JYX JYVÄSKYLÄN YLIOPISTO UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Cox, Sonja; Geiss, Stefan

Title: On Decoupling in Banach Spaces

Year: 2021

Version: Published version

Copyright: © The Author(s) 2021

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Cox, S., & Geiss, S. (2021). On Decoupling in Banach Spaces. Journal of Theoretical Probability, 34(3), 1179-1212. https://doi.org/10.1007/s10959-021-01085-6

Journal of Theoretical Probability
https://doi.org/10.1007/s10959-021-01085-6

On Decoupling in Banach Spaces

Sonja Cox[1] · **Stefan Geiss**[2]

Received: 19 June 2018 / Revised: 8 February 2021 / Accepted: 17 February 2021
© The Author(s) 2021

Abstract
We consider decoupling inequalities for random variables taking values in a Banach space X. We restrict the class of distributions that appear as conditional distributions while decoupling and show that each adapted process can be approximated by a Haar-type expansion in which only the pre-specified conditional distributions appear. Moreover, we show that in our framework a progressive enlargement of the underlying filtration does not affect the decoupling properties (in particular, it does not affect the constants involved). As a special case, we deal with one-sided moment inequalities for decoupled dyadic (i.e., Paley–Walsh) martingales and show that Burkholder–Davis–Gundy-type inequalities for stochastic integrals of X-valued processes can be obtained from decoupling inequalities for X-valued dyadic martingales.

Keywords Decoupling in Banach spaces · Regular conditional probabilities · Dyadic martingales · Stochastic integration

Mathematics Subject Classification 60E15 · 60H05 · 46B09

1 Introduction

The UMD property is crucial in harmonic and stochastic analysis in Banach spaces, see, e.g., [17,18]. A Banach space X is said to satisfy the UMD property if there exists a constant $c_{(1)} \geq 1$ such that for every X-valued martingale difference sequence

✉ Stefan Geiss
stefan.geiss@jyu.fi

Sonja Cox
s.g.cox@uva.nl

1 Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Postbus 94248, 1090 GE Amsterdam, Netherlands

2 Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35, 40014 Jyvaskyla, Finland

Published online: 14 March 2021

$(d_n)_{n=1}^N$ one has that

$$\left\| \sum_{n=1}^N d_n \right\|_{\mathcal{L}^2(\mathbb{P};X)} \leq c_{(1)} \left\| \sum_{n=1}^N \theta_n d_n \right\|_{\mathcal{L}^2(\mathbb{P};X)} \tag{1}$$

for all signs $\theta_n \in \{-1, 1\}$, i.e., one has Unconditional Martingale Differences. It follows from Maurey [27], that, in order to verify the UMD property of the Banach space X, it is sufficient to consider X-valued Haar- or dyadic martingales (dyadic martingales are also known as *Paley–Walsh martingales*).

On the other hand, McConnell [28, Theorem 2.2] (see also Hitczenko [15]) proved that the UMD property is equivalent to the existence of a $c_{(2)} \geq 1$ such that

$$\left\| \sum_{n=1}^N d_n \right\|_{\mathcal{L}^2(\mathbb{P};X)} \leq c_{(2)} \left\| \sum_{n=1}^N e_n \right\|_{\mathcal{L}^2(\mathbb{P};X)} \tag{2}$$

for all $N \in \mathbb{N}$ and all X-valued $(\mathcal{F}_n)_{n=1}^N$-martingale difference sequences $(d_n)_{n=1}^N$ and $(e_n)_{n=1}^N$ such that $\mathcal{L}(d_n \mid \mathcal{F}_{n-1}) = \mathcal{L}(e_n \mid \mathcal{F}_{n-1})$, i.e., $(e_n)_{n=1}^N$ and $(d_n)_{n=1}^N$ are *tangent*. Imposing additional assumptions on either $(e_n)_{n=1}^N$ or $(d_n)_{n=1}^N$ in (2) results in an (a priori) weaker Banach space property, e.g., imposing that $(e_n)_{n=1}^N$ in (2) is the *decoupled* tangent sequence of $(d_n)_{n=1}^N$ (see Definition 2.5) results in the *lower decoupling property for tangent martingales*. The Banach space L^1 satisfies the lower decoupling property for tangent martingales (see Cox and Veraar [9, Example 4.7]), but fails to have the UMD property (see, e.g., [17, Example 4.2.20]). The notion of decoupled tangent sequences was introduced by Kwapień and Woyczyński [23,24]. The decoupled tangent sequence $(e_n)_{n=1}^N$ of a sequence $(d_n)_{n\in\mathbb{N}}$ (adapted to a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$) is unique in distribution and replaces parts of the dependence structure of $(d_n)_{n=1}^N$ by a sequence of conditionally independent random variables. Although the definition of decoupling might not be explicit, there are canonical representations of a decoupled tangent sequence, see Kwapień and Woyczyński [24] and Montgomery-Smith [29].

Inequalities (1) and (2) describe *two-sided decoupling properties* due to the symmetry between the left- and right-hand side. The lower decoupling property for tangent martingales however is an example of a *one-sided decoupling property*. An a priori different one-sided decoupling property is obtained by considering (1) where $(\theta_n)_{n=1}^N$ is replaced by a Rademacher sequence that is independent of $(d_n)_{n=1}^N$. This one-sided decoupling property was first studied explicitly in [13] and is also satisfied by L^1. The goal of this article is to gain insight into the relation between these different kinds of one-sided decoupling properties. First, however, let us discuss some instances in the literature where decoupling inequalities play a crucial role.

The proofs by Burkholder [4] and Bourgain [2] of the equivalence of the UMD property of a Banach space X and the continuity of the X-valued Hilbert transform use that X has the UMD property if and only if it has both a lower- and an upper decoupling property. For certain applications, only a single one-sided decoupling property is

needed. For example, the lower (resp. upper) decoupling property for martingales and the type (resp. cotype) property imply martingale type (resp. cotype) and therefore by Pisier [31] an equivalent re-norming of the Banach space with a norm having a certain modulus of continuity (resp. convexity). A classical case of decoupling, studied on its own, concerns randomly stopped sums of independent random variables, see for example the results of Klass [21,22].

Another application for decoupling is stochastic integration. Indeed, only the lower decoupling property is needed to obtain sufficient conditions for the existence of stochastic integrals. This can be inferred from [12, the proof of Theorem 2], where stochastic integration of UMD Banach space-valued processes with respect to a Brownian motion is considered. One-sided decoupling is used more explicitly in [24, Section 6], where the existence of decoupled tangent processes for left quasi-continuous processes in the Skorokhod space is studied. In [34] and [9, Section 5], decoupling inequalities were used to give sufficient conditions for the existence of a Banach space-valued stochastic process with respect to a cylindrical Brownian motion. Very recently, Kallenberg [20] proved the existence of decoupled tangent semi-martingales and two-sided decoupling inequalities, and considered applications to multiple stochastic integrals. Moreover, quasi-Banach spaces fail to satisfy the UMD property, but may satisfy decoupling inequalities, see [7, Section 5.1] and e.g. [9, Example 4.7]. Section 5 contains our contribution to this topic, see also Theorem 1.7 and Remark 5.6. Before discussing this contribution, let us turn to the open problem that motivated our research:

Open Problem 1.1 *If a Banach space X has the lower decoupling property for tangent dyadic martingales, does it also have the lower decoupling property for general tangent martingales?*

We were not able to answer this question in this generality. However, our main result (Theorem 1.4) provides a reduction of this problem to simple Haar-type series and gives a partial answer (see Corollary 1.6 for a special case). The proof of Theorem 1.4 is inspired by the aforementioned work by Maurey [27]. Open Problem 1.1 can be split into two subproblems; consequently, our proof of Theorem 1.4 consists of two parts [completely solving subproblem (A) and partially solving subproblem (B)]:

(A) If a Banach space X has a lower (upper) decoupling property for X-valued sequences of random variables adapted to a (in a certain sense) natural minimal filtration $(\mathcal{F}_n)_{n=1}^{\infty}$ and with conditional distributions in a set of measures $\mathcal{P}$, does X also have a lower (upper) decoupling property for X-valued sequences adapted to *any* filtration $(\mathcal{F}_n)_{n=1}^{\infty}$ and with conditional distributions in $\mathcal{P}$?

(B) Given that X has a lower (resp. upper) decoupling property for X-valued sequences with conditional distributions in a certain $\mathcal{P}$, does X also have a lower (resp. upper) decoupling property for general X-valued sequences?

Problem (A) is of fundamental importance in stochastic integration theory as, given the driving process, the underlying filtration determines the set of integrands we may use. We now describe the content of the article in more detail:

Section 3: Theorem 3.1 provides a factorization of a random variable along regular conditional probabilities. With this result, we contribute to the results of

Montgomery-Smith [29] (see also Kallenberg [19, Lemma 3.22]). This result is the key to approximate our adapted processes in terms of Haar-like series.

Section 4: this section is devoted to our main results, Theorem 4.3 and its corollary, Theorem 1.4. It contains the key ingredients for the proofs and some examples.
To formulate Theorem 1.4, we recall some definitions. For a separable Banach space X, we denote by $\mathcal{B}(X)$ the σ-algebra generated by the norm-open sets. By $\mathcal{P}(X)$, we denote the set of all probability measures on $(X, \mathcal{B}(X))$, for $p \in (0, \infty)$ we let

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) \colon \int_X \|x\|_X^p \, \mathrm{d}\mu(x) < \infty \right\},$$

and $\delta_x \in \mathcal{P}(X)$ stands for the Dirac measure at $x \in X$. The next definition concerns a set of admissible adapted processes characterized by an assumption on the regular versions of the—in a sense—predictable projections:

Definition 1.2 Let X be a separable Banach space, $p \in (0, \infty)$, $\emptyset \neq \mathcal{P} \subseteq \mathcal{P}_p(X)$, and $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^\infty)$ be a stochastic basis. We denote by $\mathcal{A}_p(\Omega, (\mathcal{F}_n)_{n=0}^\infty; X, \mathcal{P})$ the set of $(\mathcal{F}_n)_{n=1}^\infty$-adapted sequences $(d_n)_{n=1}^\infty$ in $\mathcal{L}^p(\mathbb{P}; X)$ with the property that for all $n \geq 1$ there exists an $\Omega_{n-1} \in \mathcal{F}$ satisfying $\mathbb{P}(\Omega_{n-1}) = 1$ and $\kappa_{n-1}[\omega, \cdot] \in \mathcal{P}$ for all $\omega \in \Omega_{n-1}$, where κ_{n-1} is a regular conditional probability kernel for $\mathcal{L}(d_n \mid \mathcal{F}_{n-1})$.

The concept of regular conditional probability kernels is recalled in Sect. 2.2. Next, we introduce an extension of a given set of probability measures that is natural in our context:

Definition 1.3 For a separable Banach space X, $p \in (0, \infty)$ and $\emptyset \neq \mathcal{P} \subseteq \mathcal{P}_p(X)$ we let

$$\begin{aligned} \mathcal{P}_{p\text{-ext}} := \Big\{ \mu \in \mathcal{P}_p(X) \colon \forall j \geq 1 \, \exists \, K_j \geq 1 \text{ and } \mu_{j,1}, \ldots, \mu_{j,K_j} \in \mathcal{P} \\ \text{such that } \mu_{j,1} * \cdots * \mu_{j,K_j} \overset{w^*}{\to} \mu \text{ as } j \to \infty \\ \text{and } \left(\mu_{j,1} * \cdots * \mu_{j,K_j}\right)_{j \in \mathbb{N}} \text{ is uniformly } \mathcal{L}^p\text{-integrable} \Big\}. \end{aligned}$$

The convergence of the convolutions $\mu_{j,1} * \cdots * \mu_{j,K_j}$ toward μ in Definition 1.3 is known to be the convergence in the p-Wasserstein distance if $p \in [1, \infty)$ (cf. [6, Theorem 5.5, p. 358]).

Theorem 1.4 *Let X, Y, Z be Banach spaces, where X is separable, let $S : X \to Y$ and $T : X \to Z$ be linear and bounded, $p \in (0, \infty)$, Δ an index set, let $\Psi : [0, \infty) \to [0, \infty)$ be upper semi-continuous, and let $\Psi_\lambda : [0, \infty) \to [0, \infty)$, $\lambda \in \Delta$, be a family of lower semi-continuous functions such that*

$$\sup_{\xi \in (0,\infty)} (1 + |\xi|)^{-p} \Psi(\xi) < \infty \quad \text{and} \quad \sup_{\xi \in (0,\infty)} (1 + |\xi|)^{-p} \Psi_\lambda(\xi) < \infty \tag{3}$$

for all $\lambda \in \Delta$. Then, for a set $\mathcal{P} \subseteq \mathcal{P}_p(X)$ with $\delta_0 \in \mathcal{P}$, the following assertions are equivalent:

(i) *For every stochastic basis* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=1}^{\infty})$ *and finitely supported*[1] $(d_n)_{n=1}^{\infty} \in \mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P}_{p\text{-ext}})$, *it holds that*

$$\sup_{\lambda \in \Delta} \mathbb{E}\Psi_\lambda \left(\left\| \sum_{n=1}^{\infty} S d_n \right\|_Y \right) \leq \mathbb{E}\Psi \left(\left\| \sum_{n=1}^{\infty} T e_n \right\|_Z \right), \tag{4}$$

whenever $(e_n)_{n\in\mathbb{N}}$ *is an* $\mathbb{F}$*-decoupled tangent sequence of* $(d_n)_{n\in\mathbb{N}}$.

(ii) *For every sequence of independent random variables* $(\varphi_n)_{n=1}^{N} \subset \mathcal{L}^p(\mathbb{P}; X)$, $N \geq 1$, *satisfying* $\mathcal{L}(\varphi_n) \in \mathcal{P}$, *and every* $A_0 \in \{\emptyset, \Omega\}$, $A_n \in \sigma(\varphi_1, \ldots, \varphi_n)$, $n \in \{1, \ldots, N\}$, *it holds that*

$$\sup_{\lambda \in \Delta} \mathbb{E}\Psi_\lambda \left(\left\| \sum_{n=1}^{N} 1_{A_{n-1}} S \varphi_n \right\|_Y \right) \leq \mathbb{E}\Psi \left(\left\| \sum_{n=1}^{N} 1_{A_{n-1}} T \varphi_n' \right\|_Z \right), \tag{5}$$

where $(\varphi_n')_{n=1}^{N}$ *is an independent copy of* $(\varphi_n)_{n=1}^{N}$.

Some remarks concerning Theorem 1.4 are at place:

(1) The condition that $\delta_0 \in \mathcal{P}$ ensures that finitely supported sequences fit in our setting and is used at several instances in the proof.
(2) The condition that X is separable is mainly to simplify our presentation: after all, we can apply Theorem 1.4 whenever $(d_n)_{n\in\mathbb{N}}$ is a sequence of random variables taking values in a separable subspace X of some non-separable space $\tilde{X}$.
(3) The table below provides some typical choices for Δ, Ψ, and $(\Psi_\lambda)_{\lambda\in\Delta}$ given $p \in (0, \infty)$ (here $C \in (0, \infty)$ and f, g are $\mathbb{R}$-valued random variables):

Δ	$\Psi_\lambda(\xi)$	$\Psi(\xi)$	$\sqrt[p]{\sup_{\lambda\in\Delta} \mathbb{E}\Psi_\lambda(f)} \leq \sqrt[p]{\mathbb{E}\Psi(g)}$
$\mathrm{card}(\Delta) = 1$	ξ^p	$C^p \xi^p$	$\lVert f \rVert_{\mathcal{L}^p(\mathbb{P})} \leq C \lVert g \rVert_{\mathcal{L}^p(\mathbb{P})}$
$\mathrm{card}(\Delta) = 1$	$1_{\{\xi > \mu\}}$, $\mu \geq 0$	$C^p \xi^p$	$\sqrt[p]{\mathbb{P}(f > \mu)} \leq C \lVert g \rVert_{\mathcal{L}^p(\mathbb{P})}$
$(0, \infty)$	$\lambda^p 1_{\{\xi > \lambda\}}$	$C^p \xi^p$	$\lVert f \rVert_{\mathcal{L}^{p,\infty}(\mathbb{P})} \leq C \lVert g \rVert_{\mathcal{L}^p(\mathbb{P})}$

(4) For relevant choices for $\mathcal{P}$, see Examples 4.5–4.8; Corollary 1.6 uses Example 4.8.
(5) Theorem 1.4 remains valid if one exchanges $(d_n)_{n=1}^{N}$ with $(e_n)_{n=1}^{N}$ in (4) and $(\varphi_n)_{n=1}^{N}$ with $(\varphi_n')_{n=1}^{N}$ in (5), respectively.

Section 5: we use Theorem 1.4 to obtain relevant upper bounds for stochastic integrals, see Theorem 1.7. In order to formulate that theorem, we need the following definition:

[1] There are only finitely many n for which $d_n \not\equiv 0$.

Definition 1.5 If $p \in (0, \infty)$ and if X is a separable Banach space X, then we let $D_p(X) := \inf c$, where the infimum is taken over all $c \in [0, \infty]$ such that

$$\left\| \sum_{n=1}^{N} r_n v_{n-1} \right\|_{\mathcal{L}^p(\mathbb{P};X)} \leq c \left\| \sum_{n=1}^{N} r'_n v_{n-1} \right\|_{\mathcal{L}^p(\mathbb{P};X)} \tag{6}$$

for all $N \geq 2$, $v_0 \in X$ and $v_n := h_n(r_1, \ldots, r_n)$ with $h_n \colon \{-1, 1\}^n \to X$, $n \in \{1, \ldots, N-1\}$, where the $(r_n)_{n=1}^N$ are independent and take the values -1 and 1 with probability $1/2$, and $(r'_n)_{n=1}^N$ is an independent copy of $(r_n)_{n=1}^N$.

The process $(\sum_{k=1}^n r_k v_{k-1})_{n=1}^N$ in Definition 1.5 is a dyadic martingale. Theorem 1.4 with $X = Y = Z$, $S = T = \mathrm{Id}$, $\Delta = \{\lambda\}$, and $\Psi(\xi) = \Psi_\lambda(\xi) = \xi^p$ implies:

Corollary 1.6 *Let X be a separable Banach space, let $p \in (0, \infty)$ be such that $D_p(X) < \infty$, and let $\mathcal{P} := \{\frac{1}{2}(\delta_{-x} + \delta_{-x}) \colon x \in X\}$. Then,*

$$\left\| \sum_{n=1}^{N} d_n \right\|_{\mathcal{L}^p(\mathbb{P};X)} \leq D_p(X) \left\| \sum_{n=1}^{N} e_n \right\|_{\mathcal{L}^p(\mathbb{P};X)} \tag{7}$$

for all $N \geq 1$, every stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^N)$, every strongly p-integrable and $(\mathcal{F}_n)_{n=1}^N$ adapted sequence of random variables $(d_n)_{n=1}^N$ such that, on a set of measure one, the $\mathcal{F}_{n-1}$-conditional laws of all d_n belong to $\mathcal{P}_{p\text{-ext}}$, and every decoupled tangent sequence $(e_n)_{n=1}^N$ of $(d_n)_{n=1}^N$.

Cox and Geiss [8, Section 5] contains a characterization of $\mathcal{P}_{p\text{-ext}}$ when $\mathcal{P} := \{\frac{1}{2}(\delta_{-x} + \delta_{-x}) \colon x \in \mathbb{R}\}$. Corollary 1.6 combined with the Central Limit Theorem results in Theorem 1.7. For details, see the proof of Part (ii) of Theorem 5.2. Theorem 1.7 extends both [9, Theorem 5.4] and [12, Theorem 2]: see Remark 5.6 for details.

Theorem 1.7 *For a separable Banach space X and $p, q \in (0, \infty)$, the following assertions are equivalent:*

(i) $D_p(X) < \infty$.

(ii) *For every stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = (\mathcal{F}_t)_{t \in [0,\infty)})$, every $\mathbb{F}$-Brownian motion $W = (W(t))_{t \in [0,\infty)}$, and every simple $\mathbb{F}$-predictable X-valued process $(H(t))_{t \in [0,\infty)}$ it holds that*

$$\left\| \int_0^\infty H(t) \mathrm{d}W(t) \right\|_{\mathcal{L}^q(\mathbb{P};X)} \leq K_{p,2} D_p(X) \, \|S(H)\|_{\mathcal{L}^q(\mathbb{P})}$$

with the square function $S(H)(\omega) := \|f \mapsto \int_0^\infty f(t) H(t, \omega) \mathrm{d}t\|_{\gamma(\mathcal{L}^2((0,\infty);X)}$ and $K_{p,2}$ the constant in the $\mathcal{L}^p$-to-$\mathcal{L}^2$ Kahane–Khintchine inequality (see Sect. 5.1 for details on the γ-radonifying norm $\gamma(\mathcal{L}^2((0, \infty); X))$.

2 Preliminaries

2.1 Some General Notation

We let $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. For a vector space V and $B \subseteq V$, we set $-B := \{x \in V\colon -x \in B\}$. Given a non-empty set Ω, we let 2^Ω denote the system of all subsets of Ω and use $A \triangle B := (A \backslash B) \cup (B \backslash A)$ for $A, B \in 2^\Omega$. A system of pair-wise disjoint subsets $(A_i)_{i \in I}$ of Ω is a *partition* of Ω, where I is an arbitrary index set and $A_i = \emptyset$ is allowed, if $\bigcup_{i \in I} A_i = \Omega$. If (M, d) is a metric space we define $d\colon M \times 2^M \to [0, \infty]$ by setting $d(x, A) := \inf\{d(x, y)\colon y \in A\}$ for all $(x, A) \in M \times 2^M$. If V is a Banach space and (M, d) a metric space, then $C(M; V)$ is the space of continuous maps from M to V, and $C_b(M; V)$ the subspace of bounded continuous maps from M to V.

Banach Space-Valued random variables: For a Banach space X, we let $\mathcal{B}(X)$ denote the Borel σ-algebra generated by the norm-open sets. For $x \in X$ and $\varepsilon > 0$, we set $B_{x,\varepsilon} := \{y \in X\colon \|x - y\|_X < \varepsilon\}$. For $B \in \mathcal{B}(X)$, we let $\bar{B}$ denote the norm-closure of B, we let B^o denote the interior and $\partial B := \bar{B} \backslash B^o$. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable space (S, Σ), an $\mathcal{F}/\Sigma$-measurable mapping $\xi\colon \Omega \to S$ is called an *S-valued random variable*. For a random variable $\xi : \Omega \to S$, the law of ξ is denoted by $\mathcal{L}(\xi)(A) := \mathbb{P}(\xi \in A)$ for $A \in \Sigma$.

Lebesgue spaces: For X a separable Banach space and (S, Σ) a measurable space, we define $\mathcal{L}^0((S, \Sigma); X)$ to be the space of $\Sigma/\mathcal{B}(X)$-measurable mappings from S to X. If (S, Σ) is equipped with a σ-finite measure μ and $p \in (0, \infty)$, then we define

$$\mathcal{L}^p((S, \Sigma, \mu); X) := \left\{ \xi \in \mathcal{L}^0((S, \Sigma, \mu); X)\colon \|\xi\|^p_{\mathcal{L}^p((S,\Sigma,\mu);X)} := \int_S \|\xi\|^p_X \, \mathrm{d}\mu < \infty \right\}.$$

If there is no risk of confusion, we write for example $\mathcal{L}^p(\mu; X)$ or $\mathcal{L}^p(\Sigma; X)$ as shorthand notation for $\mathcal{L}^p((S, \Sigma, \mu); X)$, and we set $\mathcal{L}^p((S, \Sigma, \mu)) := \mathcal{L}^p((S, \Sigma, \mu); \mathbb{R})$.

Probability measures on Banach spaces:

(1) Given an index set $I \neq \emptyset$, a family $(\mu_i)_{i \in I} \subseteq \mathcal{P}_p(X)$ ($\mathcal{P}_p(X)$ was introduced in Sect. 1) is *uniformly $\mathcal{L}^p$-integrable* if

$$\lim_{K \to \infty} \sup_{i \in I} \int_{\{\|x\|_X \geq K\}} \|x\|^p_X \mathrm{d}\mu_i(x) = 0.$$

Accordingly, a family of X-valued random variables $(\xi_i)_{i \in I}$ is *uniformly $\mathcal{L}^p$-integrable* if $(\mathcal{L}(\xi_i))_{i \in I}$ is uniformly $\mathcal{L}^p$-integrable.

(2) For $\mu \in \mathcal{P}(X)$ and $\mu_n \in \mathcal{P}(X)$, $n \in \mathbb{N}$, we write $\mu_n \overset{w^*}{\to} \mu$ as $n \to \infty$ if μ_n converges weakly to μ, i.e., if $\lim_{n \to \infty} \int_X f(x) \, \mathrm{d}\mu_n(x) = \int_X f(x) \, \mathrm{d}\mu(x)$ for all $f \in C_b(X; \mathbb{R})$. Moreover, for a sequence of X-valued random variables $(\xi_n)_{n \in \mathbb{N}}$ and an X-valued random variable ξ (possibly defined on different probability spaces) we write $\xi_n \overset{w^*}{\to} \xi$ as $n \to \infty$ provided that $\mathcal{L}(\xi_n) \overset{w^*}{\to} \mathcal{L}(\xi)$ as $n \to \infty$.

We shall frequently use the following well-known result, which relates $\mathcal{L}^p$-uniform integrability and convergence of moments:

Lemma 2.1 *Let $p \in (0, \infty)$, let X be a separable Banach space, and let μ, $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_p(X)$ such that $\mu_n \xrightarrow{w^*} \mu$. Then, the following are equivalent:*

(i) $\int_X \|x\|^p \, \mathrm{d}\mu_n \to \int_X \|x\|^p \, \mathrm{d}\mu$.
(ii) $(\mu_n)_{n\in\mathbb{N}}$ *is uniformly $\mathcal{L}^p$-integrable.*

Proof Apply, e.g., [19, Lemma 4.11 (in (5) lim sup can be replaced by sup)] to the random variables $\xi, \xi_1, \xi_2, \ldots$ where $\xi = \|\zeta\|_X^p$ and $\xi_n = \|\zeta_n\|_X^p$, and where $\mathcal{L}(\zeta) = \mu$ and $\mathcal{L}(\zeta_n) = \mu_n$. □

STOCHASTIC BASIS: We use the notion of a *stochastic basis* $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, which is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$, $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}$, and where we set $\mathcal{F}_\infty := \sigma\left(\bigcup_{n\in\mathbb{N}_0} \mathcal{F}_n\right)$. For measurable spaces $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$, and $\xi = (\xi_n)_{n\in\mathbb{N}}$ a sequence of S-valued random variables on $(\Omega, \mathcal{F})$, we let $\mathbb{F}^\xi = (\mathcal{F}_n^\xi)_{n\in\mathbb{N}_0}$ denote the natural filtration generated by ξ, i.e., $\mathcal{F}_0^\xi := \{\emptyset, \Omega\}$ and $\mathcal{F}_n^\xi := \sigma(\xi_1, \ldots, \xi_n)$ for $n \in \mathbb{N}$, and $\mathcal{F}_\infty^\xi := \sigma(\xi_n : n \in \mathbb{N})$.

2.2 Stochastic Kernels

We provide some details for regular versions of conditional probabilities we shall use later.

Definition 2.2 Let X be a separable Banach space and (S, Σ) a measurable space. A mapping $\kappa\colon S \times \mathcal{B}(X) \to [0, 1]$ is a $\Sigma/\mathcal{B}(X)$-measurable kernel if and only if the following two conditions hold:

(i) For all $\omega \in S$, it holds that $\kappa[\omega, \cdot] \in \mathcal{P}(X)$.
(ii) For all $B \in \mathcal{B}(X)$, the map $\omega \to \kappa[\omega, B]$ is $\Sigma/\mathcal{B}(\mathbb{R})$-measurable.

Remark 2.3 Let the space (S, Σ) be equipped with a probability measure $\mathbb{P}$ and let $\Pi \subseteq \mathcal{B}(X)$ be a countable π-system that generates $\mathcal{B}(X)$. For two kernels $\kappa, \kappa'\colon S \to \mathcal{P}(X)$, the following assertions are equivalent:

(i) $\kappa[\omega, B] = \kappa'[\omega, B]$ for $\mathbb{P}$-almost all $\omega \in S$, for all $B \in \Pi$.
(ii) $\kappa[\omega, \cdot] = \kappa'[\omega, \cdot]$ for $\mathbb{P}$-almost all $\omega \in S$.

We need the existence of kernels describing conditional probabilities:

Theorem 2.4 [19, Theorem 6.3] *Let X be a separable Banach space, $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space, $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-algebra, and let $\xi : \Omega \to X$ be a random variable. Then, there is a $\mathcal{G}/\mathcal{B}(\mathcal{P}(X))$-measurable kernel $\kappa\colon \Omega \to \mathcal{P}(X)$ satisfying*

$$\kappa[\cdot, B] = \mathbb{P}(\xi \in B \mid \mathcal{G}) \quad a.s.$$

for all $B \in \mathcal{B}(X)$. If $\kappa' : \Omega \to \mathcal{P}(X)$ is another kernel with this property, then $\kappa' = \kappa$ a.s.

We refer to κ as a *regular conditional probability kernel* for $\mathcal{L}(\xi \mid \mathcal{G})$.

2.3 Decoupling

We briefly recall the concept of decoupled tangent sequences as introduced by Kwapień and Woyczyński [24]. For more details, we refer to [10,25] and the references therein.

Definition 2.5 Let X be a separable Banach space, let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n\in\mathbb{N}_0})$ be a stochastic basis, and let $(d_n)_{n\in\mathbb{N}}$ be an $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted sequence of X-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. A sequence of X-valued and $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted random variables $(e_n)_{n\in\mathbb{N}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called an $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$*-decoupled tangent sequence* of $(d_n)_{n\in\mathbb{N}}$ provided there exists a σ-algebra $\mathcal{H} \subseteq \mathcal{F}$ satisfying $\sigma((d_n)_{n\in\mathbb{N}}) \subseteq \mathcal{H}$ such that the following two conditions are satisfied:

(i) Tangency: For all $n \in \mathbb{N}$ and all $B \in \mathcal{B}(X)$, one has

$$\mathbb{P}(d_n \in B \mid \mathcal{F}_{n-1}) = \mathbb{P}(e_n \in B \mid \mathcal{F}_{n-1}) = \mathbb{P}(e_n \in B \mid \mathcal{H}) \quad \text{a.s.}$$

(ii) Conditional independence: For all $N \in \mathbb{N}$ and $B_1, \dots, B_N \in \mathcal{B}(X)$ one has

$$\mathbb{P}(e_1 \in B_1, \dots, e_N \in B_N \mid \mathcal{H}) = \mathbb{P}(e_1 \in B_1 \mid \mathcal{H}) \dots \mathbb{P}(e_N \in B_N \mid \mathcal{H}) \quad \text{a.s.}$$

A construction of a decoupled tangent sequence is presented in [25, Section 4.3].

Example 2.6 Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n\in\mathbb{N}_0})$ be a stochastic basis, $(\varphi_n)_{n\in\mathbb{N}}$ and $(\varphi_n')_{n\in\mathbb{N}}$ two independent and identically distributed sequences of independent, $\mathbb{R}$-valued random variables such that φ_n and φ_n' are $\mathcal{F}_n$-measurable and independent of $\mathcal{F}_{n-1}$ for all $n \in \mathbb{N}$, and let $(v_n)_{n\in\mathbb{N}_0}$ be an $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$-adapted sequence of X-valued random variables independent of $(\varphi_n')_{n\in\mathbb{N}}$. Then, $(\varphi_n' v_{n-1})_{n\in\mathbb{N}}$ is an $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$-decoupled tangent sequence of $(\varphi_n v_{n-1})_{n\in\mathbb{N}}$, where one may take

$$\mathcal{H} := \sigma((\varphi_n)_{n\in\mathbb{N}}, (v_n)_{n\in\mathbb{N}_0}).$$

Similarly, $(\varphi_n)_{n\in\mathbb{N}}$ and $(\varphi_n')_{n\in\mathbb{N}}$ could be X-valued random variables and $(v_n)_{n\in\mathbb{N}_0}$ $\mathbb{R}$-valued.

3 A Factorization for Regular Conditional Probabilities

Theorem 3.1 extends [19, Lemma 3.22] and can be viewed as a strong version of Montgomery-Smith's distributional result [29, Theorem 2.1]. Theorem 3.1 is used to prove Theorem 4.3, where it yields a refined argument for the existence of a decoupled tangent sequence. In this sense, it also contributes to [24] (cf. [10, Proposition 6.1.5]).

Theorem 3.1 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra, let $d \in \mathcal{L}^0(\mathcal{F}; \mathbb{R})$ satisfy $d(\Omega) \subseteq [0, 1)$, and let $\kappa\colon \Omega \times \mathcal{B}([0, 1)) \to [0, 1]$ be a regular conditional probability kernel for $\mathcal{L}(d \mid \mathcal{G})$. Let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}) := (\Omega \times (0, 1], \mathcal{F} \otimes \mathcal{B}((0, 1]), \mathbb{P} \otimes \lambda)$, where λ is the Lebesgue measure on $\mathcal{B}((0, 1])$. Set $[0, 0) := \emptyset$ and*

define $H\colon \bar{\Omega} \to [0,1]$, $d^0\colon \Omega \times [0,1] \to [0,1]$ *by*

$$H(\omega, s) := \kappa[\omega, [0, d(\omega))] + s\kappa[\omega, \{d(\omega)\}], \tag{8}$$

$$d^0(\omega, h) := \inf\{x \in [0,1]\colon \kappa[\omega, [0, x]] \geq h\}. \tag{9}$$

Then,

(i) H *is* $\bar{\mathcal{F}}/\mathcal{B}([0,1])$*-measurable, independent of* $\mathcal{G} \otimes \{\emptyset, (0,1]\}$*, and uniformly* $[0,1]$ *distributed,*
(ii) d^0 *is* $\mathcal{G} \otimes \mathcal{B}([0,1])/\mathcal{B}([0,1])$*-measurable, and*
(iii) *there is an* $\mathcal{N} \in \mathcal{F}$ *with* $\mathbb{P}(\mathcal{N}) = 0$ *such that* $d^0(\omega, H(\omega, s)) = d(\omega)$ *for all* $(\omega, s) \in (\Omega \backslash \mathcal{N}) \times (0,1]$.

Before we prove this theorem, let us comment on Item (i). There are two extreme cases. The first one is $\mathcal{G} := \{\emptyset, \Omega\}$. In this case, we get that $\kappa[\omega; (-\infty, x]]$ is the distribution function of the law of d and here it is known that the distribution of H is the uniform distribution on $[0,1]$. The other extreme case is $\mathcal{G} = \mathcal{F}$ and here we can take $\kappa[\omega, B] := 1_{\{d(\omega) \in B\}}$ which implies that $H(\omega, s) = s$. Our result interpolates between these two extreme cases.

Proof of Theorem 3.1 (i) For all $n \in \mathbb{N}$ and $\ell \in \{1, \ldots, 2^n\}$, let $A_{n,\ell} := [(\ell - 1)2^{-n}, \ell 2^{-n})$. Define $H_n\colon \bar{\Omega} \to [0,1]$ by

$$H_n(\omega, s) := \sum_{\ell=1}^{2^n} 1_{\{d \in A_{n,\ell}\}}(\omega)\big(\kappa[\omega, [0, (\ell-1)2^{-n})] + s\kappa[\omega, A_{n,\ell}]\big),$$

so that for all $(\omega, s) \in \bar{\Omega}$ it holds that

$$|H_n(\omega, s) - H(\omega, s)| \leq \sum_{\ell=1}^{2^n} 1_{\{d \in A_{n,\ell}\}}(\omega)(1+s)\kappa[\omega, A_{n,\ell} \backslash \{d(\omega)\}] \to 0 \quad \text{as} \quad n \to \infty.$$

The H_n are $\bar{\mathcal{F}}/\mathcal{B}([0,1])$-measurable, so H is as point-wise limit (the measurability of H can be seen directly as well). Let $n \in \mathbb{N}$, $G \in \mathcal{G}$ and $B \in \mathcal{B}([0,1])$. Because $b\int_0^1 1_{\{a+sb \in B\}}\,\mathrm{d}s = \lambda(B \cap [a, a+b])$ for $a, b \in [0,1]$ with $a + b \leq 1$ (where λ denotes the Lebesgue measure), we get

$$\begin{aligned}\bar{\mathbb{P}}((G \times (0,1]) \cap \{H_n \in B\}) &= \sum_{\ell=1}^{2^n} \int_G \lambda\big(B \cap \big[\kappa[\cdot, [0, (\ell-1)2^{-n})], \kappa[\cdot, [0, \ell 2^{-n})]\big]\big)\,\mathrm{d}\mathbb{P} \\ &= \mathbb{P}(G) \cdot \lambda(B).\end{aligned}$$

This proves that H_n is uniformly $[0,1]$ distributed and independent of $\mathcal{G} \otimes \{\emptyset, (0,1]\}$ for all $n \in \mathbb{N}$. This completes the proof of (i), as H is the point-wise limit of $(H_n)_{n \in \mathbb{N}}$ (two $\mathbb{R}$-valued random variables ξ_1, ξ_2 are independent if and only if for all $f, g \in C_b(\mathbb{R})$ it holds that $\mathbb{E}[f(\xi_1)g(\xi_2)] = \mathbb{E}[f(\xi_1)]\mathbb{E}[f(\xi_2)]$).

(ii) For all $x \in [0, 1]$, note that

$$\{d^0 \leq x\} = \{(\omega, h) \in \Omega \times [0, 1] \colon \kappa[\omega, [0, x]] - h \geq 0\} \in \mathcal{G} \otimes \mathcal{B}([0, 1]). \quad (10)$$

(iii) It follows from (10) and the definition of H that we have, for all $x \in [0, 1]$, that

$$\begin{aligned}&\{(\omega, s) \in \bar{\Omega} \colon d^0(\omega, H(\omega, s)) \leq x\}\\ &= \{(\omega, s) \in \bar{\Omega} \colon \kappa[\omega, [0, x]] \geq \kappa[\omega, [0, d(\omega))] + s\kappa[\omega, \{d(\omega)\}]\}\end{aligned}$$

can be written as $B_x \times (0, 1]$ for some unique $B_x \in \mathcal{F}$ and that we have that

$$B_x \times (0, 1] \supseteq \{(\omega, s) \in \bar{\Omega} \colon d(\omega) \leq x\} =: C_x \times (0, 1].$$

On the other hand from the fact that the image measure of the map $(\omega, s) \mapsto (\omega, H(\omega, s))$ as a map from $\overline{\Omega}$ into $\Omega \times [0, 1]$ equals $\mathbb{P} \otimes \lambda$, we obtain, for all $x \in [0, 1]$, that

$$\begin{aligned}\mathbb{P}(B_x) = \overline{\mathbb{P}}(B_x \times (0, 1]) &= \mathbb{E} \int_0^1 1_{\{d^0(\omega, h) \leq x\}} \, \mathrm{d}h \, \mathrm{d}\mathbb{P}(\omega)\\ &= \mathbb{E} \int_0^1 1_{\{\kappa[\omega, [0, x]] \geq h\}} \, \mathrm{d}h \, \mathrm{d}\mathbb{P}(\omega) = \mathbb{E}\kappa[\cdot, [0, x]] = \mathbb{P}(C_x).\end{aligned}$$

It follows that $\mathbb{P}(B_x \backslash C_x) = 0$ for all $x \in [0, 1]$. Let $\mathcal{N} := \cup_{q \in \mathbb{Q} \cap [0, 1)} (B_q \backslash C_q)$ so that $\mathbb{P}(\mathcal{N}) = 0$. Then, observing that $B_x = \cap_{q \in \mathbb{Q} \cap [x, 1)} B_q$ (this follows from $B_x \times (0, 1] = \{d^0(\cdot, H(\cdot, \cdot)) \leq x\} = \cap_{q \in \mathbb{Q} \cap [x, 1)} \{d^0(\cdot, H(\cdot, \cdot)) \leq q\} = (\cap_{q \in \mathbb{Q} \cap [x, 1)} B_q) \times (0, 1]$ and the uniqueness of the sets B_r, $r \in [0, 1])$ and $C_x = \cap_{q \in \mathbb{Q} \cap [x, 1)} C_q$ for all $x \in [0, 1)$, we have for all $(\omega, s) \in (\Omega \backslash \mathcal{N}) \times (0, 1]$ that $d^0(\omega, H(\omega, s)) = d(\omega)$. □

Corollary 3.2 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ a σ-algebra, X a separable Banach space, $d \in \mathcal{L}^0(\mathcal{F}; X)$. Let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}) := (\Omega \times (0, 1], \mathcal{F} \otimes \mathcal{B}((0, 1]), \mathbb{P} \otimes \lambda)$, where λ is the Lebesgue measure on $\mathcal{B}((0, 1])$. Then, there exist random variables $H \colon \bar{\Omega} \to [0, 1]$, $d^0 \colon \Omega \times [0, 1] \to X$ such that*

(i) *H is uniformly $[0, 1]$ distributed and independent of $\mathcal{G} \otimes \{\emptyset, [0, 1]\}$,*
(ii) *d^0 is $\mathcal{G} \otimes \mathcal{B}([0, 1])/\mathcal{B}(X)$-measurable, and*
(iii) *there is an $\mathcal{N} \in \mathcal{F}$ with $\mathbb{P}(\mathcal{N}) = 0$ such that $d(\omega) = d^0(\omega, H(\omega, s))$ for all $(\omega, s) \in (\Omega \backslash \mathcal{N}) \times (0, 1]$.*

Proof This is an immediate consequence of Theorem 3.1 and the fact that that X is Borel-isomorphic to $[0, 1)$, see, e.g., [11, Theorem 13.1.1]. □

4 A Reduction of General Decoupling to Haar-Type Series

Before we turn to our main Theorem 4.3, we discuss some properties of the extension of $\mathcal{P}$ to $\mathcal{P}_{p\text{-ext}}$ (see Definition 1.3). For this, we need

Lemma 4.1 *Assume a metric space (M, d) and a continuous map $* : M \times M \to M$ with $(x * y) * z = x * (y * z)$ for $x, y, z \in M$. Let $\emptyset \neq \mathcal{P} \subseteq M$ and*

$$\overline{\mathcal{P}}^* := \mathrm{cl}_d(\{x_1 * \cdots * x_L : x_1, \ldots, x_L \in \mathcal{P}, L \in \mathbb{N}\})$$

where the closure on the right side is taken with respect to d. Then, one has $\overline{(\overline{\mathcal{P}}^*)}^* = \overline{\mathcal{P}}^*$ and $\overline{\mathcal{P}}^*$ is the smallest d-closed set $\mathcal{Q}$ with $\mathcal{Q} \supseteq \mathcal{P}$ and $\mu * \nu \in \mathcal{Q}$ for all $\mu, \nu \in \mathcal{Q}$.

Proof The equality $\overline{(\overline{\mathcal{P}}^*)}^* = \overline{\mathcal{P}}^*$ follows from the continuity of $*$ and a standard diagonalization procedure. This also implies that $\mu * \nu \in \overline{\mathcal{P}}^*$ for all $\mu, \nu \in \overline{\mathcal{P}}^*$. Now let us assume a set $\mathcal{Q}$ as in the assertion. Then, $x_1 * \cdots * x_L \in \mathcal{Q}$ for all $x_1, \ldots, x_L \in \mathcal{P}$. As $\mathcal{Q}$ is closed we deduce $\overline{\mathcal{P}}^* \subseteq \mathcal{Q}$. □

Lemma 4.2 reveals some basic properties of $\mathcal{P}_{p\text{-ext}}$. To this end, for $p \in (0, \infty)$ we introduce on $\mathcal{P}_p(X) \subseteq \mathcal{P}(X)$ the metric

$$d_p(\mu, \nu) := d_0(\mu, \nu) + \left| \int_X \|x\|^p \, \mathrm{d}\mu(x) - \int_X \|x\|^p \, \mathrm{d}\nu(x) \right| \tag{11}$$

where d_0 is a fixed metric on $\mathcal{P}(X)$ that metricizes the w^*-convergence, see for example [30, Theorem II.6.2].

Lemma 4.2 *Let X be a separable Banach space, $p \in (0, \infty)$, and let $\mathcal{P} \subseteq \mathcal{P}_p(X)$ be non-empty. Then,*

(i) $(\mathcal{P}_{p\text{-ext}})_{p\text{-ext}} = \mathcal{P}_{p\text{-ext}}$ *and*
(ii) $\mathcal{P}_{p\text{-ext}}$ *is the smallest d_p-closed set $\mathcal{Q}$ with $\mathcal{Q} \supseteq \mathcal{P}$ and $\mu * \nu \in \mathcal{Q}$ for all $\mu, \nu \in \mathcal{Q}$.*

Proof We will verify that the convolution is continuous with respect to d_p, the assertion then follows from Lemma 4.1. To verify this, we let $\mu, \nu, \mu_n, \nu_n \in \mathcal{P}_p(X)$, $n \in \mathbb{N}$, such that $\lim_{n\to\infty} d_p(\mu, \mu_n) = \lim_{n\to\infty} d_p(\nu, \nu_n) = 0$. It is know that $\mu_n * \nu_n \xrightarrow{w^*} \mu * \nu$ as well (one can use [19, Theorem 4.30]). Because for $K > 0$ we have, with $h(x, y) := \max\{\|x\|_X, \|y\|_X\}$,

$$\begin{aligned}\int_{\{\|x+y\|_X \geq K\}} \|x + y\|_X^p \mathrm{d}\mu_n(x)\mathrm{d}\nu_n(y) &\leq 2^p \int_{\{h(x,y)\geq K/2\}} h^p(x, y)\mathrm{d}\mu_n(x)\mathrm{d}\nu_n(y) \\ &\leq 2^p \int_{\{\|x\|_X \geq K/2\}} \|x\|_X^p \mathrm{d}\mu_n(x) + 2^p \int_{\{\|y\|_X \geq K/2\}} \|y\|_X^p \mathrm{d}\nu_n(y),\end{aligned}$$

cf. [1, p. 217], by Lemma 2.1 we get that $\mu_n * \nu_n$ is uniformly $\mathcal{L}^p$-integrable and thus, again by Lemma 2.1, we obtain the convergence of the p-th moments. □

Now we formulate the main result of this section. See Definition 1.2 for the definition of $\mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P}_{p\text{-ext}})$.

Theorem 4.3 *Let X be a separable Banach space and let $\Phi_\lambda \in C(X \times X; \mathbb{R})$, $\lambda \in \Delta$, for an arbitrary non-empty index set Δ. Suppose that there exist a $p \in (0, \infty)$ and constants $C_\lambda \in (0, \infty)$, $\lambda \in \Delta$, such that*

$$|\Phi_\lambda(x, y)| \leq C_\lambda(1 + \|x\|_X^p + \|y\|_X^p)$$

for all $(x, y) \in X \times X$, and let $\mathcal{P} \subseteq \mathcal{P}_p(X)$ with $\delta_0 \in \mathcal{P}$. Then, the following assertions are equivalent:

(i) *For every stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ with $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ and every finitely supported*[2] *$(d_n)_{n\in\mathbb{N}} \in \mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P}_{p\text{-ext}})$ it holds that*

$$\sup_{\lambda\in\Delta} \mathbb{E}\Phi_\lambda\left(\sum_{n=1}^{\infty} d_n, \sum_{n=1}^{\infty} e_n\right) \leq 0, \tag{12}$$

provided that $(e_n)_{n\in\mathbb{N}}$ is an $\mathbb{F}$-decoupled tangent sequence of $(d_n)_{n\in\mathbb{N}}$.

(ii) *For every probability space $(\Omega, \mathcal{F}, \mathbb{P})$, every finitely supported sequence of independent random variables $\varphi = (\varphi_n)_{n\in\mathbb{N}}$ in $\mathcal{L}^p(\mathbb{P}; X)$ satisfying $\mathcal{L}(\varphi_n) \in \mathcal{P}$ for all $n \in \mathbb{N}$, and every $A_n \in \mathcal{F}_n^\varphi$, $n \in \mathbb{N}_0$, it holds that*

$$\sup_{\lambda\in\Delta} \mathbb{E}\Phi_\lambda\left(\sum_{n=1}^{\infty} \varphi_n 1_{A_{n-1}}, \sum_{n=1}^{\infty} \varphi_n' 1_{A_{n-1}}\right) \leq 0, \tag{13}$$

where $(\varphi_n')_{n\in\mathbb{N}}$ is an independent copy of $(\varphi_n)_{n\in\mathbb{N}}$.

Proof Proof of (i) $\Rightarrow$ (ii). In (ii), we have $(1_{A_{n-1}}\varphi_n)_{n\in\mathbb{N}} \in \mathcal{A}_p(\Omega, \mathbb{F}^{\varphi,\varphi'}; X, \mathcal{P})$ with $\mathbb{F}^{\varphi,\varphi'} = (\mathcal{F}_n^{\varphi,\varphi'})_{n\in\mathbb{N}_0}$ where $\mathcal{F}_0^{\varphi,\varphi'} := \{\emptyset, \Omega\}$ and $\mathcal{F}_n^{\varphi,\varphi'} := \sigma(\varphi_1, \varphi_1', \ldots, \varphi_n, \varphi_n')$ for $n \in \mathbb{N}$. Therefore, the implication (ii) $\Rightarrow$ (i) follows by Example 2.6. □

The implication (ii) $\Rightarrow$ (i) will be proved in Appendix A. Theorem 4.3 allows us to prove Theorem 1.4 from Sect. 1:

Proof of Theorem 1.4 The statement for general Δ follows from the case $\Delta = \{\lambda_0\}$ so that we may assume this case and let $\underline{\Psi} := \Psi_{\lambda_0}$ and $\overline{\Psi} := \Psi$. By the lower and upper semi-continuity, we can find continuous $\underline{\Psi}^\ell, \overline{\Psi}^\ell : [0, \infty) \to [0, \infty)$, $\ell \in \mathbb{N}$, such that $\underline{\Psi}^\ell(\xi) \uparrow \underline{\Psi}(\xi)$ and $C(1 + |\xi|^p) \geq \overline{\Psi}^\ell(\xi) \downarrow \overline{\Psi}(\xi)$ for all $\xi \in [0, \infty)$. Next, we set $\Phi_\ell(x, y) := \underline{\Psi}^\ell(\|Sx\|_Y) - \overline{\Psi}^\ell(\|Ty\|_Z)$, $\ell \in \mathbb{N}$. Then, the monotone convergence theorem implies that for all $\xi, \eta \in \mathcal{L}^p(X)$ the conditions $\sup_{\ell\in\mathbb{N}} \mathbb{E}\Phi_\ell(\xi, \eta) \leq 0$ and $\mathbb{E}\left[\underline{\Psi}(\|S\xi\|_Y) - \overline{\Psi}(\|T\eta\|_Z)\right] \leq 0$ are equivalent. □

Let us list some common choices of $\mathcal{P}$ in the setting of decoupling inequalities. To do so, we exploit the following lemma:

[2] Recall that this means that there is an $N \in \mathbb{N}$ with $d_n \equiv 0$ for $n > N$.

Lemma 4.4 *Let $C, p \in (0, \infty)$, let X be a separable Banach space, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\Phi \in C(X; \mathbb{R})$ be such that*

$$|\Phi(x)| \leq C(1 + \|x\|_X^p) \tag{14}$$

for all $x \in X$. Assume $\xi, \xi_n \in \mathcal{L}^p(\mathbb{P}; X)$, $n \in \mathbb{N}$, such that $\xi_n \overset{w^*}{\to} \xi$ as $n \to \infty$ and that $(\xi_n)_{n\in\mathbb{N}}$ is uniformly $\mathcal{L}^p$-integrable. Then,

$$\lim_{n\to\infty} \mathbb{E}\Phi(\xi_n) = \mathbb{E}\Phi(\xi). \tag{15}$$

Proof It follows from the uniform $\mathcal{L}^p$-integrability of $(\xi_n)_{n\in\mathbb{N}}$ and estimate (14) that $(\Phi(\xi_n))_{n\in\mathbb{N}}$ is uniformly $\mathcal{L}^1$-integrable. Moreover, note that $\Phi(\xi_n) \overset{w^*}{\to} \Phi(\xi)$ as $n \to \infty$, so that we may apply Lemma 2.1 for $p = 1$. □

Note that if $\xi_n \to \xi$ in $\mathcal{L}^p(\mathbb{P}; X)$, $\xi_n, \xi \in \mathcal{L}^p(\mathbb{P}; X)$, then the assumptions on $(\xi_n)_{n\in\mathbb{N}}$ and ξ in Lemma 4.4 are satisfied (see [19, Lemma 4.7]).

Example 4.5 (Adapted processes) If $p \in (0, \infty)$ and $\mathcal{P} = \mathcal{P}_p(X)$, then $\mathcal{P}_{p\text{-ext}} = \mathcal{P}$ by Lemma 4.4 and the space $\mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P})$ consists of all $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted processes $(d_n)_{n\in\mathbb{N}}$ in $\mathcal{L}^p(\mathbb{P}; X)$.

Example 4.6 ($\mathcal{L}^p$-martingales) If $p \in [1, \infty)$ and $\mathcal{P}$ consists of all mean zero measures in $\mathcal{P}_p(X)$, then $\mathcal{P}_{p\text{-ext}} = \mathcal{P}$ by Lemma 4.4 (one can test with $\Phi(x) := \langle x, a\rangle$, where $a \in X'$ and X' is the norm-dual) and $\mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P})$ consists of all $\mathcal{L}^p$-integrable $\mathbb{F}$-martingale difference sequences.

Example 4.7 (Conditionally symmetric adapted processes) Suppose $p \in (0, \infty)$ and $\mathcal{P}$ consists of all symmetric measures in $\mathcal{P}_p(X)$. As a measure $\mu \in \mathcal{P}(X)$ is symmetric if and only if for all $f \in C_b(X; \mathbb{R})$ it holds that $\int_X f(x)\,\mathrm{d}\mu(x) = \int_X f(-x)\,\mathrm{d}\mu(x)$, it follows that $\mathcal{P}_{p\text{-ext}} = \mathcal{P}$. Moreover, the set $\mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P})$ consists of all sequences of X-valued $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted sequences of random variables $(d_n)_{n\in\mathbb{N}}$ such that $d_n \in \mathcal{L}^p(\mathbb{P}; X)$ and d_n is $\mathcal{F}_{n-1}$-conditionally symmetric for all $n \in \mathbb{N}$, i.e., for all $n \in \mathbb{N}$ and all $B \in \mathcal{B}(X)$ it holds that $\mathbb{P}(d_n \in B \mid \mathcal{F}_{n-1}) = \mathbb{P}(d_n \in -B \mid \mathcal{F}_{n-1})$ a.s.

Example 4.8 (One- dimensional laws) If $p \in (0, \infty)$, $\emptyset \neq \mathcal{P}_0 \subseteq \mathcal{P}_p(\mathbb{R})$, and

$$\mathcal{P} = \mathcal{P}(\mathcal{P}_0, X) := \left\{\mu \in \mathcal{P}_p(X) \colon \exists \mu_0 \in \mathcal{P}_0,\ x \in X \colon \mu(\cdot) = \mu_0\big(\{r \in \mathbb{R} \colon rx \in \cdot\}\big)\right\},$$

then an X-valued random variable φ satisfies $\mathcal{L}(\varphi) \in \mathcal{P}$ if and only if there exists an $x \in X$ and a $\mathbb{R}$-valued random variable φ_0 such that $\varphi = x\varphi_0$ and $\mathcal{L}(\varphi_0) \in \mathcal{P}_0$. Moreover, $\mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P})$ contains all sequences of the form $(\varphi_n v_{n-1})_{n\in\mathbb{N}}$ where $(\varphi_n)_{n\in\mathbb{N}}$ is an $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted sequence of $\mathbb{R}$-valued random variables such that φ_n is independent of $\mathcal{F}_{n-1}$ and $\mathcal{L}(\varphi_n) \in \mathcal{P}_0$, and $v_{n-1} \in \mathcal{L}^p(\mathcal{F}_{n-1}; X)$ for all $n \in \mathbb{N}$. Finally, it holds that $\mathcal{P}((\mathcal{P}_0)_{p\text{-ext}}, X) \subseteq \mathcal{P}_{p\text{-ext}}$.

5 Decoupling for Dyadic Martingales and Stochastic Integration

In this section, we consider the case of decoupling of dyadic martingales and combine our main result, i.e., Theorem 4.3, with a standard extrapolation argument to obtain a decoupling result that is useful for the theory of stochastic integration of vector-valued stochastic processes, see Theorem 5.2.

5.1 Stochastic Integrals and γ-Radonifying Operators

Let X be a separable Banach space, let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,\infty)})$ be a stochastic basis, and let $W = (W_t)_{t\geq 0}$ be an $(\mathcal{F}_t)_{t\in[0,\infty)}$-Brownian motion, i.e., a centered $\mathbb{R}$-valued Gaussian process such that for all $0 \leq s \leq t < \infty$ it holds that W_t is $\mathcal{F}_t$-measurable, $W_t - W_s$ is independent of $\mathcal{F}_s$, and $\mathbb{E}W_sW_t = s$. We say that $H\colon [0,\infty) \times \Omega \to X$ is a *simple predictable stochastic process* if there exist $0 = t_0 < \cdots < t_N < \infty$ and random variables $v_n \in \mathcal{L}^\infty(\mathcal{F}_{t_n}; X)$, $n \in \{0, \ldots, N-1\}$, such that for all $t \in [0,\infty)$ it holds that

$$H(t,\omega) = \sum_{n=1}^{N} 1_{(t_{n-1},t_n]}(t) v_{n-1}(\omega).$$

For $H\colon [0,\infty) \times \Omega \to X$ an X-valued simple predictable process, we define the stochastic integral $\int_0^\infty H(s)\mathrm{d}W(s)$ in the usual way and we define

$$u_H : L^2((0,\infty)) \times \Omega \to X \quad \text{by} \quad u_H(f)(\omega) := \int_0^\infty f(t)H(t,\omega)\mathrm{d}t.$$

Note that for all $\omega \in \Omega$ we obtain a finite rank operator $u_H(\omega) : L^2((0,\infty)) \to X$. Given a finite rank operator $T : L^2((0,\infty)) \to X$, one can define the γ-radonifying norm $\|\cdot\|_{\gamma(L^2((0,\infty));X)}$ by

$$\|T\|_{\gamma(L^2((0,\infty));X)} := \left\| \sum_{n=1}^{\infty} \gamma_n T e_n \right\|_{\mathcal{L}^2(\mathbb{P}',X)},$$

where $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of $L^2((0,\infty))$ and $(\gamma_n)_{n\in\mathbb{N}}$ is a sequence of independent standard Gaussian random variables on some probability space $(\Omega', \mathcal{F}', \mathbb{P}')$. The γ-radonifying norm is independent of the chosen orthonormal basis. For more information about the γ-radonifying norm, see, for example, [32, Chapter 3] or the survey article [35]. For the relevance of γ-radonifying norms to the definition of vector-valued stochastic integrals, see the definition of and results on $W_p(X)$ in Definition 5.1 and Theorem 5.2, or see [36] for more details.

5.2 Decoupling Constants

In order to state our result (Theorem 5.2), we first recall that a random variable $f \in \mathcal{L}^0((\Omega, \mathcal{F}, \mathbb{P}); X)$ is conditionally symmetric given a sub-σ-algebra $\mathcal{G}$ if $\mathbb{P}(\{f \in B\} \cap G) = \mathbb{P}(\{f \in -B\} \cap G)$ for all $B \in \mathcal{B}(X)$ and $G \in \mathcal{G}$. In addition to the constant $D_p(X)$ from Definition 1.5, we introduce two more constants:

Definition 5.1 Assume a separable Banach space X and $p \in (0, \infty)$.

$\underline{W_p(X)}$: Let $W_p(X) \in [0, \infty]$ be the infimum over all $c \in [0, \infty]$ such that for every stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,\infty)})$, every $(\mathcal{F}_t)_{t\in[0,\infty)}$-Brownian motion W, and every $(\mathcal{F}_t)_{t\in[0,\infty)}$-simple predictable process $H\colon [0, \infty) \times \Omega \to X$ one has that

$$\left\| \int_0^\infty H(s)\mathrm{d}W(s) \right\|_{\mathcal{L}^p(\mathbb{P};X)} \leq c \left\| \|u_H\|_{\gamma(L^2((0,\infty));X)} \right\|_{\mathcal{L}^p(\mathbb{P})}.$$

$\underline{\mathrm{UMD}_p^{-,s}(X)}$: Let $\mathrm{UMD}_p^{-,s}(X) \in [0, \infty]$ be the infimum over all $c \in [0, \infty]$ such that for every stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n\in\mathbb{N}})$ and every finitely supported sequence of X-valued random variables $(d_n)_{n=1}^\infty$ such that $d_n \in \mathcal{L}^p(\mathcal{F}_n; X)$ and d_n is $\mathcal{F}_{n-1}$-conditionally symmetric for all $n \in \mathbb{N}$ it holds that

$$\left\| \sum_{n=1}^\infty d_n \right\|_{\mathcal{L}^p(\mathbb{P};X)} \leq c \left\| \sum_{n=1}^\infty r_n d_n \right\|_{\mathcal{L}^p(\mathbb{P}\otimes\mathbb{P}_{\mathbb{D}};X)}. \tag{16}$$

Theorem 5.2 *Let X be a separable Banach space and $p \in (0, \infty)$.*

(i) *If $D_p(X) < \infty$, then $D_q(X) < \infty$ for all $q \in (0, \infty)$.*

(ii) *If $K_{p,2}$ is the constant in the $\mathcal{L}^p$-to-$\mathcal{L}^2$ Kahane–Khintchine inequality, then*

$$W_p(X) \leq K_{p,2} D_p(X).$$

Conversely, if $W_p(X) < \infty$, then $D_p(X) < \infty$.

(iii) $D_p(X) = \mathrm{UMD}_p^{-,s}(X)$.

For the proof, we use two lemmas. For the formulation of the first one, we introduce, for $\nu \in \mathcal{P}(\mathbb{R})$ and a separable Banach space X, the notation (see also Example 4.8):

$$\mathcal{P}(\nu, X) := \big\{\mu \in \mathcal{P}(X)\colon \exists x \in X\colon \mu(\cdot) = \nu\big(\{r \in \mathbb{R} : rx \in \cdot\}\big)\big\}.$$

Lemma 5.3 *Let X be a separable Banach space, $p \in [2, \infty)$, let $\mu \in \mathcal{P}_p(\mathbb{R})$ satisfy $\int_{\mathbb{R}} r\mathrm{d}\mu(r) = 0$, $\sigma^2 := \int_{\mathbb{R}} |r|^2 \mathrm{d}\mu(r) \in (0, \infty)$, and let $\gamma \in \mathcal{P}(\mathbb{R})$ be the standard Gaussian law. Then, $\mathcal{P}(\gamma, X) \subseteq (\mathcal{P}(\mu, X))_{p\text{-ext}}$.*

Proof Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of independent, μ-distributed random variables, and let $\mu_n := \mathcal{L}((\sigma\sqrt{n})^{-1} \sum_{k=1}^n \xi_k)$. Observe that $\mathcal{L}((\sigma\sqrt{n})^{-1}\xi_1) \in \mathcal{P}(\mu, \mathbb{R})$. Moreover, it follows from, e.g., [3, Theorem 5] that $\mu_n \overset{w^*}{\to} \gamma$ and that $\int_{\mathbb{R}} |r|^p \mathrm{d}\mu_n(r) \to \int_{\mathbb{R}} |r|^p \mathrm{d}\gamma(r)$. It thus follows from Lemma 2.1 that $\gamma \in (\mathcal{P}(\mu, \mathbb{R}))_{p\text{-ext}}$ and hence $\mathcal{P}(\gamma, X) \subseteq (\mathcal{P}(\mu, X))_{p\text{-ext}}$. □

Lemma 5.4 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let X be a separable Banach space, let $p \in (0, \infty)$, let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra, and let $f \in \mathcal{L}^p(\mathcal{F}; X)$ be $\mathcal{G}$-conditionally symmetric. Then, there exists a sequence of $\mathcal{G}$-conditionally symmetric $\mathcal{F}$-simple functions $(f_n)_{n\in\mathbb{N}}$ such that* $\lim_{n\to\infty} \|f - f_n\|_{\mathcal{L}^p(\mathbb{P};X)} = 0$.

Proof Let $(g_n)_{n\in\mathbb{N}}$ be a sequence of $\sigma(f)$-simple functions such that $\lim_{n\to\infty} \|f - g_n\|_{\mathcal{L}^p(\mathbb{P};X)} = 0$. For $n \in \mathbb{N}$, let $m_n \in \mathbb{N}$ and $B_{n,k} \in \sigma(f)$, $x_{n,k} \in X$, $k \in \{1, \dots, m_n\}$, be such that $g_n = \sum_{k=1}^{m_n} x_{n,k} 1_{\{f\in B_{n,k}\}}$. Define, for $n \in \mathbb{N}$, $f_n = \frac{1}{2}\sum_{k=1}^{m_n} x_{n,k}(1_{\{f\in B_{n,k}\}} - 1_{\{-f\in B_{n,k}\}})$ and observe that f_n is $\mathcal{G}$-conditionally symmetric because f is $\mathcal{G}$-conditionally symmetric. Moreover, the conditional symmetry of f implies that $\mathcal{L}(f) = \mathcal{L}(-f)$, whence

$$\begin{aligned}
&\|f - f_n\|_{\mathcal{L}^p(\mathbb{P};X)} \\
&= \left\| \tfrac{1}{2}\left(f - \sum_{k=1}^{m_n} x_{n,k} 1_{\{f\in B_{n,k}\}}\right) - \tfrac{1}{2}\left(-f - \sum_{k=1}^{m_n} x_{n,k} 1_{\{-f\in B_{n,k}\}}\right)\right\|_{\mathcal{L}^p(\mathbb{P};X)} \\
&\le 2^{\left(\frac{1}{p}-1\right)^+} \left\| f - \sum_{k=1}^{m_n} x_{n,k} 1_{\{f\in B_{n,k}\}} \right\|_{\mathcal{L}^p(\mathbb{P};X)} = 2^{\left(\frac{1}{p}-1\right)^+} \|f - g_n\|_{\mathcal{L}^p(\mathbb{P};X)}.
\end{aligned}$$

□

Proof of Theorem 5.2 Part (i) follows from [14] and can be found in [8, Proposition B.1] for the convenience of the reader.

Part (ii): first, we check $W_p(X) \le K_{p,2} D_p(X)$. For $0 = t_0 < \cdots < t_N < \infty$ Lemma 5.3 applied to $\mu := \frac{1}{2}(\delta_{-1} + \delta_1)$ and Theorem 1.4 give (see also Corollary 1.6)

$$\left\| \sum_{n=1}^{N} (W_{t_n} - W_{t_{n-1}}) v_{n-1} \right\|_{\mathcal{L}^p(\mathbb{P};X)} \le D_p(X) \left\| \sum_{n=1}^{N} (W'_{t_n} - W'_{t_{n-1}}) v_{n-1} \right\|_{\mathcal{L}^p(\mathbb{P}\otimes\mathbb{P}';X)}$$

for all $\mathcal{L}^p$-integrable and $\mathcal{F}_{t_{n-1}}$-measurable random variables $v_{n-1} : \Omega \to X$ where $(W'_t)_{t\ge 0}$ is a Brownian motion defined on an auxiliary basis $(\Omega', \mathcal{F}', \mathbb{P}', (\mathcal{F}'_t)_{t\in[0,\infty)})$. Exploiting the Kahane–Khintchine inequality gives that

$$\begin{aligned}
&\left\| \sum_{n=1}^{N} (W'_{t_n} - W'_{t_{n-1}}) v_{n-1} \right\|_{\mathcal{L}^p(\mathbb{P}\otimes\mathbb{P}';X)} \\
&\le K_{p,2} \left(\int_\Omega \left(\int_{\Omega'} \left\| \sum_{n=1}^{N} (W'_{t_n}(\omega') - W'_{t_{n-1}}(\omega')) v_{n-1}(\omega) \right\|_X^2 \mathrm{d}\mathbb{P}'(\omega') \right)^{\frac{p}{2}} \mathrm{d}\mathbb{P}(\omega) \right)^{\frac{1}{p}}.
\end{aligned}$$

For $H := \sum_{n=1}^{N} 1_{(t_{n-1},t_n]} v_{n-1}$, the result follows by the known relation

$$\left(\int_{\Omega'} \left\| \sum_{n=1}^{N} (W'_{t_n}(\omega') - W'_{t_{n-1}}(\omega')) v_{n-1}(\omega) \right\|_X^2 \mathrm{d}\mathbb{P}'(\omega') \right)^{\frac{1}{2}} = \|u_H(\omega)\|_{\gamma(L^2((0,\infty));X)}.$$

Conversely, let us assume that $W_p(X) < \infty$. Now we use [37, Lemma 2.5] to deduce that X has finite cotype. (In [37, Lemma 2.5] it is assumed that $p \in [1, \infty)$, however in the 5th line of the proof of this lemma it is shown that $\| \int_0^\infty \phi(t) \mathrm{d}W(t) \| = N$ a.s., which implies the desired conclusion for $p \in (0, \infty)$, see also [8, Lemma 6.1].) Thus, the proof of [37, Theorem 2.2] guarantees that $D_p(X) < \infty$. Here, we exploit that [26, Proposition 9.14] works (in their notation) with $r \in (0, 1)$ as well: one starts on the left-hand side with $\mathcal{L}^r$, estimates this by $\mathcal{L}^1$, applies [26, Proposition 9.14], and uses [26, Proposition 4.7] (Khintchine's inequality for a vector-valued Rademacher series) to change $\mathcal{L}^1$ back to $\mathcal{L}^r$ on the right-hand side.

Part (iii) is divided into several steps:

PROOF of $D_p(X) \leq \mathrm{UMD}_p^{-,s}(X)$: this inequality follows from the following two observations: firstly, dyadic martingales are conditionally symmetric, and secondly if $(d_n)_{n=1}^\infty = (r_n v_{n-1})_{n=1}^\infty$ is a dyadic martingale and $(r'_n)_{n=1}^\infty$ is Rademacher sequence independent of $(r_n)_{n=1}^\infty$, then $(r'_n r_n v_{n-1})_{n=1}^\infty$ and $(r'_n v_{n-1})_{n=1}^\infty$ are equal in distribution.

PROOF of $\mathrm{UMD}_p^{-,s}(X) \leq D_p(X)$: using Lemma 5.4, we approximate each d_n in $\mathcal{L}^p(\mathcal{F}_n; X)$ so that we may assume that the d_n take finitely many values only. Let

$$\varepsilon_0 := \inf\{\|d_n(\omega)\|_X : n = 1, \ldots, N,\ \omega \in \Omega,\ d_n(\omega) \neq 0\} > 0$$

where $\inf \emptyset := 1$. Take an $x \in X$ with $0 < \|x\|_X < \varepsilon_0$ and let $r = (r_n)_{n=1}^N$ be a Rademacher sequence on a probability space $(\Omega_{\mathbb{D}}, \mathcal{F}_{\mathbb{D}}, \mathbb{P}_{\mathbb{D}})$. If we define $\tilde{d}_n : \Omega \times \Omega_{\mathbb{D}} \to X$ by $\tilde{d}_n(\omega, \omega_{\mathbb{D}}) := d_n(\omega) + r_n(\omega_{\mathbb{D}})x$, then $\tilde{d}_n(\omega, \omega_{\mathbb{D}}) \neq 0$ for all $(\omega, \omega_{\mathbb{D}}) \in \Omega \times \Omega_{\mathbb{D}}$ and $\tilde{d}_n$ is conditionally symmetric given the σ-algebra $\mathcal{F}_{n-1} \otimes \mathcal{F}^r_{\mathbb{D},n-1}$, where $(\mathcal{F}^r_{\mathbb{D},n})_{n=0}^N$ is the natural filtration of $(r_n)_{n=1}^N$. Because we may let $\|x\| \downarrow 0$, it suffices to verify the statement for $(\tilde{d}_n)_{n=1}^N$ or, in other words, we may assume without loss of generality that for all $n \in \mathbb{N}$ the range of d_n is a finite set that does not contain 0.

Note that by removing all (i.e., at most finitely many) atoms of measure zero in the σ-algebra $\mathcal{F}^d_N$ and 'updating' the definition of $(d_n)_{n=1}^N$ accordingly, we may assume that the filtration $(\mathcal{F}^d_n)_{n=1}^N$ has the property that $\mathcal{F}^d_n$ is generated by finitely many atoms of positive measure.

Bearing in mind that for all $n \in \{1, \ldots, N\}$ the random variable d_n takes only finitely many values, each nonzero, and each with positive probability, one may check that for every atom $A \in \mathcal{F}^d_{n-1}$, $n \in \{1, \ldots, N\}$, there exist disjoint sets $A^+, A^- \in \mathcal{F}^d_n$ such that $A = A^+ \cup A^-$, $\mathbb{P}(A^+) = \mathbb{P}(A^-)$, and $\mathcal{L}(d_n \mid A^+) = \mathcal{L}(-d_n \mid A^-)$. Now we introduce a Rademacher sequence $(\rho_n)_{n=1}^N$, $\rho_n : \Omega \to \{-1, 1\}$, defined as follows: for each atom A of $\mathcal{F}^d_{n-1}$ we set $\rho_n|_{A^+} \equiv 1$, and $\rho_n|_{A^-} \equiv -1$, where A^+ and A^- form

a partition of A as described above. Moreover, we let $v_n := \rho_n d_n$ so that $d_n = \rho_n v_n$. By construction, ρ_n is independent from $\mathcal{F}^d_{n-1} \vee \sigma(v_n)$. It follows from the definition of $D_p(X)$ and Theorem 1.4 (see also Example 2.6) that

$$\begin{aligned}\left\|\sum_{n=1}^N d_n\right\|_{\mathcal{L}^p(\mathbb{P};X)} &= \left\|\sum_{n=1}^N \rho_n v_n\right\|_{\mathcal{L}^p(\mathbb{P};X)} \leq D_p(X)\left\|\sum_{n=1}^N r'_n v_n\right\|_{\mathcal{L}^p(\mathbb{P}\otimes\mathbb{P}_{\mathbb{D}};X)}\\ &= D_p(X)\left\|\sum_{n=1}^N r'_n r_n v_n\right\|_{\mathcal{L}^p(\mathbb{P}\otimes\mathbb{P}_{\mathbb{D}};X)} = D_p(X)\left\|\sum_{n=1}^N r'_n d_n\right\|_{\mathcal{L}^p(\mathbb{P}\otimes\mathbb{P}_{\mathbb{D}};X)}.\end{aligned}$$

□

We conclude with some remarks regarding Theorem 5.2.

Remark 5.5 (1) Let $(h_n)_{n\in\mathbb{N}}$ be the Haar system for $\mathcal{L}^2((0,1])$ with $\operatorname{ess\,sup}(|h_n|) = 1$, and let $H_p(X) \in [0,\infty]$ be the infimum over all $c \in [0,\infty]$ such that for all finitely supported sequences $(x_n)_{n\in\mathbb{N}}$ in X one has that

$$\left\|\sum_{n=1}^\infty h_n x_n\right\|_{\mathcal{L}^p((0,1]);X)} \leq c\left\|\sum_{n=1}^\infty r_n h_n x_n\right\|_{\mathcal{L}^p((0,1]\times\mathbb{D};X)}.$$

Let $|H|_p(X) \in [0,\infty]$ be defined as above but with h_n replaced by $|h_n|$. Then, it is straightforward to see that $D_p(X) = H_p(X) = |H|_p(X)$.

(2) Garling [13] introduced the constant $\mathrm{UMD}^-_p(X)$, which is defined like the constant $\mathrm{UMD}^{-,s}_p(X)$ in Definition 5.1 but without the condition *conditionally symmetric*. In general, the constants $\mathrm{UMD}^-_p(X)$ and $\mathrm{UMD}^{-,s}_p(X)$ behave differently: it follows from Hitczenko [16, Theorem 1.1] that $\sup_{p\in[2,\infty)} D_p(\mathbb{R}) < \infty$ and thus, by Theorem 5.2, $\sup_{p\in[2,\infty)} \mathrm{UMD}^{-,s}_p(\mathbb{R}) < \infty$. On the other hand, as outlined in [9, p. 348], one has $\mathrm{UMD}^-_p(\mathbb{R}) \succeq \sqrt{p}$ as $p \to \infty$ by combining the result of Burkholder [5, Theorem 3.1] about the optimal behavior of the constant in the square function inequality and the behavior of the constant in the Khintchine inequality for Rademacher variables.

Remark 5.6 Part (ii) of Theorem 5.2 is an extension of Garling's [12, Theorem 2]: whereas Garling requires the integrands to be adapted with respect to the filtration *generated* by the Brownian motion, we can allow for any filtration. In the development of stochastic integration theory in Banach spaces (as presented in, e.g., [9,34]), the issue regarding the undesirable assumption on the filtration in [12] was known to the authors. In those articles, the problem was circumvented in two ways:

(a) In [34, Lemma 3.4], a decoupling argument due to Montgomery-Smith [29] is used to prove $W_p(X) \leq \beta_p(X)$ for $p \in (1,\infty)$, where $\beta_p(X)$ is the $\mathcal{L}^p$-UMD constant of X. This approach does not cover $p \in (0,1]$ and the UMD property seems to be too strong as $W_p(L^1) < \infty$ for $p \in (0,\infty)$ (see also [9]).

(b) In [9, Theorem 5.4], it is observed that $W_p(X) < \infty$ if $D_p^{\text{gen}}(X) < \infty$, where $D_p^{\text{gen}}(X)$ is the infimum over all $c \in [0, \infty]$ such that

$$\left\| \sum_{n=1}^{\infty} d_n \right\|_{\mathcal{L}^p(\mathbb{P};X)} \leq c \left\| \sum_{n=1}^{\infty} e_n \right\|_{\mathcal{L}^p(\mathbb{P};X)}$$

whenever $(e_n)_{n\in\mathbb{N}}$ is an $\mathbb{F}$-decoupled tangent sequence of a finitely supported $\mathcal{L}^p$-integrable X-valued $\mathbb{F}$-adapted sequence of random variables $(d_n)_{n\in\mathbb{N}}$.

The approach in [9] leads us to wonder: is that true that $D_p(X) < \infty$ implies $D_p^{\text{gen}}(X) < \infty$? (See Open Problem 1.1.) Although we could not fully answer this question, Theorem 5.2 resolves the issue regarding the filtration in [12] and thereby provides a direct approach for vector-valued stochastic integration.

Acknowledgements The first author is supported by the research program *VENI Vernieuwingsimpuls* with Project Number 639.031.549, which is financed by the Netherlands Organization for Scientific Research (NWO). The second author is supported by the project *Stochastic Analysis and Nonlinear Partial Differential Equations, Interactions and Applications* of the Academy of Finland with Project Number 298641. The authors wish to thank Mark Veraar, Peter Spreij, and an anonymous referee. The first author would also like to thank Lotte Meijer.

Funding Open access funding provided by University of Jyväskylä (JYU).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A: Proof of Theorem 4.3

Theorem 4.3 is proved in Sect. A.2, and Sect. A.1 provides some necessary tools.

A.1 Technical lemmas

We begin with Lemmas A.1 and A.2 that are obtained by an adaptation of [33, Lemma 12.8], in which the dyadic setting is considered and which simplifies the procedure originally sketched in [27].

Recall that a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *divisible* if for every $A \in \mathcal{F}$ and every $\theta \in (0, 1)$ there exists an $A_\theta \in \mathcal{F}$ such that $A_\theta \subset A$ and $\mathbb{P}(A_\theta) = \theta\mathbb{P}(A)$.

Lemma A.1 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a divisible probability space, X be a separable Banach space, $F \in \mathcal{F}$, and let $\mu \in \mathcal{P}(X)$ be of the form $\mu = \sum_{k=1}^{n} \alpha_k \delta_{x_k}$ for some $n \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_n \in (0, 1)$, and some distinct $x_1, \ldots, x_n \in X$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra generated by a finite partition $(A_i)_{i=1}^{k}$ of Ω with $\mathbb{P}(A_i) > 0$. Then, there exists an*

$\mathcal{F}$-measurable, μ-distributed random variable φ that is independent of $\mathcal{G}$, for which there exist $H_1, H_2 \in \sigma(\mathcal{G}, \varphi)$ satisfying $H_1 \subseteq F \subseteq H_2$ and

$$\mathbb{P}(H_2 \backslash H_1) \leq \left[\max_{j \in \{1,\ldots,n\}} \alpha_j\right] \min\left\{\mathbb{P}(G_2 \backslash G_1) \colon G_1, G_2 \in \mathcal{G}, G_1 \subseteq F \subseteq G_2\right\}.$$

Proof As $(\Omega, \mathcal{F}, \mathbb{P})$ is divisible, we can construct a partition $(A_{i,j})_{\substack{i \in \{1,\ldots,k\} \\ j \in \{1,\ldots,n\}}}$ of Ω with $A_{i,j} \in \mathcal{F}$ for all i, j, such that $A_i = \bigcup_{j=1}^n A_{i,j}$ for all $i \in \{1, \ldots, k\}$, and such that $\mathbb{P}(A_{i,j}) = \alpha_j \mathbb{P}(A_i)$ for all $j \in \{1, \ldots, n\}$ and all $i \in \{1, \ldots, k\}$.

The partition $(A_{i,j})_{\substack{i \in \{1,\ldots,k\} \\ j \in \{1,\ldots,n\}}}$ is assumed to satisfy some conditions with respect to the set F which we shall explain below. Before doing so, we observe that given such a partition, the random variable φ defined by $\varphi := \sum_{i=1}^k \sum_{j=1}^n x_j 1_{A_{i,j}}$ has the law μ and is independent of $\mathcal{G} = \sigma((A_i)_{i=1}^k)$, and

$$\sigma(\mathcal{G}, \varphi) = \sigma\left(\left\{A_{i,j} \colon i \in \{1, \ldots, k\}, j \in \{1, \ldots, n\}\right\}\right).$$

Let $I_0 \subseteq \{1, \ldots, k\}$ be such that $i \in I_0$ if and only if $A_i \cap F = \emptyset$, and $I_1 \subseteq \{1, \ldots, k\}$ be such that $i \in I_1$ if and only if $A_i \subseteq F$. Set $I_{\text{mix}} = \{1, \ldots, k\} \backslash (I_0 \cup I_1)$ (one or two of the sets I_0, I_1, I_{mix} may be empty). Observe that

$$\sum_{i \in I_{\text{mix}}} \mathbb{P}(A_i) = \min\left\{\mathbb{P}(G_2 \backslash G_1) \colon G_1, G_2 \in \mathcal{G}, G_1 \subseteq F \subseteq G_2\right\}. \tag{17}$$

For $i \in I_0 \cup I_1$, we simply partition the set A_i into sets $(A_{i,j})_{j=1}^n$ that satisfy $A_{i,j} \in \mathcal{F}$ and

$$\mathbb{P}(A_{i,j}) = \alpha_j \mathbb{P}(A_i)$$

for all $j \in \{1, \ldots, n\}$. For $i \in I_{\text{mix}}$ we choose the partition $(A_{i,j})_{j=1}^n$ not only such that it satisfies $A_{i,j} \in \mathcal{F}$ and

$$\mathbb{P}(A_{i,j}) = \alpha_j \mathbb{P}(A_i)$$

for all $j \in \{1, \ldots, n\}$, but also such that there is *at most one* $j \in \{1, \ldots, n\}$ such that $\emptyset \neq F \cap A_{i,j} \subsetneq A_{i,j}$. It follows from this construction and from (17) that

$$\begin{aligned}
&\min\left\{\mathbb{P}(H_2 \backslash H_1) \colon H_1, H_2 \in \sigma(\mathcal{G}, \varphi), H_1 \subseteq F \subseteq H_2\right\} \\
&\quad \leq \sum_{i \in I_{\text{mix}}} \left[\max_{j \in \{1,\ldots,n\}} \alpha_j\right] \mathbb{P}(A_i) \\
&\quad = \left[\max_{j \in \{1,\ldots,n\}} \alpha_j\right] \min\left\{\mathbb{P}(G_2 \backslash G_1) \colon G_1, G_2 \in \mathcal{G}, G_1 \subseteq F \subseteq G_2\right\}.
\end{aligned}$$

□

Lemma A.2 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a divisible probability space, X be a separable Banach space and let $\mu \in \mathcal{P}(X)$ be of the form $\mu = \sum_{k=1}^{n} \alpha_k \delta_{x_k}$ for some $n \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_n \in (0, 1)$, and some distinct $x_1, \ldots, x_n \in X$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra generated by a finite partition of atoms with positive measure. Then, for every $A \in \mathcal{F}$ and every $\varepsilon > 0$ there exists an $m \in \mathbb{N}$ and $\mathcal{F}$-measurable independent μ-distributed random variables $(\varphi_1, \ldots, \varphi_m)$ that are independent of $\mathcal{G}$ such that there exists an $A_\varepsilon \in \sigma(\mathcal{G}, \varphi_1, \ldots, \varphi_m)$ satisfying $\mathbb{E}|1_A - 1_{A_\varepsilon}| < \varepsilon$.*

Proof Let $A \in \mathcal{F}$ and $\varepsilon > 0$ be given. Define $\delta := \max_{j \in \{1,\ldots,n\}} \alpha_j \in (0, 1)$ and let $m \in \mathbb{N}$ be such that $\delta^m < \varepsilon$.

Step 1 Apply Lemma A.1 with $\mathcal{G}$ and $\mathcal{F}$ as given to find an $\mathcal{F}$-measurable, μ-distributed random variable φ_1 that is independent of $\mathcal{G}$, and sets $H_{1,1}, H_{1,2} \in \sigma(\mathcal{G}, \varphi_1)$ such that $H_{1,1} \subseteq A \subseteq H_{1,2}$ and $\mathbb{P}(H_{1,2} \backslash H_{1,1}) \leq \delta$. Define $\mathcal{G}_1 := \sigma(\mathcal{G}, \varphi_1)$ that is, by construction, a σ-algebra generated by a finite partition of sets of positive measure.

Step i, $i = 2, \ldots, m$ Apply Lemma A.1 with $\mathcal{G} := \mathcal{G}_{i-1}$, and with $\mathcal{F}$ as given, to find an $\mathcal{F}$-measurable, μ-distributed random variable φ_i that is independent of $\mathcal{G}_{i-1}$, and sets $H_{i,1}, H_{i,2} \in \sigma(\mathcal{G}_{i-1}, \varphi_i)$ such that $H_{i,1} \subseteq A \subseteq H_{i,2}$ and

$$\mathbb{P}(H_{i,2} \backslash H_{i,1}) \leq \delta \mathbb{P}(H_{i-1,2} \backslash H_{i-1,1}) \leq \delta^i.$$

Set $\mathcal{G}_i := \sigma(\mathcal{G}_{i-1}, \varphi_i)$.

We have now obtained a sequence of independent, $\mathcal{F}$-measurable, μ-distributed random variables $(\varphi_1, \ldots, \varphi_m)$ that are independent of $\mathcal{G}$, and sets $H_{m,1}, H_{m,2} \in \mathcal{G}_m = \sigma(\mathcal{G}, \varphi_1, \ldots, \varphi_m)$ such that $H_{m,1} \subseteq A \subseteq H_{m,2}$ and

$$\mathbb{P}(H_{m,2} \backslash H_{m,1}) \leq \delta^m < \varepsilon.$$

Setting $A_\varepsilon = H_{m,1}$, we obtain that $\mathbb{E}|1_A - 1_{A_\varepsilon}| < \varepsilon$. □

Lemmas A.3, A.4, and A.5 concern approximation procedures:

Lemma A.3 *Let X be a separable Banach space, $C, p \in (0, \infty)$, and let $\Phi \in C(X \times X, \mathbb{R})$ satisfy*

$$|\Phi(x, y)| \leq C(1 + \|x\|_X^p + \|y\|_X^p) \tag{18}$$

for all $(x, y) \in X \times X$, let $N \in \mathbb{N}$, and let $\mu_1, \ldots, \mu_N \in \mathcal{P}_p(X)$. Then, for all $\varepsilon > 0$ there exists a measurable mapping $P_\varepsilon \colon X \to X$ with finite range such that for every sequence of independent random variables $(\varphi_1, \ldots, \varphi_N, \varphi_1', \ldots, \varphi_N')$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathcal{L}(\varphi_n) = \mathcal{L}(\varphi_n') = \mu_n$, $n \in \{1, \ldots, N\}$, and for all $F_1, \ldots, F_N \in \mathcal{F}$ it holds that

$$\mathbb{E}\left|\Phi\left(\sum_{n=1}^{N} \varphi_n 1_{F_n}, \sum_{n=1}^{N} \varphi_n' 1_{F_n}\right) - \Phi\left(\sum_{n=1}^{N} P_\varepsilon(\varphi_n) 1_{F_n}, \sum_{n=1}^{N} P_\varepsilon(\varphi_n') 1_{F_n}\right)\right| < \varepsilon. \tag{19}$$

Moreover, if for some $n \in \{1, \ldots, N\}$ it holds that μ_n is not a Dirac measure, then P_ε may be chosen such that $\mu_n \circ P_\varepsilon^{-1}$ is not a Dirac measure.

Proof Fix $\varepsilon > 0$, set $M_p = \sum_{n=1}^{N} \int_X \|x\|_X^p \, \mathrm{d}\mu_n(x)$ and let $K \subseteq X$ be a compact set such that

$$\sup_{n\in\{1,\dots,N\}} \int_{K^c} (1 + N^p M_p + N^p \|x\|_X^p) \, \mathrm{d}\mu_n(x) < 2^{-(p-1)^+} (8CN)^{-1} \varepsilon, \tag{20}$$

where C is as in (18). (Note that such a set K exists as X is separable and hence $\mu_1, \dots, \mu_N$ are Radon measures, and moreover $\mu_1, \dots, \mu_N \in P_p(X)$.) It follows that for $K^N = K \times \ldots \times K$ (N times) one has

$$\begin{aligned}
&\int_{(K^N)^c} \left(1 + N^p \sum_{n=1}^{N} \|x_n\|_X^p\right) \mathrm{d}\mu_1(x_1) \dots \mathrm{d}\mu_N(x_N) \\
&\quad \le \sum_{j=1}^{N} \int_{\{x_j \in K^c\}} \left(1 + N^p \sum_{n=1}^{N} \|x_n\|_X^p\right) \mathrm{d}\mu_1(x_1) \dots \mathrm{d}\mu_N(x_N) \\
&\quad \le \sum_{j=1}^{N} \int_{\{x_j \in K^c\}} \left(1 + N^p \|x_j\|_X^p\right) \mathrm{d}\mu_1(x_1) \dots \mathrm{d}\mu_N(x_N) \\
&\qquad + \sum_{j=1}^{N} \sum_{n=1, n\neq j}^{N} \int_{\{x_j \in K^c\}} N^p \|x_n\|_X^p \, \mathrm{d}\mu_1(x_1) \dots \mathrm{d}\mu_N(x_N) \\
&\quad \le \sum_{j=1}^{N} \int_{\{x_j \in K^c\}} \left(1 + N^p M_p + N^p \|x_j\|_X^p\right) \mathrm{d}\mu_1(x_1) \dots \mathrm{d}\mu_N(x_N) \\
&\quad < 2^{-(p-1)^+} (8C)^{-1} \varepsilon.
\end{aligned} \tag{21}$$

As Φ is continuous, it is uniformly continuous on $K \times K$ and hence there exists a $\delta \in (0, \infty)$ such that if $x_1, y_1, x_2, y_2 \in K$ and $\|x_1 - x_2\|_X < \delta$, $\|y_1 - y_2\|_X < \delta$, then it holds that $|\Phi(x_1, y_1) - \Phi(x_2, y_2)| < \frac{\varepsilon}{2}$. Note that without loss of generality we may assume that $\delta \le N^{-1/p}$ and that $K \neq \emptyset$. Now let $M \in \mathbb{N}$ and $\{U_1, \dots, U_M\} \subseteq \mathcal{B}(X)$ be a partition of K such that for all $m \in \{1, \dots, M\}$ it holds that $U_m \neq \emptyset$ and

$$\sup_{m\in\{1,\dots,M\}} \sup_{x,y\in U_m} \|x - y\|_X < N^{-1}\delta.$$

Let $x_1, \dots, x_M \in X$ be such that $x_m \in U_m$, $m \in \{1, \dots, M\}$. Let $x_0 \in \{x \in X \colon \|x\|_X = N^{-1}\delta\} \setminus \{x_1, \dots, x_M\}$ (this will be important for the last part of the proof of the lemma). Define $P_\varepsilon \colon X \to X$ by

$$P_\varepsilon(x) = \begin{cases} x_m; & x \in U_m, \\ x_0; & x \notin K. \end{cases} \tag{22}$$

Observe that by construction for all $x \in X$ it holds that

$$\|P_\varepsilon(x)\|_X \leq \|x\|_X + N^{-1}\delta \leq \|x\|_X + N^{-(1+\frac{1}{p})} \tag{23}$$

and for all $x \in K$ it holds that

$$\|x - P_\varepsilon(x)\|_X < N^{-1}\delta. \tag{24}$$

We verify that P_ε satisfies the desired properties. Indeed, clearly P_ε has finite range. Moreover, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\varphi_1, \ldots, \varphi_N, \varphi_1', \ldots, \varphi_N')$ be random variables on this space such that $\mathcal{L}(\varphi_n) = \mathcal{L}(\varphi_n') = \mu_n$, $n \in \{1, \ldots, N\}$, and let $F_1, \ldots, F_N \in \mathcal{F}$. For simplicity of notation, define $\xi = \sum_{n=1}^N \varphi_n 1_{F_n}$, $\xi' = \sum_{n=1}^N \varphi_n' 1_{F_n}$, $\xi_\varepsilon = \sum_{n=1}^N P_\varepsilon(\varphi_n) 1_{F_n}$, and $\xi_\varepsilon' = \sum_{n=1}^N P_\varepsilon(\varphi_n') 1_{F_n}$. Define

$$K_\varphi = \{\omega \in \Omega \colon (\varphi_1, \ldots, \varphi_N, \varphi_1', \ldots, \varphi_N') \in K^{2N}\}.$$

Observe that by (24) for $\omega \in K_\varphi$ it holds that

$$\|\xi(\omega) - \xi_\varepsilon(\omega)\|_X \leq \sum_{n=1}^N \|\varphi_n - P_\varepsilon(\varphi_n)\|_X < \delta,$$

and similarly $\|\xi'(\omega) - \xi_\varepsilon'(\omega)\|_X < \delta$, whence for all $\omega \in K_\varphi$ it holds that

$$|\Phi(\xi(\omega), \xi'(\omega)) - \Phi(\xi_\varepsilon(\omega), \xi_\varepsilon'(\omega))| < \frac{\varepsilon}{2}. \tag{25}$$

By the estimate above, Assumption (18), and inequalities (23) and (21) it now follows that

$$\begin{aligned}
&\mathbb{E}|\Phi(\xi, \xi') - \Phi(\xi_\varepsilon, \xi_\varepsilon')| \\
&\quad = \int_{K_\varphi} |\Phi(\xi, \xi') - \Phi(\xi_\varepsilon, \xi_\varepsilon')| \, \mathrm{d}\mathbb{P} + \int_{K_\varphi^c} |\Phi(\xi, \xi') - \Phi(\xi_\varepsilon, \xi_\varepsilon')| \, \mathrm{d}\mathbb{P} \\
&\quad < \frac{\varepsilon}{2} + C \int_{K_\varphi^c} (2 + \|\xi\|_X^p + \|\xi'\|_X^p + \|\xi_\varepsilon\|_X^p + \|\xi_\varepsilon'\|_X^p) \, \mathrm{d}\mathbb{P} \\
&\quad \leq \frac{\varepsilon}{2} + 2C \int_{K_\varphi^c} \left(1 + N^p \left(\sum_{n=1}^N \|\varphi_n\|_X^p + \sum_{n=1}^N \|P_\varepsilon(\varphi_n)\|_X^p\right)\right) \mathrm{d}\mathbb{P} \\
&\quad \leq \frac{\varepsilon}{2} + 2^{(p-1)^+} 4C \int_{K_\varphi^c} \left(1 + N^p \sum_{n=1}^N \|\varphi_n\|_X^p\right) \mathrm{d}\mathbb{P} < \varepsilon.
\end{aligned}$$

Recalling the definition of ξ, ξ', ξ_ε and ξ_ε', this completes the proof of estimate (19).

In order to prove the final statement in the lemma, we make some minor adjustments to the proof above. Indeed, suppose that for some $n \in \{1, \ldots, N\}$ it holds that μ_n is

not a Dirac measure. It follows that there exists a compact set $F \in \mathcal{B}(X)$ such that $\mu_n(F) \in (0,1)$. Now proceed as above, but with the additional assumption that the set K satisfying (20) also satisfies $F \subseteq K$, and that the partition $\{U_1, \dots, U_M\}$ is chosen such that $U_m \cap F \in \{U_m, \emptyset\}$ for all $m \in \{1, \dots, M\}$. As $x_0, x_1, \dots, x_M$ are all distinct values by construction, this ensures that there exists a set $G \in \mathcal{B}(X)$ such that $P_\varepsilon^{-1}(G) = F$. □

Lemma A.4 *Let X be a separable Banach space, $\mu \in \mathcal{P}(X)$, and let $\mathcal{B}_{\mu\text{-cont}}(X) = \{B \in \mathcal{B}(X) \colon \mu(\partial B) = 0\}$. Then, for all $B \in \mathcal{B}(X)$ and all $\varepsilon > 0$ there exists an $B_\varepsilon \in \mathcal{B}_{\mu\text{-cont}}(X)$ such that $\mu(B \triangle B_\varepsilon) < \varepsilon$.*

Proof If $\mathcal{A} := \left\{B \in \mathcal{B}(X) \colon \forall \varepsilon > 0\, \exists B_\varepsilon \in \mathcal{B}_{\mu\text{-cont}}(X) \text{ with } \mu(B \triangle B_\varepsilon) < \varepsilon\right\}$, then we may check that $\mathcal{A}$ is a Dynkin-system and that $\mathcal{A}$ contains all closed sets. Whence the result follows by the π-λ-Theorem. □

Lemma A.5 *For $N \in \mathbb{N}$ and a separable Banach space X let $\varphi_1, \dots, \varphi_N : \Omega \to X$ and $\varphi_{1,k}, \dots, \varphi_{N,k} : \Omega_k \to X$, $k \in \mathbb{N}$, be families of independent random variables with $w^* - \lim_{k\to\infty} \varphi_{n,k} = \varphi_n \quad \text{for} \quad n \in \{1, \dots, N\}$. Let $(\varphi_n')_{n=1}^N : \Omega \to X^N$ and $(\varphi_{n,k}')_{n=1}^N : \Omega_k \to X^N$ be independent copies of $(\varphi_n)_{n=1}^N$ and $(\varphi_{n,k})_{n=1}^N$, respectively, $v_0 \in \mathbb{R}$, and $B_n \in \mathcal{B}(X^n)$ such that $\mathbb{P}((\varphi_j)_{j=1}^n \in \partial B_n) = 0$ for $n \in \{1, \dots, N-1\}$. Then, for the w^*-convergence in $X \times X$ it holds that*

$$
\begin{aligned}
&w^* - \lim_{k\to\infty} \left(\varphi_{1,k} v_0 + \sum_{n=2}^N \varphi_{n,k} 1_{\{(\varphi_{j,k})_{j=1}^{n-1} \in B_{n-1}\}}, \varphi_{1,k}' v_0 + \sum_{n=2}^N \varphi_{n,k}' 1_{\{(\varphi_{j,k})_{j=1}^{n-1} \in B_{n-1}\}} \right) \\
&\quad = \left(\varphi_1 v_0 + \sum_{n=2}^N \varphi_n 1_{\{(\varphi_j)_{j=1}^{n-1} \in B_{n-1}\}}, \varphi_1' v_0 + \sum_{n=2}^N \varphi_n' 1_{\{(\varphi_j)_{j=1}^{n-1} \in B_{n-1}\}} \right).
\end{aligned}
$$

Proof This result follows by an application of the Skorokhod theorem (see, e.g., [19, Theorem 4.30]) and the Portmanteau theorem (see, e.g., [19, Theorem 4.25]). □

In addition to the space $\mathcal{A}_p(\Omega, \mathbb{F}; X, \mathcal{P})$ introduced in Definition 1.2, we shall need the following one:

Definition A.6 Let X be a separable Banach space, $p \in (0, \infty)$, $\emptyset \neq \mathcal{P} \subseteq \mathcal{P}_p(X)$, and let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^N)$, $N \in \mathbb{N}$, be a stochastic basis. We shall denote by $\mathcal{A}_{p\text{-simple}}(\Omega, (\mathcal{F}_n)_{n=0}^N; X, \mathcal{P})$ the set of $(\mathcal{F}_n)_{n=1}^N$-adapted sequences $(d_n)_{n=1}^N$ such that for all $n \in \{1, \dots, N\}$ there exist $K_n \in \mathbb{N}$, a partition $(A_{n-1,k})_{k=1}^{K_n} \subseteq \mathcal{F}_{n-1}$ of Ω consisting of sets of positive measure, and $\mu_{n,1}, \dots, \mu_{n,K_n} \in \mathcal{P}$ such that $\sum_{k=1}^{K_n} 1_{A_{n-1,k}} \mu_{n,k}$ is a regular version of $\mathbb{P}(d_n \in \cdot \mid \mathcal{F}_{n-1})$.

To deal with this class of processes, we use the following two lemmas:

Lemma A.7 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $K \in \mathbb{N}$ and let $A_k \in \mathcal{F}$, $k \in \{1, \dots, K\}$, be such that $(A_k)_{k=1}^K$ is a partition of Ω and $\mathbb{P}(A_k) > 0$ for all $k \in \{1, \dots, K\}$. Let X be a separable Banach space and let $d \colon \Omega \to X$ be a random variable. Let $\kappa \colon \Omega \to \mathcal{P}(X)$ be a regular version of $\mathcal{P}(d \in \cdot \mid \sigma((A_k)_{k=1}^K))$, i.e.,*

$\kappa = \sum_{k=1}^{K} \mu_k 1_{A_k}$ for some $\mu_1, \ldots, \mu_K \in \mathcal{P}(X)$. Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be an auxiliary probability space and let $(d'_k)_{k=1}^{K}$ be a sequence of independent X-valued random variables on $(\Omega', \mathcal{P}', \mathcal{F}')$ satisfying $\mathcal{L}(d'_k) = \mu_k$, and let, for all $k \in \{1, \ldots, K\}$, $d_k : (\Omega, \mathcal{F}, \mathbb{P}) \times (\Omega', \mathcal{F}', \mathcal{P}') \to X$ be a random variable defined by

$$d_k(\omega, \omega') = d(\omega) 1_{A_k}(\omega) + d'_k(\omega') 1_{\Omega \setminus A_k}(\omega) \tag{26}$$

for all $(\omega, \omega') \in \Omega \times \Omega'$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra such that $\sigma((A_k)_{k=1}^{K}) \subseteq \mathcal{G}$. Define $\mathcal{G}_0 = \mathcal{G} \otimes \{\emptyset, \Omega'\}$ and for $k \in \{1, \ldots, K\}$ define

$$\mathcal{G}_k = \sigma(\mathcal{G}_0, d_1, \ldots, d_k). \tag{27}$$

Then, the following holds:

(i) *For all $(\omega, \omega') \in \Omega \times \Omega'$ it holds that $d(\omega) = \sum_{k=1}^{K} d_k(\omega, \omega') 1_{A_k}(\omega)$.*
(ii) *$(d_k)_{k=1}^{K}$ is $(\mathcal{G}_k)_{k=1}^{K}$-adapted.*
(iii) *$\mathcal{L}(d_k) = \mu_k$.*
(iv) *$(\mathcal{G}_0, \sigma(d_1), \ldots, \sigma(d_K))$ are independent if and only if for all $k \in \{1, \ldots, K\}$, all $A \in \mathcal{G}$ satisfying $A \subseteq A_k$ and $\mathbb{P}(A) > 0$, and all $B \in \mathcal{B}(X)$ it holds that*

$$\mathbb{P}(d \in B \mid A) := \frac{\mathbb{P}(\{d \in B\} \cap A)}{\mathbb{P}(A)} = \mu_k(B). \tag{28}$$

Proof Claims (i), (ii) and (iii) are trivial. Regarding claim (iv), suppose that d_k is independent of $\mathcal{G}_{k-1}$ for all $k \in \{1, \ldots, K\}$, then in particular $(d_k)_{k=1}^{K}$ is independent of $\mathcal{G}_0$. Let $k \in \{1, \ldots, K\}$, $A \in \mathcal{G}$ satisfying $A \subseteq A_k$ and $\mathbb{P}(A) > 0$, and $B \in \mathcal{B}(X)$ be given. Then,

$$\begin{aligned}\mathbb{P}(d \in B \mid A) &= \frac{\mathbb{P}(\{d \in B\} \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P} \otimes \mathbb{P}'(\{d_k \in B\} \cap (A \times \Omega'))}{\mathbb{P} \otimes \mathbb{P}'(A \times \Omega')} \\ &= \mathbb{P} \otimes \mathbb{P}'(d_k \in B) = \mu_k(B),\end{aligned}$$

where we use (i), independence, and (iii). In order to prove the reverse implication, let $B_1, \ldots, B_K \in \mathcal{B}(X)$, let $k \in \{1, \ldots, K\}$, and let $A \in \mathcal{G}$ be such that $A \subseteq A_k$ and $\mathbb{P}(A) > 0$. It holds that

$$\begin{aligned}&\mathbb{P} \otimes \mathbb{P}'((A \times \Omega') \cap \{d_1 \in B_1, \ldots, d_K \in B_K\}) \\ &\quad = \mathbb{P} \otimes \mathbb{P}' \left[(A \cap \{d \in B_k\}) \times \left(\bigcap_{\ell \in \{1, \ldots, K\} \setminus \{k\}} \{d'_\ell \in B_\ell\} \right) \right] \\ &\quad = \mathbb{P}(A) \mathbb{P}(d \in B_k \mid A) \prod_{\ell \in \{1 \ldots, K\} \setminus \{k\}} \mathbb{P}'(d'_\ell \in B_\ell) = \mathbb{P}(A) \prod_{\ell \in \{1 \ldots, K\}} \mu_\ell(B_\ell).\end{aligned} \tag{29}$$

This suffices to prove the reverse implication. □

Lemma A.9 requires a special setting:

Setting A.8 *Let X be a separable Banach space, $N \in \mathbb{N}$ and let $(d_n)_{n=1}^N \in \mathcal{A}_{p\text{-simple}}(\Omega, \mathbb{P}, (\mathcal{F}_n)_{n=0}^N; X, \mathcal{P})$, i.e., for all $n \in \{1, \dots, N\}$ we have a $K_n \in \mathbb{N}$ such that*

(i) $\mathbb{P}(d_n \in \cdot \mid \mathcal{F}_{n-1}) = \sum_{k=1}^{K_n} \mu_{n,k} 1_{A_{n-1,k}}$ *a.s., where*
(ii) $A_{n-1,1}, \dots, A_{n-1,K_n} \in \mathcal{F}_{n-1}$ *is a partition of* Ω *with* $\mathbb{P}(A_{n-1,k}) > 0$ *and* $\mu_{n,1}, \dots, \mu_{n,K_n} \in \mathcal{P}$.

We set $K_0 := 1$ *and*

$$J_0 := \{(n,k) : n \in \{0, \dots, N\}, k \in \{1, \dots, K_n\}\},$$
$$J := \{(n,k) : n \in \{1, \dots, N\}, k \in \{1, \dots, K_n\}\}.$$

On J_0 we introduce the lexicographical order $(m, j) \prec (n, k)$ if either $m < n$ and $j \in \{1, \dots, K_m\}$ or $m = n$ and $j \in \{1, \dots, k\}$. For an auxiliary probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ and independent random variables $d'_{n,k} \colon \Omega' \to X$ with $\mathcal{L}(d'_{n,k}) = \mu_{n,k}$, $n \in \{1, \dots, N\}$, and $k \in \{1, \dots, K_n\}$ let

$$d_{n,k}(\omega, \omega') := d_n(\omega) 1_{A_{n-1,k}}(\omega) + d'_{n,k}(\omega') 1_{\Omega \setminus A_{n-1,k}}(\omega).$$

We define

$$\begin{aligned}
\mathcal{G}_{0,1} &:= \mathcal{F}_0 \otimes \{\emptyset, \Omega'\}, \\
\mathcal{G}_{n,k} &:= \mathcal{G}_{n-1,K_{n-1}} \vee \sigma(d_{n,1}, \dots, d_{n,k}) \quad (n \in \{1, \dots, N\}, k \in \{1, \dots, K_n - 1\}), \\
\mathcal{G}_{n,K_n} &:= \mathcal{G}_{n-1,K_{n-1}} \vee \sigma(d_{n,1}, \dots, d_{n,K_n}) \vee \Big(\mathcal{F}_n \otimes \{\emptyset, \Omega'\}\Big).
\end{aligned}$$

Finally, we let

(i) $K := K_1 + \dots + K_N$,
(ii) $(\mathcal{H}_\ell)_{\ell=0}^K$ *be the lexicographical ordering of* $(\mathcal{G}_{n,k})_{(n,k) \in J_0}$,
(iii) $(\varphi_\ell)_{\ell=1}^K$ *be the lexicographical ordering of* $(d_{n,k})_{(n,k) \in J}$,
(iv) $(A_\ell)_{\ell=0}^{K-1}$ *be the lexicographical ordering of* $(A_{n-1,k})_{(n,k) \in J}$,
(v) $(\varphi'_\ell)_{\ell=1}^K$ *be the lexicographical ordering of* $(d'_{n,k})_{(n,k) \in J}$.

Lemma A.9 *Assume Setting A.8. Then, the following holds true:*

(i) $(\varphi_\ell)_{\ell=1}^K$ *is* $(\mathcal{H}_\ell)_{\ell=1}^K$*-adapted and* φ_ℓ *is independent from* $\mathcal{H}_{\ell-1}$.
(ii) $A_\ell \in \mathcal{H}_\ell$ *for* $\ell \in \{0, \dots, K-1\}$.
(iii) $\sum_{n=1}^N d_n = \sum_{\ell=1}^K \varphi_\ell 1_{A_{\ell-1}}$.
(iv) *If* $e_n := \sum_{k=1}^{K_n} d'_{n,k} 1_{A_{n-1,k}}$, *then*

  (a) $(e_n)_{n=1}^N$ *is a decoupled tangent sequence of* $(d_n)_{n=1}^N$ *and*
  (b) $\sum_{n=1}^N e_n = \sum_{\ell=1}^K \varphi'_\ell 1_{A_{\ell-1}}$.

Proof (i) Fix $\ell \in \{1, \dots, K\}$. By definition, φ_ℓ is $\mathcal{H}_\ell$-measurable. To show that φ_ℓ is independent from $\mathcal{H}_{\ell-1}$, it is enough to verify that $\mathcal{G}_{n-1,K_{n-1}}, d_{n,1}, \dots, d_{n,K_n}$ are independent for $n \in \{1, \dots, N\}$. Because

$$\mathcal{G}_{n-1,K_{n-1}} = \mathcal{F}_{n-1} \otimes \sigma\big(d'_{m,j} : m \in \{1, \dots, n-1\}, j \in \{1, \dots, K_m\}\big),$$

where for $n = 1$ the second factor is replaced by $\{\emptyset, \Omega'\}$, it remains to check that

$$\Big(\mathcal{F}_{n-1} \otimes \{\emptyset, \Omega'\}\Big), d_{n,1}, \dots, d_{n,K_n}$$

are independent. But this follows from Lemma A.7. (ii) and (iv) follow by construction, (iii) from $d_n(\omega) = \sum_{k=1}^{K_n} d_{n,k}(\omega, \omega') 1_{A_{n-1,k}}(\omega)$ for $n \in \{1, \dots, N\}$. □

A.2 Proof of Theorem 4.3

(ii) $\Rightarrow$ (i). First, we remark that it is sufficient to consider the case $\Delta = \{\lambda_0\}$. We do so and denote Φ_{λ_0} simply by Φ.
Step 1 We show that (ii) implies the following:

(II) For every $N \in \mathbb{N}$, every stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^N)$, every sequence of independent random variables $(\varphi_n)_{n=1}^N$ with $\mathcal{L}(\varphi_n) \in \mathcal{P}_{p\text{-ext}}$ such that φ_n is $\mathcal{F}_n$-measurable and independent of $\mathcal{F}_{n-1}$, and for all $A_n \in \mathcal{F}_n$, $n \in \{0, \dots, N-1\}$, it holds that

$$\mathbb{E}\Phi\left(\sum_{n=1}^N \varphi_n 1_{A_{n-1}}, \sum_{n=1}^N \varphi'_n 1_{A_{n-1}}\right) \le 0,$$

whenever $(\varphi'_n)_{n=1}^N$ is a copy of $(\varphi_n)_{n=1}^N$ independent of $\mathcal{F}_N$.

(a) We verify that if (ii) holds, then (ii) remains valid with $\mathcal{P}$ replaced by $\mathcal{P}_{p\text{-ext}}$.

(a.0) First, we consider $(\varphi_n)_{n=1}^N$ such that for all $n \in \{1, \dots, N\}$ there exist a $K_n \in \mathbb{N}$ and $(\mu_{n,k})_{k=1}^{K_n} \subseteq \mathcal{P}$ with $\mathcal{L}(\varphi_n) = \mu_{n,1} * \cdots * \mu_{n,K_n}$. As $A_n \in \mathcal{F}_n^\varphi$ for $n \in \{0, \dots, N-1\}$ there are $B_n \in X^n$ such that $A_n = \{(\varphi_j)_{j=1}^n \in B_n\}$ for $n \ge 1$, whereas for $n = 0$ we have $A_0 \in \{\emptyset, \Omega\}$. Now let $((\psi_{n,k})_{k=1}^{K_n})_{n=1}^N$ be independent random variables satisfying $\mathcal{L}(\psi_{n,k}) = \mu_{n,k}$. Then, it holds in distribution that

$$\varphi_1 1_{A_0} + \sum_{n=2}^N \varphi_n 1_{\{(\varphi_j)_{j=1}^{n-1} \in B_{n-1}\}} \overset{d}{=} \sum_{k=1}^{K_1} \psi_{1,k} 1_{A_0} + \sum_{n=2}^N \sum_{k=1}^{K_n} \psi_{n,k} 1_{\{(\sum_{\ell=1}^{K_j} \psi_{j,\ell})_{j=1}^{n-1} \in B_{n-1}\}}$$

and

$$\varphi_1' 1_{A_0} + \sum_{n=2}^{N} \varphi_n' 1_{\{(\varphi_j)_{j=1}^{n-1} \in B_{n-1}\}} \stackrel{d}{=} \sum_{k=1}^{K_1} \psi_{1,k}' 1_{A_0} + \sum_{n=2}^{N} \sum_{k=1}^{K_n} \psi_{n,k}' 1_{\{(\sum_{\ell=1}^{K_j} \psi_{j,\ell})_{j=1}^{n-1} \in B_{n-1}\}}$$

where $((\psi_{n,k}')_{k=1}^{K_n})_{n=1}^{N}$ is an independent copy of $((\psi_{n,k})_{k=1}^{K_n})_{n=1}^{N}$. Hence, if (ii) holds, then (ii) remains valid with $\mathcal{P}$ replaced by the set of finite convolutions of elements from $\mathcal{P}$.

(a.1) Assume that $\mathcal{L}(\varphi_n) \in \mathcal{P}_{p\text{-ext}}$ for $n \in \{1, \ldots, N\}$. By Lemmas A.4 and 4.4, we can restrict the A_n to $A_n = \{(\varphi_j)_{j=1}^{n} \in B_n\}$ with $\mathbb{P}((\varphi_j)_{j=1}^{n} \in \partial B_n) = 0$ and $B_n \in \mathcal{B}(X^n)$ for $n \in 1, \ldots, N-1$, whereas we keep $A_0 \in \{\emptyset, \Omega\}$. We find a uniformly $\mathcal{L}^p$-integrable family of independent random variables $(\varphi_{n,k})_{n=1,k=1}^{N,\infty}$ such that $\mathcal{L}(\varphi_{n,k})$ is a finite convolution of measures from $\mathcal{P}$ and $\mathcal{L}(\varphi_{n,k}) \stackrel{w^*}{\to} \mathcal{L}(\varphi_n)$ as $k \to \infty$. Lemma A.5 gives

$$\left(\varphi_{1,k} 1_{A_0} + \sum_{n=2}^{N} \varphi_{n,k} 1_{\{(\varphi_{j,k})_{j=1}^{n-1} \in B_{n-1}\}}, \varphi_{1,k}' 1_{A_0} + \sum_{n=2}^{N} \varphi_{n,k}' 1_{\{(\varphi_{j,k})_{j=1}^{n-1} \in B_{n-1}\}}\right)$$
$$\xrightarrow{w^*} \left(\varphi_1 1_{A_0} + \sum_{n=2}^{N} \varphi_n 1_{\{(\varphi_j)_{j=1}^{n-1} \in B_{n-1}\}}, \varphi_1' 1_{A_0} + \sum_{n=2}^{N} \varphi_n' 1_{\{(\varphi_j)_{j=1}^{n-1} \in B_{n-1}\}}\right)$$

as $k \to \infty$. By Lemma 4.4, we conclude step (a), i.e., (ii) is valid for $\mathcal{P}_{p\text{-ext}}$.

(b) We now prove that if (ii) is valid for $\mathcal{P}_{p\text{-ext}}$, then (II) holds.

(b.0) Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_{n=0}^{N})$ and $(\varphi_n)_{n=1}^{N}$ be as in (II), and set $\mu_n := \mathcal{L}(\varphi_n)$. If each μ_n is a Dirac measure in an $x_n \in X$, then $\mathbb{E}\Phi(\sum_{n=1}^{N} \varphi_n 1_{A_{n-1}}, \sum_{n=1}^{N} \varphi_n' 1_{A_{n-1}})$ is a weighted sum of terms $\Phi(\sum_{n \in I} x_n, \sum_{n \in I} x_n)$ with $I \subseteq \{1, \ldots, n\}$ (the empty sum is treated as zero). In this case, (ii) implies that each of these terms in non-positive, so that in what follows we assume there exists an $\ell \in \{1, \ldots, N\}$ such that μ_ℓ is *not* a Dirac measure. We will prove that for all $\varepsilon > 0$ it holds that

$$\mathbb{E}\Phi\left(\sum_{n=1}^{N} \varphi_n 1_{A_{n-1}}, \sum_{n=1}^{N} \varphi_n' 1_{A_{n-1}}\right) < 3\varepsilon, \tag{30}$$

which completes the proof of (II). By passing to the larger probability space $(\Omega, \mathcal{F}, \mathbb{P}) \otimes ([0,1], \mathcal{B}([0,1]), \lambda)$ (where λ is the Lebesgue measure), endowed with the filtration $(\mathcal{F}_n \otimes \mathcal{B}([0,1]))_{n=0}^{N}$, we may assume that $(\Omega, \mathcal{F}_n, \mathbb{P})$ is divisible for all $n \in \{0, \ldots, N\}$. Fix $\varepsilon > 0$, and let P_ε be as in Lemma A.3 with the property that $\mu_\ell \circ P_\varepsilon^{-1}$ is not a Dirac measure. Recall from Lemma A.3 that

$$\mathbb{E}\left|\Phi\left(\sum_{n=1}^{N} \varphi_n 1_{A_{n-1}}, \sum_{n=1}^{N} \varphi_n' 1_{A_{n-1}}\right) - \Phi\left(\sum_{n=1}^{N} P_\varepsilon(\varphi_n) 1_{A_{n-1}}, \sum_{n=1}^{N} P_\varepsilon(\varphi_n') 1_{A_{n-1}}\right)\right| < \varepsilon. \tag{31}$$

(b.1) Set $\mu_{\ell,\varepsilon} := \mu_\ell \circ P_\varepsilon^{-1}$ and recall that $\mu_{\ell,\varepsilon}$ is not a single Dirac measure, but a finite sum of Dirac measures. For $m \in \mathbb{N}$, we proceed as follows:

Step 0 We apply Lemma A.2 with $\mathcal{G} = \{\emptyset, \Omega\}$ and $\mathcal{F} = \mathcal{F}_0$ to find a $k_{m,1} \in \{2, 3, \ldots\}$ and a sequence of independent, $\mathcal{F}_0$-measurable, $\mu_{\ell,\varepsilon}$-distributed random variables $(\psi_{m,1}, \ldots, \psi_{m,k_{m,1}-1})$, and to find an $A_{m,0} \in \sigma(\psi_{m,1}, \ldots, \psi_{m,k_{m,1}-1})$ with $\|1_{A_0} - 1_{A_{m,0}}\|_{\mathcal{L}^p(\mathbb{P})} \leq 2^{-m}$. Set $\psi_{m,k_{m,1}} := P_\varepsilon(\varphi_1)$.
Step n; $n = 1, \ldots, N-1$ Apply Lemma A.2 with $\mathcal{G} = \sigma(\psi_{m,1}, \ldots, \psi_{m,k_{m,n}})$ (note that $\psi_{m,k_{m,1}} = P_\varepsilon(\varphi_1), \ldots, \psi_{m,k_{m,n}} = P_\varepsilon(\varphi_n)$) and $\mathcal{F} = \mathcal{F}_{n+1}$, to find a $k_{m,n+1} \in \{k_{m,n}+2, k_{m,n}+3, \ldots\}$ and independent, $\mathcal{F}_{n+1}$-measurable, $\mu_{\ell,\varepsilon}$-distributed random variables $(\psi_{m,k_{m,n}+1}, \ldots, \psi_{m,k_{m,n+1}-1})$, independent of the σ-algebra $\sigma(\psi_{m,1}, \ldots, \psi_{m,k_{m,n}})$ as well, and to find an $A_{m,n} \in \sigma(\psi_{m,1}, \ldots, \psi_{m,k_{m,n+1}-1})$ with $\|1_{A_n} - 1_{A_{m,n}}\|_{\mathcal{L}^p(\mathbb{P})} \leq 2^{-m}$. Set $\psi_{m,k_{m,n+1}} := P_\varepsilon(\varphi_{n+1})$.

(b.2) By construction,

$$\begin{aligned}
&\lim_{m\to\infty} \left\| \sum_{n=1}^N (1_{A_{m,n-1}} - 1_{A_{n-1}}) P_\varepsilon(\varphi_n) \right\|_{\mathcal{L}^p(\mathbb{P};X)} \\
&= \lim_{m\to\infty} \left\| \sum_{n=1}^N (1_{A_{m,n-1}} - 1_{A_{n-1}}) P_\varepsilon(\varphi_n') \right\|_{\mathcal{L}^p(\mathbb{P};X)} = 0,
\end{aligned}$$

hence by Lemma 4.4 there exists an $M \in \mathbb{N}$ such that, for $B_{n-1} = A_{M,n-1}$, $n \in \{1, \ldots, N\}$,

$$\begin{aligned}
&\left| \mathbb{E}\Phi\left(\sum_{n=1}^N P_\varepsilon(\varphi_n) 1_{B_{n-1}}, \sum_{n=1}^N P_\varepsilon(\varphi_n') 1_{B_{n-1}} \right) \right. \\
&\quad \left. - \mathbb{E}\Phi\left(\sum_{n=1}^N P_\varepsilon(\varphi_n) 1_{A_{n-1}}, \sum_{n=1}^N P_\varepsilon(\varphi_n') 1_{A_{n-1}} \right) \right| < \varepsilon. \qquad (32)
\end{aligned}$$

(b.3) We would like to apply (ii), verified for $\mathcal{P}_{p\text{-ext}}$ in part (a), to $\sum_{n=1}^N 1_{B_{n-1}} \varphi_n$, however, our construction of B_{n-1} only guarantees that $P_\varepsilon(\varphi_n)$ is independent of B_{n-1}, *not* that φ_n is independent of B_{n-1}. Hence, we proceed as follows. Let $(\phi_k)_{k=1}^{k_{M,N}}$ be a sequence of independent random variables such that

$$\mathcal{L}(\phi_k) := \begin{cases} \mu_n : k = k_{M,n}, \\ \mu_\ell : \text{else} \end{cases}.$$

Recall that μ_ℓ was specified in (b.0). By the Factorization Lemma for all $n \in \{1, \ldots, N\}$, there exists a $C_{n-1} \in X^{k_{M,n}-1}$ such that

$$B_{n-1} = \left\{ (P_\varepsilon(\varphi_1), \ldots, P_\varepsilon(\varphi_{k_{M,n}-1})) \in C_{n-1} \right\}.$$

Define, for $n \in \{1, \dots, N\}$,

$$\tilde{B}_{n-1} := \Big\{ \big(P_\varepsilon(\phi_1), \dots, P_\varepsilon(\phi_{k_{M,n}-1})\big) \in C_{n-1} \Big\}.$$

By construction,

$$\left(\sum_{n=1}^N 1_{B_{n-1}} P_\varepsilon(\varphi_n), \sum_{n=1}^N 1_{B_{n-1}} P_\varepsilon(\varphi_n') \right) \stackrel{d}{=} \left(\sum_{n=1}^N 1_{\tilde{B}_{n-1}} P_\varepsilon(\phi_{k_n}), \sum_{n=1}^N 1_{\tilde{B}_{n-1}} P_\varepsilon(\phi_{k_n}') \right),$$

where $(\phi_k')_{k=1}^{k_{M,N}}$ is an independent copy of $(\phi_k)_{k=1}^{k_{M,N}}$, so that

$$\begin{aligned} &\mathbb{E}\Phi\left(\sum_{n=1}^N P_\varepsilon(\varphi_n) 1_{B_{n-1}}, \sum_{n=1}^N P_\varepsilon(\varphi_n') 1_{B_{n-1}} \right) \\ &= \mathbb{E}\Phi\left(\sum_{n=1}^N P_\varepsilon(\phi_{k_{M,n}}) 1_{\tilde{B}_{n-1}}, \sum_{n=1}^N P_\varepsilon(\phi_{k_{M,n}}') 1_{\tilde{B}_{n-1}} \right). \end{aligned} \tag{33}$$

By the assumption that (ii) holds for $\mathcal{L}(\varphi_n) \in \mathcal{P}_{p\text{-ext}}$, it follows that

$$\mathbb{E}\Phi\left(\sum_{n=1}^N \phi_{k_n} 1_{\tilde{B}_{n-1}}, \sum_{n=1}^N \phi_{k_n}' 1_{\tilde{B}_{n-1}} \right) \le 0.$$

The inequality above in combination with (19) implies that

$$\mathbb{E}\Phi\left(\sum_{n=1}^N P_\varepsilon(\phi_{k_n}) 1_{\tilde{B}_{n-1}}, \sum_{n=1}^N P_\varepsilon(\phi_{k_n}') 1_{\tilde{B}_{n-1}} \right) < \varepsilon.$$

Combining the inequality above with (33), (32), and (31), we arrive at the desired estimate (30).

Step 2 Lemma A.9 (with $\mathcal{P} = \mathcal{P}_{p\text{-ext}}$) combined with statement (II) above implies that Item (i) of Theorem 4.3 holds for all $(d_n)_{n=1}^N \in \mathcal{A}_{p\text{-simple}}(\Omega, (\mathcal{F}_n)_{n=1}^N; X, \mathcal{P}_{p\text{-ext}})$ (see Definition A.6).

Step 3 Let $(d_n)_{n\in\mathbb{N}}$ be as in (i) of Theorem 4.3 with $d_n \equiv 0$ if $n > N$ for some $N \in \mathbb{N}$. By Corollary 3.2 we obtain for $n \in \{1, \dots, N\}$ random variables $d_n^0 \colon \Omega \times [0,1] \to X$ and $H_n \colon \Omega \times (0,1] \to [0,1]$ such that d_n^0 is $\mathcal{F}_{n-1} \otimes \mathcal{B}([0,1])/\mathcal{B}(X)$-measurable, H_n is independent of $\mathcal{F}_{n-1} \otimes \{\emptyset, (0,1]\}$ with $\mathcal{L}(H_n) = \lambda$, and $d_n(\omega) = d_n^0(\omega, H_n(\omega, s))$ for all $(\omega, s) \in \Omega_n \times (0,1]$ for some $\Omega_n \in \mathcal{F}_n$ of measure one. We define $\bar{\Omega} := \Omega \times (0,1]^N$, $\bar{\mathcal{F}}_0 := \mathcal{F}_0 \otimes \{\emptyset, (0,1]^N\}$, $\bar{\mathcal{F}}_n := \mathcal{F}_n \otimes \sigma(\pi_1, \dots \pi_n)$, where $\pi_n : (0,1]^N \to (0,1]$ is the projection onto the n-th coordinate, and $\bar{\mathbb{P}} := \mathbb{P} \otimes \lambda_N$ where λ_N is the Lebesgue measure on $(0,1]^N$. Then, $\bar{H}_n, \bar{d}_n, \bar{e}_n \colon \bar{\Omega} \to [0,1]$ are given by $\bar{H}_n(\omega, s) := H_n(\omega, \pi_n(s))$, $\bar{d}_n(\omega, s) := d_n^0(\omega, \bar{H}_n(\omega, s))$, and $\bar{e}_n(\omega, s) := d_n^0(\omega, \pi_n(s))$, $n \in \{1, \dots, N\}$. We get:

(i) $\{(\omega, s) \in \bar{\Omega} \colon \bar{d}_n(\omega, s) = d_n(\omega)\} \supseteq \Omega_n \times (0, 1]$.
(ii) The fact that $\bar{H}_n$ is uniformly $[0, 1]$ distributed and independent of $\bar{\mathcal{F}}_{n-1}$ and d_n^0 is $\mathcal{F}_{n-1} \otimes \mathcal{B}([0, 1])/\mathcal{B}(X)$-measurable implies that $(\bar{e}_n)_{n\in\mathbb{N}}$ is an $(\bar{\mathcal{F}}_n)_{n=1}^N$-decoupled tangent sequence of $(\bar{d}_n)_{n\in\mathbb{N}}$.
(iii) The function $\kappa_{n-1}[\omega, B] := \lambda(\{h \in [0, 1] : d_n^0(\omega, h) \in B\})$, $B \in \mathcal{B}(X)$, is a regular conditional probability for $\mathbb{P}(d_n \in \cdot \mid \mathcal{F}_{n-1})$.

Next note that for all $\varepsilon > 0$ and $n \in \{1, \dots, N\}$ there exists an $d_n^{0,\varepsilon} \colon \Omega \times [0, 1] \to X$ which satisfies

$$d_n^{0,\varepsilon}(\omega, s) = \sum_{k=1}^{m_\varepsilon} 1_{F_{n-1,k,\varepsilon}}(\omega) f_{n,k,\varepsilon}(s),$$

with $m_\varepsilon \in \mathbb{N}$, $f_{n,k,\varepsilon} \in \mathcal{L}^p([0, 1]; X)$, pair-wise disjoint $F_{n-1,1,\varepsilon}, \dots, F_{n-1,m_\varepsilon,\varepsilon} \in \mathcal{F}_{n-1}$ of positive measure, and $\|d_n^{0,\varepsilon} - d_n^0\|_{\mathcal{L}^p(\Omega\times[0,1];X)} < \varepsilon$. Moreover, we can pick $f_{n,k,\varepsilon}$ such that $\mathcal{L}(f_{n,k,\varepsilon}) \in \mathcal{P}_{p\text{-ext}}$. Indeed, let $(f_\ell)_{\ell\in\mathbb{N}} \subset \mathcal{L}^p([0, 1]; X)$ be dense. For $\varepsilon > 0$, we can choose an appropriate $\eta = \eta(p, \varepsilon) > 0$ and define

$$\begin{aligned} S_1 &:= \{\omega \in \Omega : \|d_n^0(\omega, \cdot) - f_1\|_{\mathcal{L}^p([0,1];X)} < \eta\}, \\ S_2 &:= \{\omega \in \Omega : \|d_n^0(\omega, \cdot) - f_2\|_{\mathcal{L}^p([0,1];X)} < \eta\} \backslash S_1, \dots. \end{aligned}$$

From these sets $(S_n)_{n\in\mathbb{N}}$, we extract the collection $(F_{n-1,k,\varepsilon})_{k=1}^{m_\varepsilon}$, find $\omega_k \in F_{n-1,k,\varepsilon} \cap \kappa_{n-1}^{-1}(\mathcal{P})$, and let $f_{n,k,\varepsilon} := d_n^0(\omega_k, \cdot)$. We continue and define $\bar{d}_n^\varepsilon, \bar{e}_n^\varepsilon \colon \bar{\Omega} \to [0, 1]$ by $\bar{d}_n^\varepsilon(\omega, s) := d_n^{0,\varepsilon}(\omega, \bar{H}_n(\omega, s))$, and $\bar{e}_n^\varepsilon(\omega, s) := d_n^{0,\varepsilon}(\omega, \pi_n(s))$, $n \in \{1, \dots, N\}$. By construction, we have

(i) $\|\bar{e}_n^\varepsilon - \bar{e}_n\|_{\mathcal{L}^p(\bar{\Omega};X)} = \|\bar{d}_n^\varepsilon - \bar{d}_n\|_{\mathcal{L}^p(\bar{\Omega};X)} < \varepsilon$ for all $n \in \{1, \dots, N\}$,
(ii) $(\bar{e}_n^\varepsilon)_{n=1}^N$ is an $(\bar{\mathcal{F}}_n)_{n=1}^N$-decoupled tangent sequence of $(\bar{d}_n^{\,\varepsilon})_{n\in\mathbb{N}}$,
(iii) a conditional regular conditional probability kernel for $\mathcal{L}(d_n^\varepsilon \mid \bar{\mathcal{F}}_{n-1})$ is given by $\sum_{k=1}^{m_\varepsilon} 1_{F_{n-1,k,\varepsilon}}(\omega)\mathcal{L}(f_{n,k,\varepsilon})$.

As $(d_n^{0,\varepsilon})_{n=1}^N \in \mathcal{A}_{p\text{-simple}}(\Omega, (\mathcal{F}_n)_{n=0}^N; X, \mathcal{P}_{p\text{-ext}})$, this concludes the proof that (ii) implies (i) in Theorem 4.3.

References

1. Billingsley, P.: Probability and Measure, 3rd edn. Wiley, New York (1995)
2. Bourgain, J.: Some remarks on Banach spaces in which martingale difference sequences are unconditional. Ark. Mat. **21**, 163–168 (1983)
3. Brown, B.M.: Characteristic functions, moments, and the central limit theorem. Ann. Math. Stat. **41**, 658–664 (1970)
4. Burkholder, D.L.: A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions. In: Conference on Harmonic Analysis in Honor of Antoni Zygmund, Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983, vol. I, II, pp. 270–286 (1981)
5. Burkholder, D.L.: Sharp inequalities for martingales and stochastic integrals. Astérisque, (157-158):75–94 (1988). Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987)
6. Carmona, R., Delarue, F.: Probabilistic Theory of Mean Field Games with Applications I. Springer, Berlin (2018)

7. Cioica-Licht, P., Cox, S., Veraar, M.: Stochastic integration in quasi-Banach spaces. To appear in Studia Mathematica. Preprint at arXiv:1804.08947 (2018)
8. Cox, S.G., Geiss, S.: On decoupling in Banach spaces. arXiv:1805.12377
9. Cox, S.G., Veraar, M.C.: Vector-valued decoupling and the Burkholder–Davis–Gundy inequality. Illinois J. Math. **55**(1), 343–375 (2012). 2011
10. de la Peña, V.H., Giné, E.: Decoupling. Probability and its Applications (New York). Springer, New York (1999). From dependence to independence, Randomly stopped processes. U-statistics and processes. Martingales and beyond
11. Dudley, R.M.: Real Analysis and Probability. Cambridge Studies in Advanced Mathematics (2002)
12. Garling, D.J.H.: Brownian motion and UMD-spaces. In: Probability and Banach spaces (Zaragoza, 1985), Lecture Notes in Math., vol. 1221, pp. 36–49. Springer, Berlin (1986)
13. Garling, D.J.H.: Random martingale transform inequalities. In: Probability in Banach spaces 6 (Sandbjerg, 1986), Progr. Probab., vol. 20, pp. 101–119. Birkhäuser Boston, Boston (1990)
14. Geiss, S.: BMO_ψ-spaces and applications to extrapolation theory. Studia Math. **122**(3), 235–274 (1997)
15. Hitczenko, P.: On tangent sequences of UMD-space valued random vectors. Unpublished
16. Hitczenko, P.: On a domination of sums of random variables by sums of conditionally independent ones. Ann. Probab. **22**(1), 453–468 (1994)
17. Hytönen, T., van Neerven, J., Veraar, M., Weis, L.: Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 63. Springer, Cham (2016)
18. Hytönen, T., van Neerven, J., Veraar, M., Weis, L.: Analysis in Banach spaces. Vol. II, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 67. Springer, Cham (2017). Probabilistic methods and operator theory
19. Kallenberg, O.: Foundations of Modern Probability. Probability and its Applications (New York), 2nd edn. Springer, New York (2002)
20. Kallenberg, O.: Tangential existence and comparison with applications to single and multiple integration. Probab. Math. Stat. **37**(1), 21–52 (2017)
21. Klass, M.J.: A best possible improvement of Wald's equation. Ann. Probab. **16**(2), 840–853 (1988)
22. Klass, M.J.: Uniform lower bounds for randomly stopped Banach space valued random sums. Ann. Probab. **18**(2), 790–809 (1990)
23. Kwapień, S., Woyczyński, W.A.: Tangent sequences of random variables: basic inequalities and their applications. In: Almost everywhere convergence, pp. 237–265. Academic Press, Boston (1989)
24. Kwapień, S., Woyczyński, W.A.: Semimartingale integrals via decoupling and tangent processes. Probab. Math. Stat. **12**(2), 165–200 (1991)
25. Kwapień, S., Woyczyński, W.A.: Random Series and Stochastic Integrals: Single and Multiple. Probability and its Applications. Birkhäuser Boston Inc., Boston (1992)
26. Ledoux, M., Talagrand, M.: Probability in Banach Spaces. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23. Springer, Berlin (1991). Isoperimetry and processes
27. Maurey, B.: Système de Haar. In: Séminaire Maurey-Schwartz 1974–1975: Espaces Lsup p, applications radonifiantes et géométrie des espaces de Banach, Exp. Nos. I et II, p. 26, pp. (erratum, p. 1). Centre Math., École Polytech., Paris (1975)
28. McConnell, T.R.: Decoupling and stochastic integration in UMD Banach spaces. Probab. Math. Stat. **10**(2), 283–295 (1989)
29. Montgomery-Smith, S.: Concrete representation of martingales. Electron. J. Probab. **3**(15), 15 (1998)
30. Parthasarathy, K.R.: Probability measures on metric spaces. Academic Press, London (1967)
31. Pisier, G.: Martingales with values in uniformly convex spaces. Isr. J. Math. **20**, 326–350 (1975)
32. Pisier, G.: The Volume of Convex Bodies and Banach Space Geometry. Cambridge University Press, Cambridge (1989)
33. van Neerven, J.M.A.M.: Stochastic Evolution Equations. 2008. Lecture Notes to the 10th Internet Seminar. http://fa.its.tudelft.nl/~neerven/
34. van Neerven, J.M.A.M., Veraar, M.C., Weis, L.: Stochastic integration in UMD Banach spaces. Ann. Probab. **35**(4), 1438–1478 (2007)
35. van Neerven, J.M.A.M.: γ-Radonifying operators—a survey. In: The AMSI-ANU Workshop on Spectral Theory and Harmonic Analysis. Proc. Centre Math. Appl. Austral. Nat. Univ., vol. 44, pp. 1–61. Austral. Nat. Univ., Canberra (2010)

36. van Neerven, J.M.A.M., Veraar, M., Weis, L.: Stochastic integration in Banach spaces—a survey. In: Stochastic Analysis: A Series of Lectures. Progr. Probab., vol. 68, pp. 297–332. Birkhäuser/Springer, Basel (2015)
37. Veraar, M.C.: Randomized UMD Banach spaces and decoupling inequalities for stochastic integrals. Proc. Am. Math. Soc. **135**(5), 1477–1486 (2007). (Electronic)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.