehat{Z}_n = \widehat{Z}_{n,0} + \ldots + \widehat{Z}_{n,d}$ of c_0 valued tetrahedral polynomials converges almost surely to 0, while by (5.8), $\widehat{Z}_{n,k}$ does not. By Proposition 2.4 this shows that $\mathbb{X}$ does not have the CDP, which by Theorem 2.13 contradicts (2.5) and finishes the proof of the first part of the corollary.

As for the second part, if (2.8) is satisfied for $d = 1$ and $E = \mathbb{R}$, then clearly the implication (5.1) holds and thus, by Lemma 5.1, $\mathbb{X}$ satisfies the CDP. By Theorem 2.13 this implies that (2.5) is satisfied. □

6 Proofs of results for Poisson stochastic integrals

In this section we will prove Theorem 3.2. The basic proof we will provide is again based on decoupling inequalities. After completing the argument we will also present an alternate proof based on Mehler's formula for the Poisson process. We choose to focus on the decoupling proof since it is a variation on the approach we used for independent random variables and also it seems that its adaptation to more general situations (i.e., other random measures) is more straightforward than in the case of Mehler's formula argument.

Let us start by recalling the basic definitions of multiple Wiener-Itô integrals with respect to the Poisson process. Clearly we are not able to provide here a complete exposition, so we will just present the basic formulas and constructions necessary for carrying out the proof, and refer the reader to the monograph [13] for details.

Consider a measurable space $(\mathcal{X},\mathcal{F})$ with a σ-finite intensity measure λ. In what follows we regard point processes on $(\mathcal{X},\mathcal{F})$ as random elements of the space $\mathbf{N}(\mathcal{X})$ of $\mathbb{N}\cup\{\infty\}$-valued measures on $(\mathcal{X},\mathcal{F})$, which can be written as countable sums of $\mathbb{N}$-valued measures. The measurable structure on $\mathbf{N}(\mathcal{X})$ is given by the smallest σ-field for which all maps $\mu\mapsto\mu(A)$ for $A\in\mathcal{F}$, are measurable. A point process η is a Poisson process with intensity measure λ if

(i) for every $A\in\mathcal{F}$, the random variable $\eta(A)$ has Poisson distribution with parameter $\lambda(A)$ (which we interpret as the Dirac mass at $\lambda(A)$ if $\lambda(A)\in\{0,\infty\}$),

(ii) for every positive integer m and all pairwise disjoint sets $A_1,\ldots,A_m\in\mathcal{F}$, the random variables $\eta(A_1),\ldots,\eta(A_m)$ are jointly independent.

The multiple Wiener-Itô integral $I_n\colon L_{2,s}(\mathcal{X}^n,\mathcal{F}^{\otimes n},\lambda^{\otimes n})\to L_2(\Omega,\mathbb{P})$ is defined first for integrable f with an explicit formula (6.1) below and then uniquely extended to the space $L_{2,s}(\mathcal{X}^n,\mathcal{F}^{\otimes n},\lambda^{\otimes n})$, by a standard density argument, in such a way that $I_n/\sqrt{n!}$ is an isometric embedding. For $f\colon\mathcal{X}^n\to\mathbb{R}$, integrable (not necessarily symmetric or square integrable) one defines

$$I_n(f)=\sum_{J\subset[n]}(-1)^{n-|J|}\int_{\mathcal{X}^{|J^c|}}\int_{\mathcal{X}^{|J|}}f(x_1,\ldots,x_n)\eta^{(|J|)}(dx_J)\lambda^{n-|J|}(dx_{J^c}) \tag{6.1}$$

where $x_J=(x_i)_{i\in J}$, and $\eta^{(m)}$ is the m-th factorial measure on $\mathcal{X}^m$, defined inductively by $\eta^{(1)}=\eta$,

$$\begin{aligned}&\eta^{(m+1)}(\cdot)\\&=\int_{\mathcal{X}^m}\Big(\int_{\mathcal{X}}\mathbb{1}_{\{(x_1,x_2,\ldots,x_{m+1})\in\cdot\}}\eta(dx_{m+1})-\sum_{i=1}^m\mathbb{1}_{\{(x_1,x_2,\ldots,x_m,x_i)\in\cdot\}}\Big)\eta^{(m)}(d(x_1,\ldots,x_m)).\end{aligned}$$

If η is a proper point process, i.e., if η can be represented as a countable sum of Dirac's deltas $\eta=\sum_{i=1}^{\kappa}\delta_{X_i}$ for some $\mathbb{N}\cup\{\infty\}$-valued random variable κ and $\mathcal{X}$-valued random variables X_i, then

$$\mu^{(m)}=\sum_{i_1,\ldots,i_m=1}^{\kappa}\mathbb{1}_{\{i_1\neq\ldots\neq i_m\}}\delta_{(X_{i_1},\ldots,X_{i_k})}.$$

In particular, if $B_1\ldots,B_n\subset\mathcal{X}$ are pairwise disjoint with $\lambda(B_i)<\infty$, and $B=B_1\times\ldots\times B_n$, then $I_n(\mathbb{1}_B)=\prod_{i=1}^n(\eta(B_i)-\lambda(B_i))$. One also proves that $I_n(g)$ and $I_m(f)$ are uncorrelated for $n\neq m$. The subspace of $L_2(\Omega)$ consisting of all m-fold stochastic integrals of square integrable symmetric functions in m variables is called the *m-th Wiener-Poisson chaos*. The chaos representation property asserts that these spaces form an orthogonal decomposition of the space of square integrable random variables measurable with respect to η, which we will denote by $L_2(\eta)$ (see (6.5) below for an explicit formula).

Proof of Theorem 3.2. For the proof of Theorem 3.2 it will be convenient to assume that the measure λ is non-atomic, i.e., that for every $A \in \mathcal{F}$ with $\lambda(A) > 0$, there exists $B \in \mathcal{F}$ with $B \subset A$ and $0 < \lambda(B) < \lambda(A)$. We can do it without loss of generality, since we can replace $\mathcal{X}$ with $\mathcal{X}' = \mathcal{X} \times (0,1)$, λ with $\lambda' = \lambda \otimes Leb$ (where Leb is the Lebesgue measure on the interval). The new measure is non-atomic, moreover if η' is a Poisson process on $\mathcal{X}'$, then by the Mapping Theorem (see [13, Theorem 5.1]), the image of η' under the natural projection $\pi\colon \mathcal{X}' \to \mathcal{X}$ has the same distribution as η. We may thus replace $f_{n,k}$ by $f_{n,k} \circ \pi^k$, where π^k is the corresponding natural projection from $(\mathcal{X} \times (0,1))^k$ onto $\mathcal{X}^k$ and one can then check that the joint distribution of all the stochastic integrals involved remains unchanged. It is thus indeed enough to prove Theorem 3.2 under the additional assumption that λ is non-atomic. We remark that this construction can be carried out for an arbitrary σ-finite measure λ, it does not require additional regularity properties. If η is a proper point process, then probabilistically it can be interpreted as a marking of η with a constant marking kernel given by the Lebesgue measure on the interval (see [13, Chapter 5]).

Given a square $\lambda^{\otimes k}$-integrable symmetric function $f\colon \mathcal{X}^k \to \mathbb{R}$, by Lemma C.1, we can approximate it in L_2 by a function h of the form

$$h = \sum_{i_1,\ldots,i_k=1}^{N} a_{i_1,\ldots,i_k} \mathbb{1}_{A_{i_1}\times\ldots\times A_{i_k}}, \tag{6.2}$$

where the sets $A_1, \ldots, A_N$ are pairwise disjoint subsets of $\mathcal{X}$ with $\lambda(A_i) < \infty$, the coefficients $a_{i_1,\ldots,i_k}$ are symmetric and vanish if $i_l = i_m$ for some $l \neq m$. Note also that if we have a finite family of functions of this form (perhaps with different k's and N's), we can always find their representations with the same sets $A_1, \ldots, A_N$ (first one enlarges the corresponding sequences of sets to have the same union, then one takes all possible intersections).

In the setting of Theorem 3.2, we can thus find functions $g_{n,k} \in L_{2,s}(\mathcal{X}^k, \mathcal{F}^{\otimes k}, \lambda^k)$, $k = 1, \ldots, n$, such that as $n \to \infty$,

$$\sum_{n=0}^{\infty} \sum_{k=1}^{d} \|I_n(f_{n,k}) - I_n(g_{n,k})\|_2 < \infty. \tag{6.3}$$

Define $Z_n = \mathbb{E}F_n + \sum_{k=1}^d I_k(g_{n,k})$. It follows from Chebyshev's inequality and the Borel-Cantelli lemma that for each k, $I_n(f_{n,k}) - I_n(g_{n,k})$ tends to zero almost surely as $n \to \infty$. In particular Z_n converges almost surely to F_∞. Moreover,

$$\mathbb{E} \sup_{n\in\mathbb{N}} |Z_n| \leq \mathbb{E} \sup_{n\in N} |F_n| + \mathbb{E} \sup_{n\in\mathbb{N}} |Z_n - F_n| \leq \mathbb{E}X + \sum_{n=0}^{\infty}\sum_{k=1}^{d} \|I_n(f_{n,k}) - I_n(g_{n,k})\|_2 < \infty.$$

Therefore it is enough to prove the almost sure convergence of $Z_{n,k} := I_n(g_{n,k})$. To this end we will closely follow the strategy used in the proof of Lemma 5.1.

Assume that $g_{n,k}$ is of the form

$$g_{n,k} = \sum_{i_1,\ldots,i_k=1}^{N_n} a^{(n,k)}_{i_1,\ldots,i_k} \mathbb{1}_{A_{n,i_1}\times\ldots\times A_{n,i_k}},$$

where the sets $A_{n,i}$ are pairwise disjoint and of finite measure λ and coefficients $a^{(n,k)}_{i_1,\ldots,i_k}$ are symmetric and with vanishing diagonals (as explained before we can assume that the family of sets $A_{n,1},\ldots,A_{n,N_n}$ does not depend on k). Note that

$$Z_{n,k} = I_k(g_{n,k}) = \sum_{i_1,\ldots,i_k=1}^{N_n} a^{(n,k)}_{i_1,\ldots,i_k} \prod_{j=1}^{k}(\eta(A_{n,i_j}) - \lambda(A_{n,i_j}))$$

Let $\eta_1,\ldots,\eta_d$ be independent copies of the Poisson process η and define the decoupled version of Z_n with the formula

$$Z_n^{dec} = \mathbb{E}F_n + \sum_{k=1}^{d} \frac{1}{\binom{d}{k}} \sum_{1\le r_1<\ldots<r_k\le d} \sum_{i_1,\ldots,i_k=1}^{N_n} a^{(n,k)}_{i_1,\ldots,i_k} \prod_{j=1}^{k}(\eta_{r_j}(A_{n,i_j}) - \lambda(A_{n,i_j})).$$

The almost sure convergence of Z_n can be written as the following Cauchy type condition

$$\lim_{n\to\infty} \sup_{m>n} \mathbb{P}(\sup_{n\le l\le m} |Z_n - Z_l| \ge \varepsilon) = 0$$

for all $\varepsilon > 0$, while the majorization by an integrable random variable as

$$\lim_{m\to\infty} \mathbb{E} \sup_{0\le l\le m} |Z_l| < \infty.$$

Fix m and recall from the discussion following (6.2), that there exists M and pairwise disjoint sets of finite measure λ, $B_1,\ldots,B_M$ together with symmetric coefficients $b^{(l,k)}_{i_1,\ldots,i_k}$, vanishing on diagonals, such that for every $l \le m$,

$$g_{l,k} = \sum_{i_1,\ldots,i_k=1}^{M} b^{(l,k)}_{i_1,\ldots,i_k} \mathbb{1}_{B_{i_1}\times\ldots\times B_{i_k}},$$

so that

$$Z_{l,k} = \sum_{i_1,\ldots,i_k=1}^{M} b^{(l,k)}_{i_1,\ldots,i_k} \prod_{j=1}^{k} \Big(\eta(B_{i_j}) - \lambda(B_{i_j})\Big)$$

(to simplify the notation we suppress the dependence of M and the sets B_i on m).

Thus, setting $X_i = \eta(B_i) - \lambda(B_i)$, we get for $l \le m$,

$$Z_l = \sum_{1\le i_1\neq\ldots\neq i_d\le M} h^{(l)}_{i_1,\ldots,i_d}(X_{i_1},\ldots,X_{i_d}),$$

where

$$h^{(l)}_{i_1,\ldots,i_d}(x_1,\ldots,x_d) = \frac{(M-d)!}{M!}\mathbb{E}F_n + \sum_{k=1}^{d} \frac{(d-k)!}{d!}\frac{(M-d)!}{(M-k)!} \sum_{1\le r_1\neq\ldots\neq r_k\le d} b^{(l,k)}_{i_{r_1},\ldots,i_{r_k}} x_{r_1}\cdots x_{r_k}.$$

Denote $X_i^{(j)} = \eta_j(B_i) - \lambda(B_i)$. Using the additivity of η_j and λ, one can check that for $l \le m$,

$$Z_l^{dec} = \sum_{1\le i_1\neq\ldots\neq i_d\le M} h^{(l)}_{i_1,\ldots,i_d}(X^{(1)}_{i_1},\ldots,X^{(d)}_{i_d})$$

and hence applying the decoupling inequalities of Theorem A.1 to the spaces $\ell_\infty(\{n,n+1,\ldots,m\})$ and $\ell_\infty(\{0,1,\ldots,m\})$ and functions

$$H_{i_1,\ldots,i_d}(x_1,\ldots,x_d) = (h^{(l)}_{i_1,\ldots,i_d}(x_1,\ldots,x_d) - h^{(n)}_{i_1,\ldots,i_d}(x_1,\ldots,x_d))_{l=n}^{m}$$
$$G_{i_1,\ldots,i_d}(x_1,\ldots,x_d) = (h^{(l)}_{i_1,\ldots,i_d}(x_1,\ldots,x_d))_{l=0}^{m}$$

respectively, we obtain

$$\lim_{n\to\infty}\sup_{m>n}\mathbb{P}(\sup_{n\le l\le m}|Z_n^{dec}-Z_l^{dec}|\ge\varepsilon)=0$$

for all $\varepsilon>0$, and

$$\lim_{m\to\infty}\mathbb{E}\sup_{0\le l\le m}|Z_l^{dec}|<\infty,$$

i.e., Z_n^{dec} converges almost surely and is dominated by an integrable random variable. By the Fubini Theorem, if we fix $s_1<\ldots<s_k\in[d]$, then with probability one Z_n^{dec} converges almost surely with respect to $\{\eta_i\colon i\in[d]\setminus\{s_1,\ldots,s_k\}\}$ and is almost surely dominated by some integrable random variable. Thus with probability one it converges in $L_1(\eta_i\colon i\in[d]\setminus\{s_1,\ldots,s_k\})$, and in particular $\mathbb{E}(Z_n^{dec}|\eta_{s_1},\ldots,\eta_{s_k})$ converges almost surely for every choice of $s_1,\ldots,s_k$. But

$$\begin{aligned}&\mathbb{E}(Z_n^{dec}|\eta_{s_1},\ldots,\eta_{s_k})\\&=\mathbb{E}F_n+\sum_{l=1}^{k}\frac{1}{\binom{d}{l}}\sum_{\substack{1\le r_1<\ldots<r_l\le d\\ r_1,\ldots,r_l\subset\{s_1,\ldots,s_k\}}}\sum_{i_1,\ldots,i_l=1}^{N_n}a^{(n,l)}_{i_1,\ldots,i_l}\prod_{j=1}^{l}(\eta_{r_j}(A_{n,i_j})-\lambda(A_{n,i_j})).\end{aligned}\tag{6.4}$$

From this, by induction one easily proves that $\mathbb{E}F_n$ is convergent and for any $1\le k\le d$, the sequence

$$Z_{n,k}^{dec}=\sum_{i_1,\ldots,i_k=1}^{N_n}a^{(n,k)}_{i_1,\ldots,i_k}\prod_{j=1}^{k}(\eta_j(A_{n,i_j})-\lambda(A_{n,i_j}))$$

converges almost surely. Indeed, taking $k=0$, we obtain convergence of $\mathbb{E}F_n$. Now assuming that $\mathbb{E}F_n$ converges and $Z_{n,l}^{dec}$ for $1\le l<k$ converge almost surely, by equidistribution of η_i we obtain that for any $l<k$

$$\sum_{\substack{1\le r_1<\ldots<r_l\le d\\ r_1,\ldots,r_l\subset\{1,\ldots,k\}}}\sum_{i_1,\ldots,i_l=1}^{N_n}a^{(n,l)}_{i_1,\ldots,i_l}\prod_{j=1}^{l}(\eta_{r_j}(A_{n,i_j})-\lambda(A_{n,i_j}))$$

converges almost surely, which combined with the almost sure convergence of the sequence $\mathbb{E}(Z_n^{dec}|\eta_1,\ldots,\eta_k)$ and (6.4) yields the almost sure convergence of $Z_{n,k}^{dec}$.

Now, using the decoupling inequalities in the opposite direction than before (we skip the definition of the corresponding functions h, which in this case is easier, since we deal with homogeneous polynomials), we obtain that the sequence $Z_{n,k}$ converges almost surely for each $k\le d$. By Lemma A.2 we obtain that

$$\mathbb{E}\sup_{n\in\mathbb{N}}|Z_{n,k}|\le C\mathbb{E}\sup_{n\in\mathbb{N}}|Z_n|<\infty,$$

so we also have convergence in L_1 (note that Lemma A.2 could be recovered from the above decoupling arguments, in fact this is the way it was proved in [1], but we prefer to rely on the abstract formulation, so as not to further complicate the above elementary but notationally unpleasant arguments). We have thus established that the variables $Z_{n,k}$ converge almost surely and in L_1 to some random variables $F_{\infty,k}$ and it follows from (6.3) and the subsequent discussion, that the same convergence holds for $F_{n,k}$. In particular we have $F_\infty=\mathbb{E}F_\infty+\sum_{k=1}^{d}F_{\infty,k}$.

It remains to prove that if $(F_n)_{n=0}^{\infty}$ is bounded in L_2, then $F_{\infty,k}$ can be expressed as a k-fold stochastic integral of a square integrable symmetric function. Note that by orthogonality, for each $k\le d$, $(F_{n,k})_{n=1}^{\infty}$ is bounded in L_2. Thus one can select a subsequence $(I_k(f_{n_l,k}))_{l=1}^{\infty}$, which converges weakly in L_2 to some random variable $\widetilde{F}_{\infty,k}$.

Since the k-th chaos is a closed linear subspace of L_2, it follows that $\widetilde{F}_{\infty,k} = I_k(f_{\infty,k})$ for some $f_{\infty,k} \in L_{2,s}(\mathcal{X}^k, \mathcal{F}^{\otimes k}, \lambda^{\otimes k})$. Moreover by the convergence of $F_{n,k}$ to $F_{\infty,k}$ in L_1, we obtain that for every measurable set A, $\mathbb{E}F_{\infty,k}\mathbb{1}_A = \lim_{n\to\infty} \mathbb{E}F_{n,k}\mathbb{1}_A = \mathbb{E}\widetilde{F}_{\infty,k}\mathbb{1}_A$, which shows that $F_{\infty,k} = \widetilde{F}_{\infty,k}$ almost surely and ends the proof of the theorem. □

Remark 6.1. Let us note that variants of the above argument can be repeated to prove the almost sure convergence in more general situations, e.g., for square integrable random fields for which one defines multiple Wiener-Itô integrals by the L_2 theory, for tetrahedral polynomial chaos based on sequences of independent random variables (as investigated in the previous section) or for U-statistics, as in all these settings we can apply the general decoupling inequality in a similar manner.

An alternate proof of Theorem 3.2. We will focus on the proof of almost sure convergence since the other parts of the theorem are its relatively straightforward consequences (as can be seen from the first proof given above).

The argument we will present is based on Mehler's formula for the Poisson process and is a direct counterpart of the proof of Theorem 1.2 due to Poly and Zheng. It is based on the notions related to the analysis of the Ornstein-Uhlenbeck semigroup on the Poisson space and the corresponding Mehler's formula. We refer the reader to [13, 20] for a comprehensive description of the theory. Here we will just introduce the basic elements, required for the argument.

For any measurable function f on the space $\mathbf{N}(\mathcal{X})$, any $\mu \in \mathbf{N}(\mathcal{X})$ and any $x \in \mathcal{X}$ we define $D_x f(\mu) = D^1_x f(\mu) = f(\mu + \delta_x) - f(\mu)$ and inductively $D_{x_1,\ldots,x_k} f(\mu) = D^1_{x_1} D^{k-1}_{x_2,\ldots,x_k} f(\mu)$. We also set $D^0 f = f$. For a random variable $F = f(\eta)$ with f as above we define $D^k_{x_1,\ldots,x_k} F = D_{x_1,\ldots,x_k} f(\eta)$. One shows that up to a set of $\mathbb{P} \otimes \lambda^{\otimes k}$ measure zero, this definition does not depend on the choice of the representative f.

We also define the symmetric functions $T_k f \colon \mathcal{X}^k \to \mathbb{R}$ with the formula

$$T_k f(x_1, \ldots, x_k) = \mathbb{E} D_{x_1,\ldots,x_k} f(\eta).$$

For $F = f(\eta) \in L_2(\eta)$, the functions $T_k f$ are square integrable with respect to $\lambda^{\otimes k}$ and we have the chaos representation (see [13, Theorem 18.10]), namely the equality

$$F = \sum_{k=0}^{\infty} \frac{1}{k!} I_k(T_k f), \tag{6.5}$$

with the series converging in $L_2(\eta)$. Note that thanks to orthogonality of the Wiener-Poisson chaoses and the isometry properties of I_n, the expansion $F = \sum_{k=0}^{\infty} I_k(g_k)$ with $g_k \in L_{2,s}(\mathcal{X}^k, \mathcal{F}^{\otimes k}, \lambda^{\otimes k})$ is unique.

If η is proper, i.e., it can be almost surely represented as a sum of Dirac's deltas, $\eta = \sum_{k=1}^{\kappa} \delta_{X_n}$ (where κ is an $\mathbb{N} \cup \{\infty\}$-valued random variable), we also define the t-thinning of η ($t \in [0, 1]$) as

$$\eta_t = \sum_{n=1}^{\kappa} \mathbb{1}_{\{U_n \le t\}} \delta_{X_n},$$

where $U_1, U_2, \ldots$ are independent random variables distributed uniformly on $[0, 1]$, independent of η. By the Thinning Theorem [13, Corollary 5.9], η_t is a Poisson process with intensity $t\lambda$.

Finally one defines a family of operators P_t, $t \in [0, 1]$ on $L_2(\eta)$ with the formula

$$P_t F = \mathbb{E}(F(\eta_t + \eta'_{1-t})|\eta), \tag{6.6}$$

where η'_{1-t} is a Poisson process with intensity $(1-t)\lambda$, independent of the pair (η, η_t). Note that by the Superposition Theorem [13, Theorem 3.3], $\eta_t + \eta'_{1-t}$ is again a Poisson process with intensity λ. In particular $\mathbb{E}P_t F = \mathbb{E}F$.

Mehler's formula ([12], [13, Lemma 20.1]) asserts now that for any $F \in L_2(\eta)$ and $t \in [0,1]$,

$$D^k_{x_1,\ldots,x_k}(P_t F) = t^k P_t D_{x_1,\ldots,x_k} F, \ \lambda^{\otimes k}\text{-a.e., } \mathbb{P}\text{-a.s.} \tag{6.7}$$

As a consequence

$$\mathbb{E}D^n_{x_1,\ldots,x_k}(P_t F) = t^k \mathbb{E}D_{x_1,\ldots,x_k} F, \ \lambda^{\otimes k}\text{-a.e.} \tag{6.8}$$

In the setting of Theorem 3.2, we can assume without loss of generality that η is proper (since we can always find a proper Poisson process with the same distribution as η, cf. [13, Corollary 3.7]).

We can also assume that η and η'_{1-t} are defined on a product probability space $\Omega = \Omega_\eta \times \Omega_{\eta'}$ with measure $\mathbb{P}_\eta \otimes \mathbb{P}_{\eta'}$ and that they depend respectively only on the first and second coordinate. Since $\widetilde{\eta} = \eta_t + \eta'_{1-t}$ has the same distribution as η, if we define $\widetilde{F}_n = \mathbb{E}F_n + \sum_{k=1}^d \widetilde{I}_k(f_{n,k})$, where $\widetilde{I}_k$ is the k-fold Wiener-Itô integral with respect to $\widetilde{\eta}$, then $\widetilde{F}_n$ also converges almost surely and $\sup_n |\widetilde{F}_n|$ is integrable. Thus, by the Fubini theorem, it follows that $\mathbb{P}_\eta$-almost surely, the sequence $\widetilde{F}_n$ converges almost surely with respect to $\mathbb{P}_{\eta'}$ and is uniformly integrable. In particular, using the definition (6.6) we obtain that $P_t F_n = \int \widetilde{F}_n d\mathbb{P}_{\eta'}$ converges almost surely as $t \to \infty$.

On the other hand (6.8) and the chaos representation property (6.5) imply that

$$P_t F_n = \mathbb{E}F_n + \sum_{k=1}^d t^k I_k(f_{n,k}).$$

Analogously as in the original argument by Poly and Zheng in the proof of Theorem 1.2, using the fact that the right-hand side above converges almost surely for sufficiently many $t \in [0,1]$, we obtain that $I_k(f_{n,k})$ converges almost surely for each $k \le d$. □

A Decoupling and related inequalities

In this section we gather basic facts concerning decoupling inequalities for U-statistics that are used throughout the article.

Let us start with the by now classical decoupling inequality due to de la Peña and Montgomery-Smith.

Theorem A.1 ([4, Theorem 1]). *Let d be a positive integer and for $n \ge d$ let $(X_i)_{i=1}^n$ be a sequence of independent random variables with values in a measurable space $(S, \mathcal{S})$ and let $(X_i^{(j)})_{i=1}^n$ $j = 1, \ldots, d$ be d independent copies of this sequence. Let E be a separable Banach space and for each $(i_1, \ldots, i_d) \in [n]^d$ with pairwise distinct coordinates let $h_{i_1,\ldots,i_d} \colon S^d \to E$ be a measurable function. There exists a numerical constant C_d, depending only on d such that for all $t > 0$,*

$$\mathbb{P}\Big(\Big\| \sum_{1 \le i_1 \ne \ldots \ne i_d \le n} h_{i_1,\ldots,i_d}(X_{i_1}, \ldots, X_{i_d})\Big\| > t\Big) \le C_d \mathbb{P}\Big(\Big\| \sum_{1 \le i_1 \ne \ldots \ne i_d \le n} h_{i_1,\ldots,i_d}(X^{(1)}_{i_1}, \ldots, X^{(d)}_{i_d})\Big\| > t/C_d\Big).$$

As a consequence for all $p \ge 1$,

$$\Big\| \sum_{1 \le i_1 \ne \ldots \ne i_d \le n} h_{i_1,\ldots,i_d}(X_{i_1}, \ldots, X_{i_d})\Big\|_p \le C'_d \Big\| \sum_{1 \le i_1 \ne \ldots \ne i_d \le n} h_{i_1,\ldots,i_d}(X^{(1)}_{i_1}, \ldots, X^{(d)}_{i_d})\Big\|_p,$$

where C'_d is another numerical constant depending only on d.

If moreover the functions $h_{i_1,\ldots,i_d}$ are symmetric in the sense that, for all $x_1,\ldots,x_d \in S$ and all permutations $\pi\colon [d] \to [d]$, $h_{i_1,\ldots,i_d}(x_1,\ldots,x_d) = h_{i_{\pi_1},\ldots,i_{\pi_d}}(x_{\pi_1},\ldots,x_{\pi_d})$, then for all $t > 0$,

$$\mathbb{P}\Big(\Big\|\sum_{1\le i_1\neq\ldots\neq i_d\le n} h_{i_1,\ldots,i_d}(X^{(1)}_{i_1},\ldots,X^{(d)}_{i_d})\Big\| > t\Big) \le \widetilde{C}_d\mathbb{P}\Big(\Big\|\sum_{1\le i_1\neq\ldots\neq i_d\le n} h_{i_1,\ldots,i_d}(X_{i_1},\ldots,X_{i_d})\Big\| > t/\widetilde{C}_d\Big),$$

where $\widetilde{C}_d$ is a constant depending only on d. As a consequence for some numerical constant $\widetilde{C}'_d$, depending only on d, and all $p \ge 1$,

$$\Big\|\sum_{1\le i_1\neq\ldots\neq i_d\le n} h_{i_1,\ldots,i_d}(X^{(1)}_{i_1},\ldots,X^{(d)}_{i_d})\Big\|_p \le \widetilde{C}'_d\Big\|\sum_{1\le i_1\neq\ldots\neq i_d\le n} h_{i_1,\ldots,i_d}(X_{i_1},\ldots,X_{i_d})\Big\|_p.$$

Another result used in our proofs is the following reverse triangle inequality for tetrahedral chaoses, obtained for the first time by Kwapień [9] in the symmetric setting, which easily gives the general case (see also [1], where an alternate proof in the general case, based on Theorem A.1 is presented). We remark that this lemma can be also obtained by methods used by Poly and Zheng in their proof of Theorem 1.1.

Lemma A.2. *For $j = 0,1,\ldots,d$ let $(a^j_{i_1,\ldots,i_j})_{1\le i_1,\ldots,i_j\le n}$ be a k-indexed symmetric array of real numbers (or more generally elements of some normed space), such that $a^j_{i_1,\ldots,i_j} = 0$ if $i_k = i_l$ for some $1 \le k < l \le j$ (for $j = 0$ we have just a single number $a^0_\emptyset$). Let $X_1,\ldots,X_n$ be independent mean zero random variables. Then there exists a constant $C_d \in (0,\infty)$, depending only on d, such that for all $p \ge 1$,*

$$\sum_{j=0}^d \Big\|\sum_{i_1,\ldots,i_j=1}^n a^j_{i_1,\ldots,i_j} X_{i_1}\cdots X_{i_j}\Big\|_p \le C_d\Big\|\sum_{j=0}^d\sum_{i_1,\ldots,i_j=1}^n a^j_{i_1,\ldots,i_j} X_{i_1}\cdots X_{i_j}\Big\|_p.$$

B Proof of Proposition 5.5

We will now prove the characterization of the convergence in probability to one, given in Proposition 5.5.

Proof. Assume first that conditions (i), (ii) are satisfied. By (ii) we get $\mathbb{P}(\max_{i\le k_n}|X_{n,i}| > \tau) \to 0$ and as a consequence $\sum_{k=0}^{k_n} X_{k,n}\mathbb{1}_{\{|X_{n,k}|>\tau\}}$ converges in probability to zero. On the other hand, by (i), (ii) and Chebyshev's inequality, $\sum_{k=0}^{k_n} X_{n,k}\mathbb{1}_{\{|X_{n,k}|\le\tau\}}$ converges in probability to one, which ends the proof of the first implication (note that we did not use the asymptotic smallness condition (5.4)).

Assume now that $\sum_{k=0}^{k_n} X_{n,k}$ converges in probability to one. Denote $\mathbb{X}_n = (X_{n,k})_{k=0}^{k_n}$ and let $\mathbb{X}' = (X'_{n,k})_{k=0}^{k_n}$ be an independent copy of $\mathbb{X}$. We have $\sum_{k=0}^{k_n}(X_{n,k} - X'_{n,k}) \to 0$ in probability. Since $X_{n,k} - X'_{n,k}$ is symmetric we also have $S_n := \sum_{k=0}^{k_n}\varepsilon_k|X_{n,k} - X'_{n,k}| \to 0$ in probability, where $\varepsilon_k, k \in \mathbb{N}$ are i.i.d. Rademacher variables independent of $(X_{n,k}), (X'_{nk})$. Consider the event

$$A_n = \{\max_{k\le k_n}|X_{n,k}| > \tau\} = \bigcup_{k=0}^{k_n} A_{n,k},$$

where $A_{n,k} = \{\forall_{0\le i<k}|X_{i,n}| \le \tau, |X_{n,k}| > \tau\}$. Note that by independence and (5.4) for large n, on $A_{n,k}$, $\mathbb{P}(|X_{n,k} - X'_{n,k}| \ge \tau/2|\mathbb{X}) \ge 1/2$. Moreover, by symmetry of the

Rademacher variables $\mathbb{P}(|S_n| \geq |X_{n,k} - X'_{n,k}||\mathbb{X}, \mathbb{X}') \geq 1/2$. Therefore we get

$$\mathbb{P}(|S_n| \geq \tau/2) \geq \sum_{k=0}^{k_n} \mathbb{P}(\{|S_n| \geq \tau/2\} \cap A_{n,k}) \geq \frac{1}{4} \sum_{k=0}^{k_n} \mathbb{P}(A_{n,k}) = \mathbb{P}(A_n)/4.$$

As a consequence $\mathbb{P}(A_n) \to 0$ as $n \to \infty$. A standard estimate

$$\mathbb{P}(A_n) \geq \frac{1}{2} \min\Big(\sum_{k=0}^{k_n} \mathbb{P}(|X_{n,k}| > \tau), 1\Big)$$

shows that

$$\sum_{k=0}^{k_n} \mathbb{P}(|X_{n,k}| > \tau) \to 0 \tag{B.1}$$

as $n \to \infty$.

Define now $Z_{n,k} = (X_{n,k}\mathbb{1}_{\{|X_{n,k}| \leq \tau\}} - X'_{n,k}\mathbb{1}_{\{|X'_{n,k}| \leq \tau\}})$ and $\widetilde{S}_n = \sum_{k=0}^{k_n} Z_{n,k}$. We have $\widetilde{S}_n \to 0$ in probability. Moreover, $\mathbb{E}Z_{n,k} = 0$ and so by independence,

$$\mathbb{E}\widetilde{S}_n^4 = \sum_{k=0}^{k_n} \mathbb{E}Z_{n,k}^4 + 3\mathbb{E} \sum_{1 \leq i \neq j \leq k_n} \mathbb{E}Z_{n,i}^2 \mathbb{E}Z_{n,j}^2 \leq 4\tau^2 \mathbb{E}\widetilde{S}_n^2 + 3(\mathbb{E}\widetilde{S}_n^2)^2.$$

By the Paley-Zygmund inequality (see, e.g., [3, Corollary 3.3.2]),

$$\mathbb{P}\Big(|\widetilde{S}_n| \geq \frac{1}{2}(\mathbb{E}\widetilde{S}_n^2)^{1/2}\Big) \geq \frac{9}{16} \frac{(\mathbb{E}\widetilde{S}_n^2)^2}{\mathbb{E}\widetilde{S}_n^4} \geq \frac{9}{16} \frac{(\mathbb{E}\widetilde{S}_n^2)^2}{4\tau^2 \mathbb{E}\widetilde{S}_n^2 + 3(\mathbb{E}\widetilde{S}_n^2)^2}.$$

This shows that $\mathbb{E}\widetilde{S}_n^2 \to 0$ as $n \to \infty$ (since otherwise the right hand side above would be separated from zero along a subsequence). But $\mathbb{E}\widetilde{S}_n^2 = 2\sum_{k=0}^{k_n} \mathrm{Var}(X_{n,k}\mathbb{1}_{\{|X_{n,k}| \leq \tau\}})$ which together with (B.1) proves (ii). The convergence asserted in (i) is now an immediate consequence of (ii) and the convergence $\sum_{k=0}^{k_n} X_{n,k}\mathbb{1}_{\{|X_{n,k}| \leq \tau\}} \to 1$ in probability. □

C Density of simple functions vanishing on the diagonal

We will now prove a lemma concerning approximation properties in the space of square integrable symmetric functions on a d-fold product of a measurable space, endowed with a σ-finite measure, which is used in the proof of Theorem 3.2. Variations of this lemma appear in the literature, e.g., as a tool for defining multiple Wiener-Itô integrals in the Gaussian case. A proof of one of the versions can be found, e.g., in [18] (pages 8–9), see also [7, Proposition E.16]. Since we have not been able to find in the literature a formulation which would correspond exactly to our needs, we state one here together with a full proof.

Let us recall that a measure λ on a measurable space $(\mathcal{X}, \mathcal{F})$ is called non-atomic if for every $A \in \mathcal{F}$ with $\lambda(A) > 0$, there exists $B \in \mathcal{F}$ with $B \subset A$ and $0 < \lambda(B) < \lambda(A)$. The measure λ is called σ-finite if $\mathcal{X}$ can be represented as a countable union of sets of finite λ measure.

Lemma C.1. *Let λ be a σ-finite, non-atomic measure on a measurable space $(\mathcal{X}, \mathcal{F})$ and denote by $L_{2,s}(\mathcal{X}^d, \mathcal{F}^{\otimes d}, \lambda^{\otimes d})$ the space of square integrable functions on $\mathcal{X}^d$, symmetric in their arguments (treated as a subspace of $L_2(\mathcal{X}^d, \mathcal{F}^{\otimes d}, \lambda^{\otimes d})$). Let $\mathcal{E}$ be the space of all functions of the form*

$$g = \sum_{i_1,\ldots,i_d=1}^{N} a_{i_1,\ldots,i_d} \mathbb{1}_{A_{i_1} \times \cdots \times A_{i_d}}, \tag{C.1}$$

where

(i) $N \in \mathbb{N}$,

(ii) $A_1, \ldots, A_N \in \mathcal{F}$ *are pairwise disjoint and of finite* λ *measure,*

(iii) *the coefficients* $a_{i_1,\ldots,i_d}$ *are symmetric under permutations of indices and such that* $a_{i_1,\ldots,i_d} = 0$ *whenever there exist* $k \neq l$ *such that* $i_k = i_l$.

Then $\mathcal{E}$ *is dense in* $L_{2,s}(\mathcal{X}^d, \mathcal{F}^{\otimes d}, \lambda^{\otimes d})$ *with respect to the* L_2 *norm.*

Proof. Step 1. Since we will be working with fixed d, to simplify the notation, let us abbreviate $L_2(\mathcal{X}^d, \mathcal{F}^{\otimes d}, \lambda^{\otimes d})$ to L_2 and $L_{2,s}(\mathcal{X}^d, \mathcal{F}^{\otimes d}, \lambda^{\otimes d})$ to $L_{2,s}$. Let $K_n \subset \mathcal{X}$ be an increasing sequence of measurable sets with $\lambda(K_n) < \infty$, such that $\bigcup_{n=1}^{\infty} K_n = \mathcal{X}$. Since $f\mathbb{1}_{K_n^d} \to f$ in $L_{2,s}$ as $n \to \infty$, without loss of generality we may and will assume that $\lambda(\mathcal{X}) < \infty$.

Step 2. Consider first a function of the form $f = \mathbb{1}_C$ where $C \in \mathcal{F}^{\otimes d}$ (not necessarily symmetric). We will show that it can be arbitrarily well approximated in L_2 by functions of the form

$$g = \sum_{i_1,\ldots,i_d=1}^{N} b_{i_1,\ldots,i_d} \mathbb{1}_{A_{1,i_1} \times \cdots \times A_{d,i_d}},$$

where $N \in \mathbb{N}$, $b_{i_1,\ldots,i_d} \in \mathbb{R}$, $A_{j,i} \in \mathcal{F}$ for $1 \leq i \leq N$, $1 \leq j \leq d$. Indeed, the class $\mathcal{C}$ of all sets C with this property is a λ-system in Dynkin's sense, i.e., it contains $\mathcal{X}$ and as one can easily check it is closed under complements and countable unions of pairwise disjoint sets. Since it trivially contains all product sets and the class of product sets is closed under finite intersections (i.e., it is a π-system), it follows from the application of the π-λ-theorem (see [8, Theorem 1.1]) that $\mathcal{C} = \mathcal{F}^{\otimes d}$.

Step 3. Now, any $f \in L_2$ can be arbitrarily well approximated by functions of the form $\sum_{i=1}^{n} a_i \mathbb{1}_{C_i}$, where $C_i \in \mathcal{F}^{\otimes d}$, $a_i \in \mathbb{R}$, and thus (by Step 2) also by functions of the form

$$g = \sum_{i_1,\ldots,i_d=1}^{N} b_{i_1,\ldots,i_d} \mathbb{1}_{A_{1,i_1} \times \cdots \times A_{d,i_d}}, \tag{C.2}$$

where $N \in \mathbb{N}$ and $A_{j,i} \in \mathcal{F}$ for $1 \leq j \leq d$, $1 \leq i \leq N$ and $b_{i_1,\ldots,i_d} \in \mathbb{R}$. Without loss of generality we can assume that for all $1 \leq j \leq d$, $\bigcup_{i=1}^{N} A_{j,i} = \mathcal{X}$. Moreover by considering all possible intersections of sets from the family $\{A_{j,i} \colon 1 \leq j \leq d, 1 \leq i \leq N\}$ and their complements we can represent any function g as in (C.2) as

$$g = \sum_{i_1,\ldots,i_d=1}^{n} c_{i_1,\ldots,i_d} \mathbb{1}_{A_{i_1} \times \cdots \times A_{i_d}} \tag{C.3}$$

where $n \in \mathbb{N}$, $c_{i_1,\ldots,i_d} \in \mathbb{R}$, the sets $A_1, \ldots, A_n \in \mathcal{F}$ are pairwise disjoint and their union is $\mathcal{X}$.

Step 4. Recall now the well-known Darboux-type property of non-atomic measures: for every $A \in \mathcal{F}$ and every $\alpha \in [0, \lambda(A)]$, there exists $B \in \mathcal{F}$, $B \subset A$, such that $\lambda(B) = \alpha$ (see, e.g., [2, Exercise 2.17]). It follows that for every g of the form (C.3), for every positive integer m, there exist $n_m \in \mathbb{N}$, coefficients $c^{(m)}_{i_1,\ldots,i_d} \in \mathbb{R}$, $i_1, \ldots, i_d \leq n_m$, and pairwise disjoint sets $A_{m,i} \in \mathcal{F}$, $1 \leq i \leq n_m$ such that $\bigcup_{i=1}^{n_m} A_{m,i} = \mathcal{X}$,

$$\max_{1 \leq i \leq n_m} \lambda(A_{m,i}) < \frac{1}{m}, \tag{C.4}$$

$$g = \sum_{i_1,\ldots,i_d=1}^{n_m} c^{(m)}_{i_1,\ldots,i_d} \mathbb{1}_{A_{m,i_1} \times \ldots \times A_{m,i_d}}. \tag{C.5}$$

and

$$\max_{i_1,\ldots,i_d \leq n_m} |c^{(m)}_{i_1,\ldots,i_d}| \leq M := \max_{i_1,\ldots,i_d} |c_{i_1,\ldots,i_d}|.$$

Let now

$$g_m = \sum_{i_1,\ldots,i_d=1}^{n_m} \widetilde{c}^{(m)}_{i_1,\ldots,i_d} \mathbb{1}_{A_{m,i_1}\times\ldots\times A_{m,i_d}},$$

where $\widetilde{c}^{(m)}_{i_1,\ldots,i_d} = c^{(m)}_{i_1,\ldots,i_d}$ if the indices $i_1,\ldots,i_d$ are pairwise distinct and $\widetilde{c}^{(m)}_{i_1,\ldots,i_d} = 0$ otherwise. Let $I_m = \{(i_1,\ldots,i_d) \in [n_m]^d \colon \exists_{1\le k<l\le d}\, i_k = i_l\}$ and observe that

$$\begin{aligned}\|g - g_m\|_2^2 &= \sum_{(i_1,\ldots,i_d)\in I_m} |c^{(m)}_{i_1,\ldots,i_d}|^2 \lambda(A_{i_1})\cdots\lambda(A_{i_d}) \\ &\le M^2 \binom{d}{2} \sum_{i=1}^{n_m} \lambda(A_i)^2 \Big(\sum_{i=1}^{n_m}\lambda(A_i)\Big)^{d-2} \le M^2\binom{d}{2}\frac{1}{m}\lambda(\mathcal{X})^{d-1} \to 0,\end{aligned}$$

as $m\to\infty$.

Together with previous steps, this shows that every $f \in L_2$ can be arbitrarily well approximated by functions g of the form

$$g = \sum_{i_1,\ldots,i_d=1}^{N} e_{i_1,\ldots,i_d} \mathbb{1}_{A_{i_1}\times\ldots\times A_{i_d}} \tag{C.6}$$

where $N\in\mathbb{N}$, $A_1,\ldots,A_N \in \mathcal{F}$ are pairwise disjoint, $\bigcup_{i\le N} A_i = \mathcal{X}$ and the coefficients $e_{i_1,\ldots,i_d} \in \mathbb{R}$ satisfy $e_{i_1,\ldots,i_d} = 0$ whenever $i_k = i_l$ for some $k \ne l$.

Step 5. Assume now that $f \in L_{2,s}$ and for $\varepsilon > 0$ let g be a function of the form (C.6) such that $\|f - g\|_2 < \varepsilon$.

Define $\widetilde{g}$ as

$$\widetilde{g} = \sum_{i_1,\ldots,i_d=1}^{N} a_{i_1,\ldots,i_d} \mathbb{1}_{A_{i_1}\times\cdots\times A_{i_d}},$$

with $a_{i_1,\ldots,i_d} = \frac{1}{d!}\sum_{\sigma\in S_d} e_{i_{\sigma(1)},\ldots,i_{\sigma(d)}}$, where S_d is the set of all permutations of $[d]$. Clearly, $\widetilde{g} \in \mathcal{E}$. Note that for every $(x_1,\ldots,x_d) \in \mathcal{X}^d$ and $\sigma \in S_d$,

$$g(x_{\sigma(1)},\ldots,x_{\sigma(d)}) = \sum_{i_1,\ldots,i_d=1}^{N} e_{i_{\sigma(1)},\ldots,i_{\sigma(d)}} \mathbb{1}_{A_{i_1}\times\ldots\times A_{i_d}}(x_1,\ldots,x_d).$$

Moreover, by symmetry of f, we have $\lambda^{\otimes d}$-almost everywhere,

$$f(x_1,\ldots,x_d) = \frac{1}{d!}\sum_{\sigma\in S_d} f(x_{\sigma(1)},\ldots,x_{\sigma(d)}).$$

Thus, by the triangle inequality and the fact that the measure $\lambda^{\otimes d}$ is invariant under permutation of coordinates, we obtain

$$\begin{aligned}\|f - \widetilde{g}\|_2 &\le \frac{1}{d!}\sum_{\sigma\in S_d}\Big(\int_{\mathcal{X}^d} (f(x_{\sigma(1)},\ldots,x_{\sigma(d)}) - g(x_{\sigma(1)},\ldots,x_{\sigma(d)}))^2 \lambda(dx_1)\cdots\lambda(dx_d)\Big)^{1/2} \\ &= \|f - g\|_2 \le \varepsilon,\end{aligned}$$

which ends the proof. □

References

[1] Radosław Adamczak and Rafał Latała, *Tail and moment estimates for chaoses generated by symmetric random variables with logarithmically concave tails*, Ann. Inst. Henri Poincaré Probab. Stat. **48** (2012), no. 4, 1103–1136. MR-3052405

[2] Patrick Billingsley, *Probability and measure*, third ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1995, A Wiley-Interscience Publication. MR-1324786

[3] Víctor H. de la Peña and Evarist Giné, *Decoupling*, Probability and its Applications (New York), Springer-Verlag, New York, 1999. MR-1666908

[4] Victor H. de la Peña and S. J. Montgomery-Smith, *Bounds on the tail probability of U-statistics and quadratic forms*, Bull. Amer. Math. Soc. (N.S.) **31** (1994), no. 2, 223–227. MR-1261237

[5] Evarist Giné, *A consequence for random polynomials of a result of de la Peña and Montgomery-Smith*, High dimensional probability (Oberwolfach, 1996), Progr. Probab., vol. 43, Birkhäuser, Basel, 1998, pp. 103–110. MR-1652323

[6] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis, *Analysis in Banach spaces. Vol. II*, vol. 67, Springer, Cham, 2017. MR-3752640

[7] Svante Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR-1474726

[8] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR-1876169

[9] Stanisław Kwapień, *Decoupling inequalities for polynomial chaos*, Ann. Probab. **15** (1987), no. 3, 1062–1071. MR-893914

[10] Stanisław Kwapień and Jerzy Szulga, *Hypercontraction methods in moment inequalities for series of independent random variables in normed spaces*, Ann. Probab. **19** (1991), no. 1, 369–379. MR-1085342

[11] Stanisław Kwapień and Wojbor A. Woyczyński, *Random series and stochastic integrals: single and multiple*, Probability and its Applications, Birkhäuser Boston, Inc., Boston, MA, 1992. MR-1167198

[12] Günter Last, Giovanni Peccati, and Matthias Schulte, *Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization*, Probab. Theory Related Fields **165** (2016), no. 3-4, 667–723. MR-3520016

[13] Günter Last and Mathew Penrose, *Lectures on the Poisson process*, Institute of Mathematical Statistics Textbooks, vol. 7, Cambridge University Press, Cambridge, 2018. MR-3791470

[14] Michel Ledoux and Michel Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 23, Springer-Verlag, Berlin, 1991. MR-1102015

[15] Terry R. McConnell and Murad S. Taqqu, *Decoupling inequalities for multilinear forms in independent symmetric random variables*, Ann. Probab. **14** (1986), no. 3, 943–954. MR-841595

[16] Ivan Nourdin and Giovanni Peccati, *Normal approximations with Malliavin calculus*, Cambridge Tracts in Mathematics, vol. 192, Cambridge University Press, Cambridge, 2012. MR-2962301

[17] Ivan Nourdin, Giovanni Peccati, and Xiaochuan Yang, *Restricted hypercontractivity on the Poisson space*, Proc. Amer. Math. Soc. **148** (2020), no. 8, 3617–3632. MR-4108865

[18] David Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR-2200233

[19] Ryan O'Donnell, *Analysis of Boolean functions*, Cambridge University Press, New York, 2014. MR-3443800

[20] Giovanni Peccati and Matthias Reitzner (eds.), *Stochastic analysis for Poisson point processes*, Bocconi & Springer Series, vol. 7, Bocconi University Press, Springer, [Cham], 2016. MR-3444831

[21] V. V. Petrov, *Sums of independent random variables*, Springer-Verlag, New York-Heidelberg, 1975, Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82. MR-0388499

[22] Guillaume Poly and Guangqu Zheng, *Almost sure convergence on chaoses*, Proc. Amer. Math. Soc. **147** (2019), no. 9, 4055–4065. MR-3993797

[23] Luca Pratelli and Pietro Rigo, *On the almost sure convergence of sums*, arXiv e-prints: arXiv:2009.03391 (2020).

[24] Matthias Reitzner, *Poisson point processes: large deviation inequalities for the convex distance*, Electron. Commun. Probab. **18** (2013), no. 96, 7. MR-3151752

[25] Jerzy Szulga, *Introduction to random chaos*, Chapman & Hall, London, 1998. MR-1681553

Acknowledgments. The author would like to express his gratitude to prof. Krzysztof Oleszkiewicz for instructive conversations and encouragement to investigate problems discussed in this article, to the Anonymous Referee and the Associated Editor for a careful reading and useful comments on the initial submission, and to Bartłomiej Polaczyk for helpful remarks concerning the revised version.