SOME APPLICATIONS OF THE MALLIAVIN CALCULUS TO SUB-GAUSSIAN AND NON-SUB-GAUSSIAN RANDOM FIELDS

by

Andrew B. Vizcarra
Purdue University

Frederi G. Viens
Purdue University

Technical Report #06-01

Department of Statistics
Purdue University

February 2006

Some applications of the Malliavin calculus to sub-Gaussian and non-sub-Gaussian random fields.

Andrew B. Vizcarra[1] Frederi G. Viens[2]

[1]Dept. Mathematics; avizcarr@math.purdue.edu,

[2] Dept. Statistics and Dept. Mathematics; viens@purdue.edu,

Purdue University
150 N. University St.
West Lafayette, IN 47907-2067, USA.

February 1, 2006

Abstract

We introduce a boundedness condition on the Malliavin derivative of a random variable to study sub-Gaussian and other non-Gaussian properties of functionals of random fields, with particular attention to the estimation of suprema. We relate the boundedness of nth Malliavin derivatives to a new class of "sub-nth Gaussian chaos" processes. An expected supremum estimation, extending the Dudley-Fernique theorem, is proved for such processes. Sub-nth Gaussian chaos concentration inequalities for the supremum are obtained, using Malliavin derivative conditions; for $n = 1$, this generalizes the Borell-Sudakov inequality to a class of sub-Gaussian processes, with a particularly simple and efficient proof; for $n = 2$ a natural extension to sub-2nd Gaussian chaos processes is established; for $n \geq 3$ a slightly less efficient Malliavin derivative condition is needed.

Key words and phrases:.Stochastic analysis, Malliavin derivative, Wiener chaos, sub-Gaussian process, concentration, suprema of processes, Dudley-Fernique theorem, Borell-Sudakov inequality.

AMS 2000 MSC codes: primary 60H07; secondary 60G15, 60G17.

1 Introduction

Gaussian analysis, and in particular the Malliavin calculus, are powerful and versatile tools in contemporary probability theory and stochastic analysis. The latter has applications ranging from other areas of probability theory, to physics, to finance, to name a few; a very short selection of references might include [2], [5], [6], [7], [12], [13], [14], [15], [16], [17], [22]. We will not attempt to give an overview of such a wide array of areas. Instead, this article presents a new way of using Malliavin derivatives to uncover sub-Gaussian and other non-Gaussian properties of functionals of random fields, with particular attention to the estimation of suprema.

After introducing some standard material on Wiener chaoses and the Malliavin derivative in what we hope is a streamlined and didactic way (Section 2), we introduce the fundamental lemma that serves as a basis and a springboard for non-Gaussian results: it is the observation that if a random variable X has a Malliavin derivative whose norm in $L^2[0,1]$ is almost surely bounded, then X is sub-Gaussian (Lemma 3.3). In Section 3, this lemma is exploited to analyze sub-Gaussian processes. Even though the proofs of the results therein are quite elementary, we believe they may have far-reaching consequences in probability and its applications. For example, even though it is not stated so explicitly, Lemma 3.3 is the key ingredient in the new proofs of existence of Lyapunov exponents for the continuous space stochastic Anderson model and the Brownian directed polymer in a Gaussian environment, obtained respectively in [8] and [18]; these existence results had been open problems for many years (see e.g. [4]). Lemma 3.3, and its application to sub-Gaussian deviations of the supremum of a sub-Gaussian random field (Theorem 3.6, which is a generalization of the so-called Borell-Sudakov inequality, see [1]), are techniques applied in [21] for statistical estimation problems for non-linear fractional Brownian functionals.

Inspired by the power of such applications, we postulate that in order to generalize the concept of sub-Gaussian random variables, one would be well-advised to investigate the properties of random fields whose nth Malliavin derivatives are bounded. Our study chooses to define the concept of *sub-nth Gaussian chaos* (or *sub-nth chaos*, for short) random fields slightly differently, in order to facilitate the study of such processes' concentration properties as well as those of their suprema. This is done in Section 4, which also includes an analysis of the relation between the sub-nth chaos property and boundedness of nth Malliavin derivatives. Our proofs in Section 4 are inspired by some of the techniques that worked well in the sub-Gaussian case of Section 3; yet when $n \geq 3$, many technical difficulties arise, and our work opens up as many new problems as it solves in that case.

While we prefer to provide full statements of our results in the main body of this paper, we include here some typical consequences of our work under a simplifying assumption which is nonetheless relevant for some applications, leaving it to the reader to check that the results now given do follow from our theorems.

Assumption Let n be a positive integer. Let X be a centered separable random field on an index set I. Assume that there exists a non-random metric δ on $I \times I$ such that almost surely, for all $x, y \in I$, for all $0 \leq s_n \leq \cdots \leq s_2 \leq s_1 \leq 1$,

$$\left| D_{s_n} \cdots D_{s_2} D_{s_1} \left(X(x) - X(y) \right) \right| \leq \delta(x, y) . \tag{1}$$

Conclusions Let $N(\varepsilon)$ be the smallest number of balls of radius ε in the metric δ needed to cover I. There is a constant C_n depending only on n such that, if the assumption above holds, the following conclusions hold:

Sub-nth Gaussian chaos property: (see Theorem 4.7)

$$\mathbf{E}\left[\exp\left(\frac{1}{C_n} \left| \frac{X(x) - X(y)}{\delta(s, y)} \right|^{2/n} \right) \right] \leq 2;$$

sub-nth Gaussian chaos extension of the Dudley-Fernique upper bound : (see Theorem 4.5)

$$\mu := \mathbf{E}\left[\sup_{x \in I} X(x) \right] \leq C_n \int_0^\infty \left(\log N(\varepsilon) \right)^{n/2} d\varepsilon;$$

sub-nth Gaussian chaos extension of the Borell-Sudakov concentration inequality : (see Corollary 4.15). With

$$\sigma = \operatorname*{ess\,sup}_{\omega \in \Omega} \left\{ \sup \left| D_{s_n} \cdots D_{s_2} D_{s_1} X(x) \right| : x \in I; 0 \leq s_n \leq \cdots \leq s_2 \leq s_1 \leq 1 \right\},$$

for all $\varepsilon > 0$, for u large enough,

$$\mathbf{P}\left[\left| \sup_{x \in I} X(x) - \mu \right| > u \right] \leq 2(1 + \varepsilon) \exp\left(-\frac{1}{(1 + \varepsilon)} \left(\frac{u}{\sigma} \right)^{2/n} \right).$$

It should be noted that in the sub-2nd-Gaussian chaos case ($n=2$), we prove (Theorem 4.5, Theorem 4.7 case $n=2$, Corollary 4.12) the three "Conclusions" above hold under the considerably weaker condition: almost surely,

$$\int\cdots\int_{[0,1]^n} |D_{s_n}\cdots D_{s_2}D_{s_1}(X(x)-X(y))|^2\, ds_1 ds_2\cdots ds_n \leq \delta(x,y). \tag{2}$$

When $n\geq 3$, the conditions we need to draw the above conclusions are intermediate between (1) and (2). However, we conjecture that the conclusions should hold under conditions much closer to (2). When $n=1$, the Dudley-Fernique theorem has been known for many years (see [11]) if one assumes the conclusion of Lemma 3.3; our interpretation of this Lemma appears to be new, although its proof below clearly shows it is a translation of Ustunel's [22, Theorem 9.1.1]; however, our proof of the Borel-Sudakov inequality (Theorem 3.6) under the hypotheses of Lemma 3.3 is new, and the inequality itself might be new for any class of non-Gaussian processes insofar as it does not seem to appear in the literature.

In addition to the obvious practical significance of results such as the "Conclusions" above, we think the reader familiar with classical proofs of such results as the Borell-Sudakov inequality and the Dudley-Fernique theorem, will appreciate the power of Malliavin derivatives: they provide, in Section 3 ($n=1$), stronger results with elegant, simpler proofs. We hope that beyond the issue of sharpening the results in Section 4 ($n\geq 3$) to come closer to Condition (2), this paper will encourage the reader to use our Malliavin-derivative based concentration inequalities in sub-Gaussian and non-sub-Gaussian settings, such as to study the almost-sure moduli of continuity of random fields to extend classical results (see [1] or [20]).

We wish to thank the three organizers of the Fifth Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, Switzerland, 2005) for providing the impetus for the research which led to this paper.

2 Preliminairies

In this didactic section, we present some basic facts about Wiener chaoses and the Malliavin calculus, largely with only sketches of proofs, to be used in the remainder of the article, and as a general quick reference guide. Excellent and complete treatment of these results and many more can be found for instance in the monographs [16] and [22]; both have been a constant source of inspiration for us.

We begin with a Brownian motion $W=\{W(t):t\in[0,1\}$ defined on a complete probability space $(\Omega,\mathcal{F},\mathbf{P})$ and adapted to a filtration $(\mathcal{F}_t)_{t\in[0,1]}$ satisfying the usual conditions (see [9]). With dr representing the Lebesgue measure, the Wiener integral $W(f)=\int_0^1 f(r)\,dW(r)$ of a non-random $f\in\mathcal{H}:=L^2([0,1],dr)$ is a centered Gaussian random variable with variance $||f||_{\mathcal{H}}^2=\int_0^1 f^2(r)\,dr$; the set $\mathcal{H}_1$ of all Wiener integrals $W(f)$ when f ranges over all of $\mathcal{H}$ is a set of jointly Gaussian random variables called the first Wiener chaos of W, or Gaussian space of W, whose entire finite-dimensional distributions are thus defined via the formula $\mathbf{E}W(f)W(g)=\langle f;g\rangle_{\mathcal{H}}=\int_0^1 f(r)g(r)\,dr$. The Wiener integral coincides with the Itô integral on $\mathcal{H}_1$, which can be seen via several different procedures, including the fact that both can be approximated in $L^2(\Omega)$ by the same Riemann sums. To construct chaoses of higher order, one may for example use iterated Itô integration. Denote $I_0(f)=f$ for any non-random constant f. Assume by induction that for any $g\in\mathcal{H}^{\otimes n}$, for almost every $(t,\omega)\in L^2([0,1]\times\Omega,dr d\mathbf{P})$,

$$I_n(g)=n!\int_0^1\int_0^{s_1}\cdots\int_0^{s_{n-1}} g(s_1,s_2,\cdots,s_n)\,dW(s_n)\cdots dW(s_2)\,dW(s_1) \tag{3}$$

has been defined. Given a symmetric function $f\in\mathcal{H}^{\otimes n+1}$, let

$$g_t(s_1,s_2,\cdots,s_n)=f(t,s_1,s_2,\cdots,s_n)\,\mathbf{1}_{s_1\leq t}.$$

We thus see that the function $t \mapsto I_n(g_t)$ is a square-integrable $(\mathcal{F}_t)_{t\in[0,1]}$-martingale. We may then define $I_{n+1}(f)$ to be the Itô integral $(n+1)\int_0^1 I_n(g_t)\,dW(t)$. The set $\mathcal{H}_{n+1}$ spanned by $I_{n+1}(f)$ for all symmetric f in $\mathcal{H}^{\otimes n+1}$ is the $(n+1)$-th Wiener chaos of W.

Remark 2.1 *It holds that $L^2(\Omega)$ is the direct sum – with respect to the inner product defined by expectations of products of r.v.'s – of all the Wiener chaoses. Specifically for any $X \in L^2(\Omega)$, there exists a sequence of non-random symmetric functions $f_n \in \mathcal{H}^{\otimes n} = L^2([0,1]^n)$ with $\sum_{n=0}^{\infty} |f_n|^2_{\mathcal{H}^{\otimes n}} < \infty$ such that $X = \sum_{n=0}^{\infty} I_n(f_n)$; moreover $\mathbf{E}[I_n(f_n) I_m(f_m)] = \delta_{m,n} n! |f_n|^2_{\mathcal{H}^{\otimes n}}$ where $\delta_{m,n}$ equals 0 if $m \neq n$ and 1 if $m = n$.*

Remark 2.2 *(see [16]) The n-th Wiener chaos $\mathcal{H}_n = I_n(\mathcal{H}^{\otimes n})$ coincides with the closed linear subspace of $L^2(\Omega)$ generated by all the random variables of the form $H_n(W(h))$ where $h \in \mathcal{H}$, $|h|_{\mathcal{H}} = 1$, and H_n is the n-th Hermite polynomial, defined by $H_0 \equiv 1, H_1(x) = x$, and $H_{n+1}(x) = (n+1)^{-1}(xH_n(x) - H_{n-1}(x))$. Moreover, $H_n' = H_{n-1}$.*

We believe the easiest way to understand the Malliavin derivative operator is using the following three-step "constructive" presentation; in fact, the essence of the construction of this operator only requires steps 1 and 2(a), as one can arguably see from step 3.

1. We define an operator D from $\mathcal{H}_1$ into $\mathcal{H}$ by the formula

$$D_r W(f) = f(r).$$

   Thus the Malliavin derivative finds the integrand which a centered Gaussian r.v. in $\mathcal{H}_1$ is formed from as a Wiener integral. If $X = W(f) + \mu$ where μ is non-random, $D.X = f$, consistent with the fact that the derivative is linear and kills constants.

2. We extend D by a consistency with the chain rule.

   (a) For any m-dimensional Gaussian vector $G = (G_i)_{i=1}^m \in (\mathcal{H}_1)^m$, for any $\Phi \in C^1(\mathbf{R}^m)$ such that $X = \Phi(G) \in L^2(\Omega)$, in order to be consistent with the appellation "derivative", one must set

$$D_r X = \sum_{i=1}^{m} \frac{\partial \Phi}{\partial g_i}(G)\, D_r G_i = \nabla \Phi(G) \cdot D_r G; \tag{4}$$

   that is to say, the chain rule must hold. It is a simple matter to check that the above requirement (4) can be satisfied for all X of this form, defining D uniquely on them.

   (b) Equivalently, by the chain rule in $C^1(\mathbf{R}^n)$, one can state that formula (4) holds for all Y of the form $Y = \Psi(X_1, \cdots, X_n)$ with $\Psi \in C^1(\mathbf{R}^n)$ and all X_i's as in part 2.a, if we replace $D_r G$ by $D_r X$: $D_r Y = \nabla\Psi(X) \cdot D_r X$ holds for any X, Y and Ψ such that the right hand side is in $L^2(\Omega)$.

3. The following argument can now be used to define D on a much larger set of random variables. For a fixed random variable $Z \in L^2(\Omega)$, we consider the orthogonal chaos decomposition $Z = \sum_{n=0}^{\infty} I_n(f_n)$ of Remark 2.1. From Remark 2.2, $I_n(f_n)$ can be further approximated in $L^2(\Omega)$: $I_n(f_n) = \sum_{j=1}^{\infty} X_j$ where $X_j = H_n(W(h_j))$ where H_n is the nth Hermite polynomial and $h_j \in \mathcal{H}$. By step 2.a, $D_r X_j$ is defined for almost all r, as it is trivial to see that $D_r X_j \in L^2(\Omega)$ for any r such that $h_j(r)$ is finite. More to the point, since $h_j \in \mathcal{H}$, we can say that $D.X_j \in L^2(\Omega) \times \mathcal{H}$. We now need to have a criterion that allows us to justify that $D.I_n(f_n)$ exists in the same space $L^2(\Omega) \times \mathcal{H}$ as a limit in that space of the sums of all the Malliavin derivatives $D.X_j$. It turns out that no additional criterion is needed beyond the fact that the symmetric f_n is in $\mathcal{H}^{\otimes n}$. Indeed, using the relation $H_n' = H_{n-1}$, one proves that the series $\sum_j D_r X_j$ converges to $nI_{n-1}(f_n(\cdot, r))$ in $L^2(\Omega) \times \mathcal{H}$. To complete the program of defining $D.Z$ on as

wide a space of Z's as possible, since from Remark 2.1 we have $\int_0^1 \mathbf{E}\left|nI_{n-1}\left(f_n\left(\cdot,r\right)\right)\right|^2 dr = nn!\left|f_n\right|^2_{\mathcal{H}^{\otimes n}}$, we immediately get that $D.Z$ exists in $L^2(\Omega) \times \mathcal{H}$ and has orthogonal decomposition in that space given by

$$D_r Z = \sum_{n=1}^{\infty} nI_{n-1}\left(f_n\left(\cdot,r\right)\right)$$

as soon as

$$\sum_{n=1}^{\infty} nn!\left|f_n\right|^2_{\mathcal{H}^{\otimes n}} < \infty. \tag{5}$$

Remark 2.3 *The set of all $Z \in L^2(\Omega)$ such that (5) holds is called the (Gross-)Sobolev space $\mathbf{D}^{1,2}$ with respect to W and its Malliavin derivative. It is a Hilbert space with respect to the inner product $\langle Z, Z'\rangle = \mathbf{E}[ZZ'] + \int_0^1 \mathbf{E}[D_r Z D_r Z']\,dr$.*

Remark 2.4 (General Chain Rule for Malliavin derivatives) *Combining relation (4) from Step 2a and Step 3 above, for any $Z \in \left(\mathbf{D}^{1,2}\right)^m$, for any $\Phi \in C^1(\mathbf{R}^m)$ such that $\nabla\Phi(Z) \in L^2(\Omega)$, we get $\Phi(Z) \in \mathbf{D}^{1,2}$ and the general chain rule formula*

$$D_r Z = \nabla\Phi(Z) \cdot D_r Z. \tag{6}$$

3 Sub-Gaussian theory

In this section we develop the concept of sub-Gaussian random variables and processes/fields (a stochastic process defined on an index set that is not a subset of $\mathbf{R}_+$ is normally called a *random field*). We define sufficient Malliavin derivative conditions implying these concepts, and we investigate extensions of the familiar concentration inequalities known as the Dudley-Fernique theorems (on the expected supremum of a process) and the Borel-Sudakov inequalities (on the deviation from this expectation).

Definition 3.1 *A centered random variable X is said to be sub-Gaussian relative to the scale σ if for all $\lambda > 0$,*

$$\mathbf{E}[\exp \lambda X] \leq \exp \lambda^2\sigma^2/2. \tag{7}$$

Remark 3.2 *The interpretation of σ^2 above is that of an upper bound on X's variance. More specifically, the following two statements imply (7) and are implied by it, with different universal constants c in each implication*

$$\mathbf{E}\left[\exp\left(X^2/\left(c\sigma^2\right)\right)\right] \leq 2, \tag{8}$$

and for all $u > 0$,

$$\mathbf{P}\left[|X| > u\right] \leq 2\exp\left(-\frac{u^2}{2c\sigma^2}\right).$$

For instance, (7) implies (8) with $c = 5$. Consult lemma 4.6 for more general results than these implications, and their proofs.

We will use the following fundamental lemma, whose consequences are far-reaching.

Lemma 3.3 *Let X be a centered random variable in $\mathbf{D}^{1,2}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ of the previous section. Assume there exists a non-random constant M such that, $\mathbf{P}$-almost surely,*

$$\int_0^1 |D_r X|^2\,dr \leq M^2. \tag{9}$$

Then X is sub-Gaussian relative to $\sigma^2 = M^2$.

Proof. The following result is due to Üstünel [22, Theorem 9.1.1]: if (9) holds then $\mathbf{P}[|X| > u] \leq 2\exp\left(-u^2/(2M^2)\right)$. The lemma is thus just a translation of this theorem using the definition of sub-Gaussian random variables. ■

In the previous section, we saw that in $(\Omega, \mathcal{F}, \mathbf{P})$ a Gaussian random variable is one such that its Malliavin derivative is non-random. The above lemma states that a class of sub-Gaussian centered random variables is obtained by requiring only that their Malliavin derivatives have an almost-surely bounded norm in $\mathcal{H} = L^2[0,1]$. The reader will check that, equivalently, condition (9) says that $D.X \in L^\infty(\Omega, \mathcal{H})$, and ess $\sup|D.X|^2_{\mathcal{H}}$ is the smallest $M > 0$ satisfying (9) almost surely.

Definition 3.4 *A pseudo-metric is a symmetric function δ on $I \times I$ such that $\delta(s,u) \leq \delta(s,t) + \delta(t,u)$.*

The axiom $\delta(s,t) = 0 \Longrightarrow s = t$ need not hold for pseudo metrics. Examples of pseudo-metrics are the canonical metrics δ_Z of all centered Gaussian fields Z on I: $\delta_Z(s,t) := \sqrt{\mathbf{E}\left[(Z(t) - Z(s))^2\right]}$.

Definition 3.5 *A centered process (random field) X on an arbitrary index set I is said to be sub-Gaussian relative to the pseudo-metric δ on I if for any $s,t \in I$, the random variable $X(t) - X(s)$ is sub-Gaussian relative to the scale $\sigma = \delta(s,t)$.*

Our first theorem is the extension to the class of sub-Gaussian processes defined via condition (9) of the so-called Borell-Sudakov inequality. The classical version of this inequality states that for a centered separable Gaussian field on an index set I, if $\mu := \mathbf{E}\sup_I X < \infty$, then $\mathbf{P}[|\sup_I X - \mu| > u] \leq 2\exp\left(-u^2/\left(2\sigma^2\right)\right)$ where $\sigma^2 = \sup_{t\in I}\mathrm{Var}[X(t)]$.

Theorem 3.6 *Let X be a separable random field on I such that all finite-dimensional vectors of X are formed of almost-surely distinct components. Assume $\mu := \mathbf{E}[\sup_I X] < \infty$. Assume for each $t \in I$, $X(t) \in \mathbf{D}^{1,2}$, and there exist a constant $\sigma^2(t)$ such that almost surely*

$$\int_0^1 |D_r X(t)|^2\, dr \leq \sigma^2(t). \tag{10}$$

Then the random variable $\sup_I X - \mu$ is sub-Gaussian relative to $\sigma^2 = \sup_{t\in I}\sigma^2(t)$. In other words

$$\mathbf{P}\left[\left|\sup_I X - \mu\right| > u\right] \leq 2\exp\left(-\frac{u^2}{2\sigma^2}\right).$$

Proof. *Step 1: setup.*

Separability of X means that its distribution only requires knowledge of X on a countable subset of I, i.e. we can assume I is countable in the expression $\sup_I X$. Hence, by the dominated convergence theorem, the problem reduces to the case of finite I. Thus we assume $I = \{1, 2, \cdots, N\}$ where N is a positive integer and $X = \{X_1, X_2, \cdots, X_N\}$. Now let

$$s_n = \max\{\sigma(1), \sigma(2), \cdots, \sigma(n)\},$$

and

$$S_n = \max\{X_1, X_2, \cdots, X_N\}.$$

Since $\Phi(x,y) = \max(x,y) = x\mathbf{1}_{x\geq y} + y\mathbf{1}_{x<y}$, thus we have $S_{n+1} = \Phi(X_{n+1}, S_n)$ where $\partial\Phi/\partial x(x,y) = \mathbf{1}_{x\geq y}$ and $\partial\Phi/\partial y(x,y) = \mathbf{1}_{x\leq y}$.

Step 2: explicit extension of the chain rule

Unfortunately Φ is not of class C^1, so to keep our proof rigorous, since we will need to use the chain rule formula (6) with Φ, we indicate how to extend it for our purposes. We claim the following.

Lemma 3.7 *The chain rule (6) holds with Z any vector of random variables in $\mathbf{D}^{1,2}$, for any Φ that is of class C^1 off of a finite union T of hyperplanes, with $\nabla\Phi$ bounded, with $\Phi(Z) \in \mathbf{D}^{1,2}$, and with $Z \notin T$ almost surely.*

See the appendix for a proof of this result which is spelled out for the situation we need.

Step 3: induction.

We prove the theorem by induction on n. Our induction hypothesis need only be that $S_n \in \mathbf{D}^{1,2}$ and almost surely,

$$\int_0^1 |D_r S_n|^2 \, dr \leq s_n^2. \tag{11}$$

Indeed, this inequality is satisfied with $n = 1$ by hypothesis since $S_1 = X_1$; when $n = N$, Lemma 3.3 applied to $S_N = \sup_I X$ proves that this induction hypothesis implies the statement of the theorem. Therefore, we only need to prove that if $S_n \in \mathbf{D}^{1,2}$ and (11) holds for some $n \in \{1, \cdots, N-1\}$, then $S_{n+1} \in \mathbf{D}^{1,2}$ and (11) holds for $n+1$. Since $S_{n+1} = \Phi(X_{n+1}, S_n)$, and by hypothesis $X_{n+1} \neq S_n$ almost surely, we can apply the above lemma: for almost every $r \in [0,1]$,

$$\begin{aligned} D_r S_{n+1} &= \frac{\partial \Phi}{\partial x}(X_{n+1}, S_n) D_r X_{n+1} + \frac{\partial \Phi}{\partial y}(X_{n+1}, S_n) D_r S_n \\ &= \mathbf{1}_{X_{n+1} \geq S_n} D_r X_{n+1} + \mathbf{1}_{X_{n+1} \leq S_n} D_r S_n \\ &= \mathbf{1}_{X_{n+1} > S_n} D_r X_{n+1} + \mathbf{1}_{X_{n+1} < S_n} D_r S_n. \end{aligned}$$

The last equality holds a.s. again because the X_i's are distinct almost surely. Therefore, since the product of the two terms in the last line above is zero, using the induction hypothesis (11) and the assumption $\|D_\cdot X_{n+1}\|^2_{L^2[0,1]} \leq \sigma^2(n+1)$, we obtain

$$\begin{aligned} \int_0^1 |D_r S_{n+1}|^2 \, dr &= \mathbf{1}_{X_{n+1} \geq S_n} \int_0^1 |D_r X_{n+1}|^2 \, dr + \mathbf{1}_{X_{n+1} < S_n} \int_0^1 |D_r S_n|^2 \, dr \\ &\leq \sigma^2(n+1) \mathbf{1}_{X_{n+1} \geq S_n} + s_n^2 \mathbf{1}_{X_{n+1} < S_n} \\ &\leq s_{n+1}^2 \mathbf{1}_{X_{n+1} \geq S_n} + s_{n+1}^2 \mathbf{1}_{X_{n+1} < S_n} \\ &= s_{n+1}^2. \end{aligned}$$

By induction, the proof of the theorem is complete. ■

Remark 3.8 *The assumption in the previous theorem that any vector of X's have almost surely distinct components can be easily satisfied using a now classical result on the existence of densities of random vectors. From [16, Theorem 2.1.2] we learn that we only need to check that the matrix of Malliavin derivatives' inner products $\left(\langle D_\cdot X_n; D_\cdot X_{n'}\rangle\right)_{n,n'=1}^N$ is almost surely invertible, since this implies that the law of X has a density. Thus the theorem's two assumptions can be phrased in terms of Malliavin derivatives, one as a boundedness condition, the other as a non-degeneracy condition. The latter is of course much weaker than the former.*

The only assumption on the non-diagonal correlations of X in the above theorem is the finiteness of μ, which has evidently little or nothing to do with the sub-Gaussian property at the process level. The main Malliavin derivative boundedness hypothesis is only a set of one-dimensional distributional hypotheses, which represents a significant improvement over assuming that the entire vector X is jointly Gaussian.

When comparing Lemma 3.3 and Theorem 3.6, one may wonder whether the Borell-Sudakov inequality holds under the weaker hypothesis that each $X(t)$ is sub-Gaussian. This represents a gap which we are not able to fill at this time. It is instructive to note that the main issue here is that the converse of Lemma 3.3 is false: if the r.v. X is sub-Gaussian relative to the scale σ, it does not imply that (9) holds for the scale σ, or even for any other scale. To see this, consider a random variable $X = \int_0^1 u(s) \, dW(s)$ where u is adapted to $(\mathcal{F}_s)_{s\in[0,1]}$ and is such that $\int_0^1 u^2(s)\,ds$ is almost surely bounded by σ^2. For any λ, using the exponential martingale $\mathcal{E}(\lambda M)_t$ based on the martingale $M_t = \int_0^t u(s)\,dW(s)$, we immediately get,

$$\begin{aligned}\mathbf{E}[\exp\lambda X] &= \mathbf{E}\left[\mathcal{E}(\lambda M)_1 \exp\frac{\lambda^2}{2}\int_0^1 u^2(s)\,ds\right] \\ &\leq \exp\left(\frac{\lambda^2\sigma^2}{2}\right)\mathbf{E}\left[\mathcal{E}(\lambda M)_1\right] = \exp\left(\frac{\lambda^2\sigma^2}{2}\right),\end{aligned}$$

which means X is sub-Gaussian relative to the scale σ. But for the specific case of $u(s) = f(W_s)$ where f is bounded and in $C^2(\mathbf{R})$, we can easily find examples of f where (9) does not hold for any scale. Using the formula $D_rX = u(r) + \int_r^1 D_r u(s)\,dW(s)$(see [16]), which in our example yields $D_rX = f(W_r) + \int_r^1 D_r W_s f'(W_s)\,dW_s$, and then using Itô's formula, we get

$$\begin{aligned}\int_0^1 |D_rX|^2\,dr &= \int_0^1 \left|f(W_r) + \int_r^1 f'(W_s)\,dW(s)\right|^2 dr \\ &= \int_0^1 \left|f(W_1) - \frac{1}{2}\int_r^1 f''(W_s)\,ds\right|^2 dr \qquad (12) \\ &= f^2(W_1) - f(W_1)\int_0^1 s f''(W_s)\,ds + \frac{1}{4}\int_0^1\int_0^1 f''(W_s) f''(W_{s'})\min(s,s')\,ds\,ds' \qquad (13)\end{aligned}$$

Even if f is bounded, it is simple to construct examples where f'' is not bounded: e.g. $f(w) = \sin(w^2)$, with $f''(w) = -2\cos(w^2) - 4w^2\sin(w^2)$. Whether in line (12) or line (13), we see that the expression above can take arbitrarily large values with positive probability.

In order to use the Borell-Sudakov inequality efficiently, it is necessary to be able to estimate the expected supremum effectively. We recall here the classical result of Dudley (upper bound) and Fernique (lower bound) for Gaussian processes.

Theorem 3.9 *Let Z be a separable Gaussian field on an index set I. Let $\delta_Z(s,t) = \left(\mathbf{E}\left[(X(s) - X(t))^2\right]\right)^{1/2}$ be its canonical metric. Let the* metric entropy *$N(\varepsilon)$ be the smallest number of balls of radius ε in the pseudo-metric δ_Z needed to cover I. There exist two positive universal constants K and K' such that*

$$\mathbf{E}\left[\sup_I X\right] \leq K\int_0^\infty \sqrt{\log N(\varepsilon)}\,d\varepsilon, \qquad (14)$$

and, if I is a subset of a group G and the law of X, defined on G, is translation invariant (e.g. $I \subset \mathbf{R}^d$ and $\delta(s,t)$ depends only on $|s-t|$, i.e. X is homogeneous *or* stationary*)*

$$\mathbf{E}\left[\sup_I X\right] \geq K'\int_0^\infty \sqrt{\log N(\varepsilon)}\,d\varepsilon.$$

For what classes of processes does a result of the same type as the lower bound above hold? This an open problem which we will not tackle in this paper. Yet the Dudley upper bound (14) of this theorem is true, with the same $N(\varepsilon)$, for all processes which are sub-Gaussian relative to the same pseudo-metric δ_Z. This result even extends beyond the sub-Gaussian case, as we are about to see in the next section, which is why we omit the proof that Theorem 3.9 holds for sub-Gaussian processes. Another reason for omiting the proof is that the result is now classical (see [11]). For the sake of completeness, we still record the statement here.

Remark 3.10 *If X is sub-Gaussian on I, as in Definition 3.5, relative to the pseudo-metric δ, then with the notation of Theorem 3.9, (14) holds.*

4 Sub-Gaussian chaos processes

One of the difficulties with Wiener chaos expansions such as $X = \sum_{n=0}^{\infty} I_n(f_n)$ (defined in Remark 2.1) is that they often mask fundamental properties of processes. In particular, a typical sub-Gaussian random variable has components of all orders in its chaos expansion, so that any estimation done term by term using this expansion will miss the sub-Gaussian property, while the entire sum of the expansion, being sub-Gaussian, is thus more akin to its term of order $n = 1$. In this section we introduce a concept which generalizes this idea to higher values of n. We use it to derive a Dudley-type theorem (Subsection 4.1). Then we attempt to relate the concept to iterated Malliavin derivative calculations (Subsection 4.2), and derive an extension of the Borell-Sudakov concentration inequality as a consequence (Subsection 4.3).

Definition 4.1 *Let n be a positive integer. A centered random variable X is said to have the sub-nth-Gaussian-chaos property (or is a sub-nth chaos r.v., or is a sub-Gaussian chaos r.v. of order n, etc...) relative to the scale M if*

$$\mathbf{E}\left[\exp\left(\left(\frac{X}{M}\right)^{2/n}\right)\right] \leq 2.$$

Obviously, when $n = 1$, such an X is sub-Gaussian relative to the scale $\sqrt{5}M$. Our definition is similar to the definition of an Orlicz norm of X, although the only intersection between the concepts appears to occur for $n = 1$ or 2, since Orlicz norms have a requirement of convexity of their Young function, which is not the case here for $n > 2$ (see [11], or [19]).

Remark 4.2 *From Definition 3.1 and Remark 3.2, we get the following equivalent definitions of the sub-nth-Gaussian chaos property, up to universal multiplicative scale constants c: for all $\lambda, u > 0$*

$$\mathbf{E}\left[\exp\lambda\left(|X|^{1/n} - \mathbf{E}|X|^{1/n}\right)\right] \leq \exp c\lambda^2 M^2/2$$

and

$$\mathbf{P}\left[|X|^{1/n} > u\right] \leq 2\exp\left(-\frac{u^2}{2cM^2}\right).$$

Definition 4.3 *Let δ be a pseudo-metric on a set I. A centered random field X on I is said to be a sub-nth-Gaussian chaos field with respect to δ if for any $s, t \in I$, the random variable $X(t) - X(s)$ has the sub-nth-Gaussian chaos property relative to the scale $\delta(s,t)$.*

Definition 4.4 *Let δ and X be as in the previous definition. We use the notation N_δ, and we say that N_δ is a metric entropy for X, if N_δ is the smallest number of balls of radius ε in the pseudo-metric δ needed to cover I.*

4.1 Expected suprema

As announced in the previous section, we now prove a Dudley upper bound for sub-nth-chaos processes.

Theorem 4.5 *For each fixed positive integer n, there exists a universal constant C_n depending only on n such that if X defined on I is a separable sub-nth-Gaussian chaos field with respect to the pseudo-metric δ, then with N_δ a metric entropy for X,*

$$\mathbf{E}\sup_{t\in I} X(t) \leq C_n \int_0^\infty (\log N_\delta(\varepsilon))^{n/2} d\varepsilon.$$

This theorem is a new result for $n > 2$; it has been established in [23] for $n \leq 2$ using convexity of the Orlicz space's Young function. Our proof of this theorem, which works for any integer $n \geq 1$, requires the first two inequalities of the following lemma, which is established in the Appendix.

Lemma 4.6 *For every integer n, there exists a universal constant v_n such that, for any sub-nth-chaos r.v. X relative to the scale δ, the following inequalities hold: for every $u > 0$,*

$$\mathbf{P}[|X| > u] \leq 2\exp\left(-\left(\frac{u}{\delta}\right)^{2/n}\right);$$

$$\mathbf{E}\left[X^2\right] \leq (v_n\delta)^2,$$

and for every $\lambda > 0$

$$\mathbf{E}\left[\exp\left(\lambda\left|\frac{X}{v_n\delta}\right|^{1/n}\right)\right] \leq v_n e^{\lambda^2/2}.$$

The converse also holds. Namely, with possibly some other universal constant $v_n' > 1$, each of the three inequalities above implies that X is a sub-nth-chaos r.v. relative to the scale $M = v_n'\delta$.

Proof of Theorem 4.5. Our proof is patterned from Michel Ledoux's notes [10] on "Isoperimetry and Gaussian Analysis", although here no Young function convexity is used, and indeed we do not have the restriction $n \leq 2$. We may and do assume that T is finite (see Step 1 of proof of Theorem 3.6). If the right-hand side of the conclusion of the theorem is infinite, there is nothing to prove. Therefore we may assume that $\sup_I X$ is integrable.

Step 1: chaining argument. Let $q > 1$ be fixed and let ℓ_0 be the largest integer ℓ in $\mathbb{Z}$ such that $N_\delta(q^{-\ell}) = 1$. For every $\ell \geq \ell_0$, we consider a family of cardinality $N(\ell) := N_\delta(q^{-\ell})$ of balls of radius $q^{-\ell}$ covering T. One may therefore construct a partition $\mathcal{A}_\ell$ of T of cardinality $N(\ell)$ on the basis of this covering with sets of diameter less than $2q^{-\ell}$. In each A of $\mathcal{A}_\ell$, fix a point of T and denote by T_ℓ the collection of these points. For each t in T, denote by $A_\ell(t)$ the element of $\mathcal{A}_\ell$ that contains t. For every t and every ℓ, let then $s_\ell(t)$ be the element of T_ℓ such that $t \in A_\ell(s_\ell(t))$. Note that $\delta(t, s_\ell(t)) \leq 2q^{-\ell}$ for every t and $\ell \geq \ell_0$. Also note that

$$\delta(s_\ell(t), s_{\ell-1}(t)) \leq 2q^{-\ell} + 2q^{-\ell+1} = 2(q+1)q^{-\ell}.$$

Hence, by the second inequality in the previous lemma, the series $\sum_{\ell>\ell_0}\left(X_{s_\ell(t)} - X_{s_{\ell-1}(t)}\right)$ converges in $L^1(\Omega)$, and also $s_\ell(t)$ converges to t in $L^1(\Omega)$ as $\ell \to \infty$. By the telescoping property of the the above sum, we thus get that almost surely for every t,

$$X_t = X_{s_0} + \sum_{\ell>\ell_0}\left(X_{s_\ell(t)} - X_{s_{\ell-1}(t)}\right) \tag{15}$$

where $s_{\ell_0}(t) := s_0$ may be chosen independent of $t \in T$.

Step 2: Applying the lemma. Let c_ℓ be a constant that will be chosen in the next step. It follows from the decomposition (15) above, and the identity $\mathbf{E}X_{s_0} = 0$, that

$$\begin{aligned}
&\mathbf{E}\left(\sup_{t\in T} X_t\right)\\
&= \mathbf{E}\left[X_{s_0} + \sup_{t\in I}\sum_{\ell>\ell_0}\left(X_{s_\ell(t)} - X_{s_{\ell-1}(t)}\right)\right]\\
&\leq \sum_{\ell>\ell_0} c_\ell + \mathbf{E}\left(\sup_{t\in T}\sum_{\ell>\ell_0}\left|X_{s_\ell(t)} - X_{s_{\ell-1}(t)}\right| 1_{\{|X_{s_\ell(t)} - X_{s_{\ell-1}(t)}|>c_\ell\}}\right)\\
&\leq \sum_{\ell>\ell_0} c_\ell + \mathbf{E}\left(\sum_{\ell>\ell_0}\sum_{(u,v)\in H_\ell}|X_u - X_v| 1_{\{|X_u - X_v|>c_\ell\}}\right)\\
&\leq \sum_{\ell>\ell_0} c_\ell + \sum_{\ell>\ell_0}\sum_{(u,v)\in H_\ell}\mathbf{E}\left(|X_u - X_v| 1_{\{|X_u - X_v|>c_\ell\}}\right)
\end{aligned}$$

where $H_\ell = \{(u,v) \in T_\ell \times T_{\ell-1}; \delta(u,v) \leq 2(q+1)q^{-\ell}\}$. Using Holder's inequality, we get

$$\mathbf{E}\left(\sup_{t\in T} X_t\right) \leq \sum_{\ell>\ell_0} c_\ell + \sum_{\ell>\ell_0} \sum_{(u,v)\in H_\ell} \left(\mathbf{E}|X_u - X_v|^2\right)^{1/2} \left(\mathbf{P}(|X_u - X_v| > c_\ell)\right)^{1/2}.$$

Using Lemma 4.6 now, and applying a uniform upper bound for all $(u,v) \in H_\ell$, we get

$$\begin{aligned}\mathbf{E}\left(\sup_{t\in T} X_t\right) &\leq \sum_{\ell>\ell_0} c_\ell + \sum_{\ell>\ell_0} \sum_{(u,v)\in H_\ell} v_n \delta(u,v) \left(2\exp\left(-\left(\frac{c_\ell}{\delta(u,v)}\right)^{2/n}\right)\right)^{1/2} \\ &\leq \sum_{\ell>\ell_0} c_\ell + \sum_{\ell>\ell_0} v_n \mathrm{Card}(H_\ell) 2(q+1)q^{-\ell} \left(2\exp\left(-\left(\frac{c_\ell}{2(q+1)q^{-\ell}}\right)^{2/n}\right)\right)^{1/2}\end{aligned}$$

Step 3: Choosing c_ℓ. Since $\mathrm{Card}(H_\ell) \leq N(\ell)^2$, it is now apparent that a convenient choice for c_ℓ, in order to exploit the summability of $q^{-\ell}$ without having to worry about the size of $\mathrm{Card}(H_\ell)$, is $c_\ell = 2(q+1)q^{-\ell}(4\log N(\ell))^{n/2}$. We thus obtain

$$\begin{aligned}\mathbf{E}\left(\sup_{t\in T} X_t\right) &\leq \sum_{\ell>\ell_0} c_\ell + \sum_{\ell>\ell_0} N(\ell)^2\, 2(q+1)q^{-\ell} v_n \exp(-2\log N(\ell)) \\ &\leq \sum_{\ell>\ell_0} 2(q+1)q^{-\ell}(4\log N(\ell))^{n/2} + \sum_{\ell>\ell_0} 2^{3/2}(q+1)q^{-\ell} v_n.\end{aligned}$$

Step 4: Conclusion. Now, since for $\ell > \ell_0$, $\log N(\ell) \geq \log 2$, then $(\log N(\ell))^{n/2} \geq (\log 2)^{n/2}$ for $n \geq 1$. It follows that

$$\mathbf{E}\left(\sup_{t\in T} X_t\right) \leq k_n(q+1) \sum_{\ell>\ell_0} q^{-\ell}(\log N(\ell))^{n/2}$$

where $k_n = 2\cdot 4^{n/2} + 2^{3/2} v_n \log^{-n/2} 2$. By comparing our series to an integral, since N_δ is decreasing, we get

$$\begin{aligned}\mathbf{E}\left(\sup_{t\in T} X_t\right) &\leq k_n \left(\frac{q+1}{1-q^{-1}}\right)(1-q^{-1}) \sum_{\ell>\ell_0} q^{-\ell}(\log N(\ell))^{n/2} \\ &\leq k_n \left(\frac{q+1}{1-q^{-1}}\right) \sum_{\ell>\ell_0} \int_{q^{-\ell-1}}^{q^{-\ell}} (\log N_\delta(\varepsilon))^{n/2} d\varepsilon \\ &\leq C_n \int_0^\infty (\log N_\delta(\varepsilon))^{n/2} d\varepsilon\end{aligned}$$

where $C_n = k_n \left(\frac{q(q+1)}{q-1}\right)$. The theorem is proved with $C_n = \left(2\sqrt{2}+3\right) k_n$. ∎

4.2 Malliavin derivative conditions

A connection between the above definition of sub-nth-chaos r.v.'s and Malliavin derivatives is provided by the following.

Theorem 4.7 *Let X be a random variable in $\mathbf{D}^{n,2}$. That is to say, X has n iterated Malliavin derivatives, and the nth derivative $D^{(n)}_{s_n,\dots,s_2,s_1}X = D_{s_n}\left(D_{s_{n-1}}\left(\cdots D_{s_2}\left(D_{s_1}X\right)\cdots\right)\right)$ is a member of $L^2\left(\Omega \times \mathcal{H}^{\otimes n}\right)$. With the notation $X = \sum_{m=0}^{n-1} I_m(f_m) + X_n$ where each f_m is a non-random symmetric function in $\mathcal{H}^{\otimes m}$, X is a sub-nth-Gaussian chaos random variable in the following two cases.*

Case $m = 2$ *Assume*

$$M_2 := \underset{\omega\in\Omega}{ess\ sup}\left(\int_0^1\int_0^{s_1}\left|D^{(2)}_{s_2,s_1}X\right|^2 ds_2 ds_1\right)^{1/2} < \infty.$$

Then $X - \mathbf{E}X$ *is a sub-2nd-Gaussian chaos random variable relative to the scale* $\pi\sqrt{10}M_2$.

Proposition 4.8 Case $m \geq 3$ *Let*

$$G_n(x) = \sum_{k=n+1}^{\infty}\left(\frac{2k}{n}\sqrt{\frac{e}{2}}\right)^{2k/n}\frac{1}{k!}x^k. \tag{16}$$

Assume that almost surely,

$$M_2 := \underset{\omega\in\Omega}{ess\ sup}\left(\int_0^1\int_0^{s_1}\cdots\int_0^{s_{n-1}}\left|\mathbf{E}\left[D^{(n)}_{s_n,\cdots,s_2,s_1}X|\mathcal{F}_{s_n}\right]\right|^2 ds_n\cdots ds_2 ds_1\right)^{1/2} < \infty \tag{17}$$

and assume there exists M_G *non random such that almost surely*

$$\int_0^1\int_0^{s_1}\cdots\int_0^{s_{n-1}} G_n\left(|M_G|^{-2}\left|\mathbf{E}\left[D^{(n)}_{s_n,\cdots,s_2,s_1}X|\mathcal{F}_{s_n}\right]\right|^{2/n}\right) ds_n\cdots ds_2 ds_1 \leq 1/2. \tag{18}$$

Then X_n *is a sub-nth-Gaussian chaos random variable relative to any scale* $M \geq \max\left(\log^{-n/2}(3/2)\,M_2; M_G\right)$.

In particular, with K_u *a universal constant, the following choice of* M *is satisfactory if it is finite:*

$$M = \underset{\omega\in\Omega}{ess\ sup}\frac{\left(\log\int_0^1\int_0^{s_1}\cdots\int_0^{s_{n-1}} n!\exp\left(\left|\mathbf{E}\left[D^{(n)}_{s_n,\cdots,s_2,s_1}X|\mathcal{F}_{s_n}\right]\right|^{6/n}\right)ds_n\cdots ds_2 ds_1\right)^{n/6}}{\log^{n/6}\left(1+\frac{n!}{K_u}\right)}.$$

Note that f may be taken to be symmetric in the above theorem. Also note that this theorem is presumably inefficient for $n \geq 3$, since the case $n = 2$ has a much more natural conclusion. In fact one may conjecture that up to a universal constant, Condition (17) by itself is sufficient to ensure that X_n is a sub-nth-Gaussian chaos random variable relative to M_2; yet we have not found a proof of this fact in general. Our result in the above theorem in the case $n = 2$ matches this conjecture in that case, up to the multiplicative universal constant $\pi\sqrt{10}$ which is presumably not sharp; the proof is self-contained, and of independent interest, but does not seem to allow passage to $m \geq 3$; the proof is also intriguing in that it seems to make rather wasteful use of the hypothesis of boundedness of $\left|D^{(2)}X\right|_{\mathcal{H}^{\otimes 2}}$, and one may wonder whether examples can be found where X is a sub-2*nd*-chaos r.v. without $\left|D^{(2)}X\right|_{\mathcal{H}^{\otimes 2}}$ being bounded. The conjecture does holds for the special case of n-th Wiener chaos random variables, i.e. $X = I_n(f_n)$ for some non-random $f \in \mathcal{H}^{\otimes n}$; we have not found an elementary proof of this fact; nevertheless it is a consequence of the proof of a result by C. Borell in [3], where the isoperimetric inequality is used (see Lemma 4.16 below). Lastly, note that the proof of Theorem 4.7 for $n = 2$ does not seem to extend to $n \geq 3$, while it is not possible to adapt the proof for $n \geq 3$ to the case $n = 2$ because, in the latter case, the function G_2 would not have an infinite radius of convergence.

The classical Clark-Ocone representation will be needed to prove Theorem 4.7.

Remark 4.9 (Clark Ocone Representation) *Any random variable* X *in* $\mathbf{D}^{1,2}$ *can be written as* $X = \mathbf{E}X + \int_0^1\mathbf{E}\left[D_sX|\mathcal{F}_s\right]dW_s$.

By iterating this proposition, we obtain the following, whose proof is in the appendix.

Lemma 4.10 *Let* $X \in \mathbf{D}^{n,2} \subset L^2(\Omega)$ *with the Wiener chaos decomposition* $X = \sum_{n=0}^{\infty} I_n(f_n)$ *where* $f_n \in \mathcal{H}^{\otimes n}$ *and* f_n *is symmetric. Then*

$$X = \sum_{m=0}^{n-1} I_m(f_m) + \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} \mathbf{E}(D^{(n)}_{s_n,\cdots,s_2,s_1} X|\mathcal{F}_{s_n}) dW_{s_n} \cdots dW_{s_2} dW_{s_1}.$$

Proof of Theorem 4.7, "Case $n \geq 3$". From Lemma 4.10, where the functions $(f_m)_{m=0}^{n-1}$ are identified, we have that

$$X_n = X - \sum_{m=0}^{n-1} I_m(f_m) = \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} u(s_1, s_2, \cdots, s_n)\, dW_{s_n} \cdots dW_{s_2} dW_{s_1}$$

where the stochastic process $u(s_1, s_2, \cdots, \cdot)$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$, and $u \in L^\infty(\Omega; L^2(\mathcal{H}^{\otimes n}))$, that is to say, with the non-random number $\|u\|_{\infty,2} := (1/n!)\,\|u\|_{L^\infty(\Omega;L^2(\mathcal{H}^{\otimes n}))}$, almost surely,

$$\int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} |u(s_1, s_2, \cdots, s_n)|^2\, ds_n \cdots ds_2 ds_1 \leq \|u\|^2_{\infty,2}. \tag{19}$$

Let now

$$U = \left| \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} u(s_1, s_2, \cdots, s_n)\, dW_{s_n} \cdots dW_{s_2} dW_{s_1} \right|^{1/n}. \tag{20}$$

Intuitively, since one way to construct an nth Wiener chaos r.v. is to take a polynomial of degree n and apply it to a Gaussian r.v., the definition of U should presumably give us a sub-Gaussian r.v. In any event, to prove the theorem, we only need to show that

$$\mathbf{E}\exp\left(\frac{U^2}{L^2}\right) \leq 2,$$

where $L = M^{1/n}$.

Step 1: Taylor expansion. For simplicity we use the notation $V = U/L$. We simply evaluate the Taylor expansion of the exponential above in the following way, where for the terms with $k = 1, \cdots, n$, we used Jensen's inequality:

$$\begin{aligned}
\mathbf{E}\exp V^2 = \mathbf{E}\sum_{k=0}^{\infty} \frac{V^{2k}}{k!} &= \mathbf{E}\left[1 + \sum_{k=1}^{n} \frac{V^{2k}}{k!} + \sum_{k=n+1}^{\infty} \frac{V^{2k}}{k!}\right] \\
&\leq 1 + \sum_{k=1}^{n} \frac{\mathbf{E}\left[V^{2n}\right]^{k/n}}{k!} + \sum_{k=n+1}^{\infty} \frac{\mathbf{E}\left[V^{2k}\right]}{k!} \\
&\leq 1 + \sum_{k=1}^{n} \frac{\mathbf{E}\left[V^{2n}\right]^{k/n}}{k!} + \sum_{k=n+1}^{\infty} \frac{\mathbf{E}\left[V^{2k}\right]}{k!} \\
&\leq 1 + \sum_{k=1}^{n} \frac{\mathbf{E}\left[V^{2n}\right]^{k/n}}{k!} + \sum_{k=n+1}^{\infty} \frac{\mathbf{E}\left[V^{2k}\right]}{k!}.
\end{aligned} \tag{21}$$

Step 2: Moments evaluations. We have that $V^{2k} = (V^n)^{2k/n} = \left(\left|Y^{(n)}(1)\right| / L^n\right)^{2k/n}$ where $\left(Y^{(n)}(t)\right)_{t\in[0,1]}$ is the $(\mathcal{F}_t)_{t\in[01]}$-martingale defined by

$$Y^{(n)}(t) = \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} u(s_1, s_2, \cdots, s_n)\, dW_{s_n} \cdots dW_{s_2} dW_{s_1}.$$

The bracket of $Y^{(n)}$ thus satisfies

$$\left\langle Y^{(n)}\right\rangle(t)=\int_0^t ds_1\left|\int_0^{s_1}\cdots\int_0^{s_{n-1}}u(s_1,s_2,\cdots,s_n)\,dW_{s_n}\cdots dW_{s_2}\right|^2.$$

We begin by evaluating the moments in the tail of the series (21). By the Burkholder-Davis-Gundy inequality, for any $k\geq n+1$, since then we have $p:=2k/n>2$,

$$\mathbf{E}\left[V^{2k}\right]\leq c(2k/n)\,\mathbf{E}\left[\left\langle Y^{(n)}\right\rangle^{k/n}(1)\right]/L^{2k}.$$

where $c(2k/n)$ is the efficient constant defined in Proposition 5.1 in the Appendix. Let us evaluate the moments of this bracket by induction. We begin by defining a sequence of martingales: for $t\leq s_1$,

$$Y_{s_1}^{(n-1)}(t)=\int_0^t\cdots\int_0^{s_{n-1}}u(s_1,s_2,\cdots,s_n)\,dW_{s_n}\cdots dW_{s_2};$$

for $t\leq s_2\leq s_1$,

$$Y_{s_1,s_2}^{(n-2)}(t)=\int_0^t\cdots\int_0^{s_{n-1}}u(s_1,s_2,\cdots,s_n)\,dW_{s_n}\cdots dW_{s_3};$$

more generally for $t\leq s_j\leq s_{j-1}\leq\cdots\leq s_1$,

$$Y_{s_1,s_2,\cdots,s_j}^{(n-j)}(t)=\int_0^t\cdots\int_0^{s_{n-1}}u(s_1,s_2,\cdots,s_n)\,dW_{s_n}\cdots dW_{s_{j+1}};$$

the last iteration is for $t\leq s_{n-1}\leq s_{n-2}\leq\cdots\leq s_1$,

$$Y_{s_1,s_2,\cdots,s_{n-1}}^{(1)}(t)=\int_0^t u(s_1,s_2,\cdots,s_n)\,dW_{s_n}.$$

We now have, iterating the use of the Burkholder-Davis-Gundy inequality, and using Jensen's inequality for the measures ds_{j+1}/s_j on $[0,s_j]$ for each $j=0,\cdots,n-1$,

$$\begin{aligned}
&c(p)\,\mathbf{E}\left[\left\langle Y^{(n)}\right\rangle^{p/2}(1)\right]\\
&=c(p)\,\mathbf{E}\left[\left(\int_0^1 ds_1\left|Y_{s_1}^{(n-1)}(s_1)\right|^2\right)^{p/2}\right]\\
&\leq c(p)\int_0^1 ds_1\mathbf{E}\left[\left|Y_{s_1}^{(n-1)}(s_1)\right|^p\right]\\
&\leq c(p)^2\int_0^1 ds_1\mathbf{E}\left[\left(\int_0^{s_1}ds_2\left|Y_{s_1,s_2}^{(n-2)}(s_2)\right|^2\right)^{p/2}\right]\\
&\leq c(p)^2\int_0^1 ds_1 s_1^{p/2-1}\mathbf{E}\left[\int_0^{s_1}ds_2\left|Y_{s_1,s_2}^{(n-2)}(s_2)\right|^p\right]\\
&\vdots\\
&\leq c(p)^{n-1}\int_0^1 ds_1 s_1^{p/2-1}\int_0^{s_1}ds_2 s_2^{p/2-1}\cdots\int_0^{s_{n-3}}ds_{n-2}s_{n-2}^{p/2-1}\int_0^{s_{n-2}}ds_{n-1}\mathbf{E}\left[\left|Y_{s_1,s_2,\cdots,s_{n-1}}^{(1)}(s_{n-1})\right|^p\right]\\
&=c(p)^{n-1}\int_0^1 ds_1 s_1^{p/2-1}\int_0^{s_1}ds_2 s_2^{p/2-1}\cdots\\
&\cdots\int_0^{s_{n-3}}ds_{n-2}s_{n-2}^{p/2-1}\int_0^{s_{n-2}}ds_{n-1}\mathbf{E}\left[\left(\int_0^{s_{n-1}}u^2(s_1,s_2,\cdots,s_{n-1},s_n)\,ds_n\right)^{p/2}\right].
\end{aligned}\tag{22}$$

Now for the first terms in the series (21), note that an immediate calculation from (20) (as in Step 2, with $p=2$, so that $c(2)=1$) yields

$$\begin{aligned}\mathbf{E}\left[U^{2n}\right] &\leq \operatorname*{ess\,sup}_{\omega\in\Omega} \int\cdots\int_{0\leq s_n\leq\cdots\leq s_1\leq 1} \left|u(\bar{s})\right|^2 ds_1\cdots ds_n \\ &= \|u\|_{\infty,2}^2 .\end{aligned}$$

Thus we have

$$\begin{aligned}\sum_{k=1}^{n} \frac{\mathbf{E}\left[V^{2n}\right]^{k/n}}{k!} &\leq \sum_{k=1}^{n} \frac{1}{k!}\left(\frac{\mathbf{E}\left[U^{2n}\right]}{L^{2n}}\right)^{k/n} \\ &\leq \sum_{k=1}^{n} \frac{1}{k!}\left(\frac{\|u\|_{\infty,2}^2}{L^{2n}}\right)^{k/n} .\end{aligned} \tag{23}$$

Step 3. General conclusion. We first deal with the terms in (23). Let us show that we can find a constant $c(n)$ depending only on n such that if $L^{2n}\geq c(n)\|u\|_{\infty,2}^2$, then the term in (23) is bounded by $1/2$. Indeed, since

$$\begin{aligned}\sum_{k=1}^{n} \frac{1}{k!}\left(\frac{\|u\|_{\infty,2}^2}{L^{2n}}\right)^{k/n} &\leq \sum_{k=1}^{n} \frac{1}{k!} c(n)^{-k/n} \\ &= \sum_{k=1}^{n} \frac{1}{k!}\left[c(n)^{-1/n}\right]^k \\ &\leq e^{c(n)^{-1/n}}-1,\end{aligned}$$

it is sufficient to have $e^{c(n)^{-1/n}}=3/2$, i.e. $c(n)=\log^{-n}(3/2)$, or in other words

$$L^{2n}\geq \log^{-n}(3/2)\,\|u\|_{\infty,2}^2 . \tag{24}$$

Under this constraint, with inequality (21), we thus get

$$\mathbf{E}\left[\exp V^2\right]\leq 1+1/2+T \tag{25}$$

where the tail term T of the Taylor expansion, is dealt with as follows. We apply line (22) above with $p=2k/n$ and then sum over all $k\geq n+1$. Thus one last use of Jensen's inequality, and the upper bound (19) on $|u|$, with the shorthand notation $u(\bar{s}):=u(s_1,s_2,\cdots,s_{n-1},s_n)$, yield

$$\begin{aligned}T:=\sum_{k=n+1}^{\infty}\frac{\mathbf{E}\left[V^{2k}\right]}{k!} &\leq \sum_{k=n+1}^{\infty}\frac{c(2k/n)^{n-1}}{k!L^{2k}}\mathbf{E}\left[\left\langle Y^{(n)}\right\rangle^{k/n}(1)\right] \\ &\leq \mathbf{E}\sum_{k=n+1}^{\infty}\frac{c(2k/n)^{n-1}}{k!L^{2k}}\int_0^1 ds_1 s_1^{k/n-1}\int_0^{s_1} ds_2 s_2^{k/n-1}\cdots \\ &\cdots\int_0^{s_{n-3}} ds_{n-2}s_{n-2}^{k/n-1}\int_0^{s_{n-2}} ds_{n-1}\left(\int_0^{s_{n-1}} u^2(\bar{s})\,ds_n\right)^{k/n} \\ &\leq \mathbf{E}\sum_{k=n+1}^{\infty}\frac{c(2k/n)^{n}}{k!L^{2k}}\int\cdots\int_{0\leq s_n\leq\cdots\leq s_1\leq 1} ds_1\cdots ds_n\left|u(\bar{s})\right|^{2k/n} \\ &\leq \operatorname*{ess\,sup}_{\omega\in\Omega}\sum_{k=n+1}^{\infty}\frac{c(2k/n)^{n}}{k!L^{2k}}\int\cdots\int_{0\leq s_n\leq\cdots\leq s_1\leq 1} ds_1\cdots ds_n\left|u(\bar{s})\right|^{2k/n} .\end{aligned} \tag{26}$$

Below for any $m \geq 2$, $\|f\|_m$ denotes the L^m norm of any function f on the simplex $0 \leq s_n \leq \cdots \leq s_1 \leq 1$ with respect to Lebesgue measure. Hence we can write from (25) and from (26):

$$\begin{aligned}
&\mathbf{E} \exp V^2 \\
&\leq 3/2 + \text{ess sup}_{\omega\in\Omega} \sum_{k=n+1}^{\infty} \frac{c\,(2k/n)^n}{k! L^{2k}} \|u\|_{2k/n}^{2k/n} \\
&= 3/2 + \operatorname*{ess\,sup}_{\omega\in\Omega} \int \cdots \int_{0\leq s_n\leq\cdots\leq s_1\leq 1} ds_1 \cdots ds_n \sum_{k=n+1}^{\infty} \frac{c\,(2k/n)^n}{k!} \left(\frac{|u\,(s_1, s_2, \cdots, s_{n-1}, s_n)|^{2/n}}{L^2} \right)^k .
\end{aligned}$$

Now the estimate in Proposition 5.1 in the appendix tells us that

$$\frac{c\,(2k/n)^n}{k!} \leq \left(\frac{2k}{n} \sqrt{\frac{e}{2}} \right)^{2k/n} \frac{1}{k!},$$

so that, with $G_n\,(x)$ as in the statement of the theorem,

$$\mathbf{E} \exp V^2 \leq 3/2 + \operatorname*{ess\,sup}_{\omega\in\Omega} \int \cdots \int_{0\leq s_n\leq\cdots\leq s_1\leq 1} ds_1 \cdots ds_n G_n \left(\frac{|u\,(\bar{s})|^{2/n}}{L^2} \right).$$

Choosing L^2 such that the last term above is less than $1/2$, and with the constraint (24), the statement following line (18) in the theorem now follows immediately.

Step 4. Analytic conclusion. To finish the proof of the theorem in the case $n \geq 3$, we only need to study the function G_n more specifically. Using the Stirling-type formula which is valid for all $k \geq 1$, $k! \geq k^k 3^{-k}$, and using the fact that $n \geq 3$, we get easily

$$G_n\,(x) \leq \sum_{k=n+1}^{\infty} \left(\frac{4}{k} \right)^{k/3} x^k.$$

For any integer $m \geq 2$, consider the three values $k = 3m$, $3m+1$, or $3m+2$. We then obtain $k^{k/3} \geq (3m)^m$. On the other hand, for these same values of k, with $x > 1$, we get $x^k \leq \left(x^3\right)^m x^2$. Thus

$$G_n\,(x) \leq 4^{2/3} x^2 \sum_{m=1}^{\infty} \left(\frac{4}{3} \right)^m m^{-m} \left(x^3\right)^m.$$

Using again the Stirling-type formula, valid for all $m \geq 1$, $2^m m^{-m} \leq 1/m!$, we get

$$\begin{aligned}
G_n\,(x) &\leq 4^{2/3} \sum_{m=1}^{\infty} \frac{1}{m!} \left(\frac{2}{3} x^3 \right)^m \\
&= 4^{2/3} x^2 \left(\exp\left(\frac{2}{3} x^3 \right) - 1 \right).
\end{aligned}$$

Thus for $x > 1$,

$$G_n\,(x) \leq 9 \cdot 4^{2/3} \left(\exp\left(x^3\right) - 1 \right),$$

even though the universal constant $9 \cdot 4^{2/3}$ may not be optimal. When $0 < x < 1$, on the other hand, a similar inequality is found, with a different universal constant; we use the notation K_u for the maximum of the two constants. We may now rewrite the left-hand side of (18), which we call Γ, using the last inequality above:

$$
\begin{aligned}
n!\Gamma &:= \int_{[0,1]^n} G_n\left(L^{-2}\left|u\left(\bar{s}\right)\right|^{2/n}\right) ds_n \cdots ds_2 ds_1 \\
&\leq -K_u + K_u \int_{[0,1]^n} \exp\left(L^{-6}\left|u\left(\bar{s}\right)\right|^{6/n}\right) ds_n \cdots ds_2 ds_1 \\
&= -K_u + K_u \int_{[0,1]^n} \left[\exp\left(\left|u\left(\bar{s}\right)\right|^{6/n}\right)\right]^{1/L^6} ds_n \cdots ds_2 ds_1 .
\end{aligned}
$$

We now make a temporary assumption that $L \geq 1$. This allows us to use Jensen's inequality in the above time integral over the simplex:

$$
n!\Gamma \leq -K_u + K_u \left(\int_{[0,1]^n} \exp\left(\left|u\left(\bar{s}\right)\right|^{6/n}\right) ds_n \cdots ds_2 ds_1\right)^{1/L^6} .
$$

Hence, since we only need to satisfy the condition (18), i.e. $\Gamma \leq 1/2$ almost surely, we only need to have

$$
L^6 \geq \frac{\log\left(\int_{[0,1]^n} \exp\left(\left|u\left(\bar{s}\right)\right|^{6/n}\right) ds_n \cdots ds_2 ds_1\right)}{\log\left(1+\frac{n!}{2K_u}\right)} \tag{27}
$$

almost surely. Jensen's inequality can then be used to check that this last expression is always larger than the right-hand side of 17. The last statement of the theorem is thus proved if the essential supremum $(L^*)^6$ of the right-hand side of (27) happens to be greater than 1. If it is not, we leave it to the reader to check that the same conclusion holds by repeating the above calculation (Steps 3 and 4) for the random variable $\tilde{U} = U/L^*$, thereby allowing us not to require $L \geq 1$. ■

Proof of Theorem 4.7, "Case $n = 2$". The proof is based on Lemma 3.3, applied to the random variable

$$
Y = \left(\int_0^1 \left|D_r X\right|^2 ds\right)^{1/2} = \left|D_{\cdot} X\right|_{\mathcal{H}} .
$$

The first step is to prove the following: almost surely,

$$
\left|D_{\cdot} Y\right|_{\mathcal{H}}^2 \leq M_2^2 = \operatorname*{ess\ sup}_{\omega\in\Omega} \left|D_{\cdot}^{(2)} X\right|_{\mathcal{H}^{\otimes 2}}^2 = \operatorname*{ess\ sup}_{\omega\in\Omega} \left(\int_0^1 \int_0^{s_1} \left|D_{s_2,s_1}^{(2)} X\right|^2 ds_2 ds_1\right)^{1/2} .
$$

Indeed we have

$$
\begin{aligned}
\left|D_{\cdot} Y\right|_{\mathcal{H}}^2 &= \int_0^1 \left|D_t \sqrt{\int_0^1 \left|D_r X\right|^2 dr}\right|^2 dt \\
&= \int_0^1 \left|\frac{D_t \int_0^1 \left|D_r X\right|^2 dr}{2\sqrt{\int_0^1 \left|D_r X\right|^2 dr}}\right|^2 dt \\
&= \int_0^1 \frac{\left|\int_0^1 \left(D_r X\right)\left(D_{t,r}^{(2)} X\right) dr\right|^2}{\int_0^1 \left|D_r X\right|^2 dr} dt \\
&\leq \int_0^1 \frac{\int_0^1 \left|D_r X\right|^2 dr \cdot \int_0^1 \left|D_{t,r}^{(2)} X\right|^2 dr}{\int_0^1 \left|D_r X\right|^2 dr} dt \\
&= \int_0^1 \int_0^1 \left|D_{t,r}^{(2)} X\right|^2 dr dt \leq M_2^2 .
\end{aligned}
$$

Thus we can consider that $Z = Y - \mathbf{E}Y$ is a random variable satisfying the hypotheses of Lemma 3.3. We can thus conclude that Z is sub-Gaussian relative to the scale M_2. In particular we get, from Remark 3.2,

$$\mathbf{E}\left[\exp\left(\frac{Z^2}{5M_2^2}\right)\right] \leq 2.$$

Because we will need to find a smaller constant than 2 above, we restate this as

$$\mathbf{E}\left[\exp\left(\frac{Z^2}{10M_2^2}\right)\right] \leq \sqrt{2}. \tag{28}$$

We now invoke an exponential Poincaré inequality of Üstünel [22, Theorem 9.2.3(i)]: for any random centered variable V in $\mathbf{D}^{1,2}$,

$$\mathbf{E}\left[\exp V\right] \leq \mathbf{E}\left[\exp\left(\frac{\pi^2}{8}\left|D.V\right|_{\mathcal{H}}^2\right)\right].$$

Applying this to $V = Z/c$ for some constant $c > 0$, we get

$$\begin{aligned}
\mathbf{E}\left[\exp\left(\frac{X}{c}\right)\right] &\leq \mathbf{E}\left[\exp\left(\frac{\pi^2}{8c^2}\left|D.X\right|_{\mathcal{H}}^2\right)\right] \\
&= \mathbf{E}\left[\exp\left(\frac{\pi^2}{8c^2}\left(Z + \mathbf{E}\left|D.X\right|_{\mathcal{H}}\right)^2\right)\right] \\
&\leq \mathbf{E}\left[\exp\left(\frac{\pi^2}{4c^2}Z^2\right)\right]\exp\left(\frac{\pi^2}{4c^2}\mathbf{E}\left[\left|D.X\right|_{\mathcal{H}}^2\right]\right).
\end{aligned} \tag{29}$$

Now if we choose $\pi^2/\left(4c^2\right) = 1/\left(10M_2^2\right)$, from (28), the first term in the last line above is bounded above by $\sqrt{2}$. In order to control the second term, we use the Clark-Ocone representation to get

$$\begin{aligned}
\mathbf{E}\left[\left|D.X\right|_{\mathcal{H}}^2\right] &= \mathbf{E}\left[\left|\int_0^1\int_0^s \mathbf{E}\left[D_{r,s}^{(2)}X|\mathcal{F}_r\right]dW_r dW_s\right|^2\right] \\
&= \mathbf{E}\left[\int_0^1\int_0^s \left|\mathbf{E}\left[D_{r,s}^{(2)}X|\mathcal{F}_r\right]\right|^2 drds\right] \\
&\leq \mathbf{E}\int_0^1\int_0^s \left|D_{r,s}^{(2)}X\right|^2 drds \\
&\leq M^2.
\end{aligned}$$

Certainly, the above choice for c implies $\pi^2/\left(4c^2\right) \leq \left(\log\sqrt{2}\right)/M_2^2$. From (29) we now get

$$\mathbf{E}\left[\exp\left(\frac{X}{\pi\sqrt{5/2M}}\right)\right] \leq 2. \tag{30}$$

The last step in the proof is to allow the use of $|X|$ instead of X above. Since we have no information about the symmetry of X, we proceed as follows. Since X and $-X$ satisfy the same hypotheses, we have that (30) holds for X replaced by $-X$. Now we can write, with $X' = X/\left(\pi\sqrt{10}M\right)$, and using the notation $p = \mathbf{P}\left[X' \geq 0\right]$

$$\begin{aligned}
\mathbf{E}\left[\exp\left(|X'|\right)\right] &= \mathbf{E}\left[\exp\left(X'\right)\mathbf{1}_{X'\geq 0}\right] + \mathbf{E}\left[\exp\left(-X'\right)\mathbf{1}_{X'<0}\right] \\
&\leq \sqrt{p}\sqrt{\mathbf{E}\left[\exp\left(2X'\right)\right]} + \sqrt{1-p}\sqrt{\mathbf{E}\left[\exp\left(-2X'\right)\right]} \\
&\leq \sqrt{2}\left(\sqrt{p} + \sqrt{1-p}\right) \\
&\leq 2.
\end{aligned}$$

This finishes the proof of Case $n = 2$ of the theorem. ∎

4.3 Concentration: the sub-nth chaos property for suprema

We now prove the core of a Borell-Sudakov-type inequality for sub-nth chaos random fields.

Proposition 4.11 *Let X be a separable random field on an index set I such that all finite-dimensional vectors of X are formed of almost-surely distinct components. Assume $\mu := \mathbf{E}[\sup_I X] < \infty$. Assume $X(t) \in \mathbf{D}^{n,2}$ for each $t \in I$. Assume there exist non-random constants $\sigma(t)$ for each $t \in I$ such that almost surely*

$$\left\| D^{(n)}_{\cdot} X(t) \right\|_2^2 := \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} \left| D^{(n)}_{s_n,\cdots,s_2,s_1} X(t) \right|^2 ds_n \cdots ds_2 ds_1 \leq \sigma^2(t).$$

Then $\sup_{t \in I} X(t) \in \mathbf{D}^{1,2}$ and

$$\left\| D^{(n)}_{\cdot} \sup_{t \in I} X(t) \right\|_2^2 \leq \sup_{t \in I} \sigma^2(t).$$

Proof. As in the proof of Theorem 3.6, we can assume without loss of generality that $I = \{1, 2, \cdots N\}$. Here we have $n \geq 2$. Using the same strategy as in the proof of Theorem 3.6, we denote $X_m = X(m)$ and define $S_m = \max\{X_1, X_2, \cdots, X_m\}$, so that $S_{m+1} = \max\{X_m, S_m\}$. In order to prove that $\max_I X \in \mathbf{D}^{n,2}$, the approximation technique used in the proof of Theorem 3.6 can again be used. We omit the details, only to say that $\mathbf{1}_{X_{m+1} > Sm}$ can be approximated in $\mathbf{D}^{1,2}$ by a smooth function of $X_{m+1} - S$ whose Malliavin derivative tends to 0 for almost every (ω, s) in $L^2(\Omega) \times \mathcal{H}$ because $X_{m+1} - S_m \neq 0$ a.s. In particular, $D_{\cdot}\mathbf{1}_{X_{m+1} > Sm} = 0$ in $L^2(\Omega) \times \mathcal{H}$, and for any $k \leq n$, the kth-order Malliavin derivative of $\mathbf{1}_{X_{m+1} > Sm}$ is 0 in $L^2(\Omega) \times \mathcal{H}^{\otimes k}$ as well.

This justifies the following computation, where equalities hold in $L^2(\Omega) \times \mathcal{H}^{\otimes n}$:

$$\begin{aligned}
D^{(n)}_{s_n,\cdots,s_2,s_1} S_{m+1} &= D^{(n-1)}_{s_n,\cdots,s_2} \left(D_{s_1} X_{m+1} \mathbf{1}_{X_{m+1} > Sm} + D_{s_1} S_m \mathbf{1}_{X_{m+1} < Sm} \right) \\
&= D^{(n-2)}_{s_n,\cdots,s_3} \left(\left[D_{s_2} D_{s_1} X_{m+1} \right] \mathbf{1}_{X_{m+1} > Sm} + \left[D_{s_2} D_{s_1} S_m \right] \mathbf{1}_{X_{m+1} < Sm} \right) \\
&\vdots \\
&= \left[D^{(n)}_{s_n,\cdots,s_2,s_1} X_{m+1} \right] \mathbf{1}_{X_{m+1} > Sm} + \left[D^{(n)}_{s_n,\cdots,s_2,s_1} S_m \right] \mathbf{1}_{X_{m+1} < Sm}.
\end{aligned} \tag{31}$$

Now, still following the strategy of the proof of Theorem 3.6, we let $\sigma_m^{*2} = \max\{\sigma^2(1); \cdots; \sigma^2(m)\}$, and we assume by induction that $\left\| D^{(n)}_{\cdot} S_m \right\|_2^2 \leq \sigma_m^{*2}$ almost surely. Our hypothesis and equality (31) implies that almost surely

$$\begin{aligned}
\left\| D^{(n)}_{\cdot} S_{m+1} \right\|_2^2 &= \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} \left| D^{(n)}_{s_n,\cdots,s_2,s_1} S_{m+1} \right|^2 ds_n \cdots ds_2 ds_1 \\
&= \mathbf{1}_{X_{m+1} > Sm} \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} \left| D^{(n)}_{s_n,\cdots,s_2,s_1} X_{m+1} \right|^2 ds_n \cdots ds_2 ds_1 \\
&\quad + \mathbf{1}_{X_{m+1} < Sm} \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} \left| D^{(n)}_{s_n,\cdots,s_2,s_1} S_m \right|^2 ds_n \cdots ds_2 ds_1 \\
&= \mathbf{1}_{X_{m+1} > Sm} \left\| D^{(n)}_{\cdot} X_{m+1} \right\|_2^2 + \mathbf{1}_{X_{m+1} < Sm} \left\| D^{(n)}_{\cdot} S_m \right\|_2^2 \\
&\leq \sigma^2(m+1) \mathbf{1}_{X_{m+1} > Sm} + \sigma_m^{*2} \mathbf{1}_{X_{m+1} < Sm} \\
&\leq \sigma_{m+1}^{*2}.
\end{aligned}$$

Since $\left\| D^{(n)}_{\cdot} X_1 \right\|_2^2 \leq \sigma_1^{*2} = \sigma^2(1)$ by hypothesis, induction implies the conclusion of the proposition when $m = N$. ∎

Combining the results of Theorem 4.7 and Proposition 4.11, we state the extension of the Borell-Sudakov inequality in two separate results, depending on whether $n = 2$ or $n \geq 3$.

Corollary 4.12 *Let X and μ be as in Proposition 4.11 with $n = 2$. Then $\sup_I X - \mu$ is a sub-2nd-chaos random variable. It can be decomposed as*

$$\sup_I X - \mu = \int_0^1 f(s)\, dW_s + X_2$$

where $f \in \mathcal{H}$ and X_2 is a sub-2nd-chaos r.v. relative to the scale $M = \pi\sqrt{10}\sup_{t\in I}\sigma^2(t)$. In particular we get the following extension of the Borell-Sudakov inequality: for any $u > 0$,

$$\mathbf{P}\left[|\sup_I X - \int_0^1 f(s)\, dW_s - \mu| > u\right] = \mathbf{P}\left[|X_2| > u\right] \leq 2\exp\left(-\frac{u}{M}\right). \tag{32}$$

Proof. The first statement follows immediately from the conclusion of Proposition 4.11 as applied to "Case $n = 2$" in Theorem 4.7. The second statement is an immediate consequence of the tail estimate in Lemma 4.6. ∎

The presence of the function G_n in Theorem 4.7 case $n \geq 3$ makes it impossible to apply Proposition 4.11 directly. Moreover, the conditional expectation in that same portion of the theorem causes further difficulties, making it necessary to impose slightly stronger conditions on $D^{(n)}X$ than in that theorem, in order to derive a Borell-Sudakov extension.

Proposition 4.13 *Let X and μ be as in Proposition 4.11 with $n \geq 3$. Recall the function G_n defined in "Case $n = 3$" of Theorem 4.7. Assume moreover that for any $t \in I$ and for any $s_n \in [0,1]$, there exists a non-random value $M(t)$ not dependent on s_n, such that, almost surely*

$$\int_{s_n}^1 \int_{s_n}^{s_1} \cdots \int_{s_n}^{s_{n-2}} G_n\left(M(t)^{-2}\left|D^{(n)}_{s_n,\cdots,s_2,s_1}X(t)\right|^{2/n}\right) ds_{n-1}\cdots ds_2 ds_1 \leq 1/2 \tag{33}$$

and

$$M(t) \geq \sqrt{2e}\left\|D^{(n)}X(t)\right\|_{\mathcal{H}}.$$

Then the random variable $\sup_I X - \mu$ is a sub-nth-chaos r.v. It can be decomposed as $\sup_I X - \mu = \sum_{m=1}^{n-1} I_m(f_m) + X_n$ where each f_m is a non-random symmetric function in $\mathcal{H}^{\otimes m}$, and X_n is a sub-nth-Gaussian chaos random variable relative to the scale

$$M = \sup_{t\in I} M(t).$$

In particular, the extension (32) of the Borell-Sudakov inequality holds for X_n with this M, namely

$$\mathbf{P}\left[|\sup_I X - \sum_{m=0}^{n-1} I_m(f_m) - \mu| > u\right] = \mathbf{P}\left[|X_n| > u\right] \leq 2\exp\left(-\left(\frac{u}{M}\right)^{2/n}\right). \tag{34}$$

Remark 4.14 *The hypothesis of this proposition is clearly satisfied if there exist constants $\sigma(t)$ such that almost surely, for all $s_1, s_2, \cdots, s_n$, $\left|D^{(n)}_{s_n,\cdots,s_2,s_1}X(t)\right| \leq \sigma(t)$. Then there is a constant k_n depending only on n such that we may take $M = k_n \sup_{t\in I}\sigma(t)$.*

Proof of Proposition 4.13. Here, we may not apply Proposition 4.11 directly. Instead, we return to its proof, and use the notation therein. Let $T_{n-1}(s_n) = \left\{(s_i)_{i=1}^{n-1} : s_1 \geq s_2 \geq \cdots \geq s_{n-1} \geq s_n\right\}$, a simplex for any fixed $s_n \in [0,1]$. Let $M_m = M(m)$. Also use the shorthand notation $\bar{s} = (s_i)_{i=1}^{n-1}$. By hypothesis we have

$$\int \cdots \int_{T_{n-1}(s_n)} G_n\left((M_m)^{-2}\left|D^{(n)}_{s_n,\bar{s}}X(m)\right|^{2/n}\right) d\bar{s} \leq \frac{1}{2}. \tag{35}$$

We also define

$$M_m^* = \max\{M_1, M_2, \cdots, M_m\}.$$

Then, since G_n is an increasing function, we have, from line (31),

$$\int_{T_{n-1}(s_n)}\cdots\int G_n\left(\left(M_{m+1}^*\right)^{-2}\left|D_{s_n,\bar{s}}^{(n)}S_{m+1}\right|^{2/n}\right)d\bar{s}$$

$$\begin{aligned}
&= \mathbf{1}_{X(m+1)>S_m}\int_{T_{n-1}(s_n)}\cdots\int G_n\left(\left(M_{m+1}^*\right)^{-2}\left|D_{s_n,\bar{s}}^{(n)}X(m+1)\right|^{2/n}\right)d\bar{s}\\
&+ \mathbf{1}_{X(m+1)<S_m}\int_{T_{n-1}(s_n)}\cdots\int G_n\left(\left(M_{m+1}^*\right)^{-2}\left|D_{s_n,\bar{s}}^{(n)}S_m\right|^{2/n}\right)d\bar{s}\\
&\leq \mathbf{1}_{X(m+1)>S_m}\int_{T_{n-1}(s_n)}\cdots\int G_n\left(\left(M_{m+1}\right)^{-2}\left|D_{s_n,\bar{s}}^{(n)}X(m+1)\right|^{2/n}\right)d\bar{s}\\
&+ \mathbf{1}_{X(m+1)<S_m}\int_{T_{n-1}(s_n)}\cdots\int G_n\left(\left(M_m^*\right)^{-2}\left|D_{s_n,\bar{s}}^{(n)}S_m\right|^{2/n}\right)d\bar{s}.
\end{aligned}$$

Thus, if we assume that

$$\int_{T_{n-1}(s_n)}\cdots\int G_n\left(\left(M_m^*\right)^{-2}\left|D_{s_n,\bar{s}}^{(n)}S_m\right|^{2/n}\right)d\bar{s} \leq \frac{1}{2}, \tag{36}$$

using (35), we obtain that (36) holds at rank $m+1$, and thus, by induction, for all $m \leq N$.

The definition (16) of G_n shows that the function $x \mapsto G_n\left(|x|^{2/n}\right)$ is convex for all x. Let $M = M_N^* = \max\{M_1, M_2, \cdots, M_N\}$. We may now write, using Jensen, and (36) for $m = N$,

$$\begin{aligned}
&\int_0^1 ds_n \int_{T_{n-1}(s_n)}\cdots\int G_n\left(M^{-2}\left|\mathbf{E}\left[D_{s_n,\bar{s}}^{(n)}S_N|\mathcal{F}_{s_n}\right]\right|^{2/n}\right)d\bar{s}\\
&\leq \int_0^1 ds_n \int_{T_{n-1}(s_n)}\cdots\int \mathbf{E}\left[G_n\left(M^{-2}\left|D_{s_n,\bar{s}}^{(n)}S_N\right|^{2/n}\right)|\mathcal{F}_{s_n}\right]d\bar{s}\\
&= \int_0^1 ds_n \mathbf{E}\left[\int_{T_{n-1}(s_n)}\cdots\int d\bar{s}G_n\left(M^{-2}\left|D_{s_n,\bar{s}}^{(n)}S_N\right|^{2/n}\right)|\mathcal{F}_{s_n}\right]\\
&\leq \frac{1}{2}.
\end{aligned}$$

This establishes Condition (18) of Theorem 4.7. We omit the details needed to check the other conditions of this Theorem. Inequality (34) is again only a consequence of Lemma 4.6. ■

The presence of the Wiener chaos correction terms $\sum_{m=0}^{n-1} I_m(f_m)$ in the statement of the generalizations (32) and (34) of Borell-Sudakov are somewhat of an annoyance, because these inequalities' proofs present no way of calculating the magnitude of the non-random functions $\{f_m : m = 1, \cdots, n-1\}$. We propose an additional result which shows that asymptotically, these functions are irrelevant.

Corollary 4.15 *With the hypotheses and notation as in Corollary 4.12 or Proposition 4.13, we have for any $\varepsilon > 0$, for u large enough,*

$$\mathbf{P}\left[\left|\sup_I X - \mu\right| > u\right] \leq 2(1+\varepsilon)\exp\left(-\frac{1}{(1+\varepsilon)}\left(\frac{u}{M}\right)^{2/n}\right).$$

More concisely, we can write

$$\lim_{u\to\infty}\frac{1}{u^{2/n}}\log\mathbf{P}\left[\left|\sup_I X - \mu\right| > u\right] \leq -\frac{1}{M^{2/n}}.$$

Proof. First note that, for any $r \in (0,1)$

$$\begin{aligned}
\mathbf{P}\left[\left|\sup_I X - \mu\right| > u\right] &= \mathbf{P}\left[\left|X_n + \sum_{m=1}^{n-1} I_m(f_m)\right| > u\right] \\
&\leq \mathbf{P}\left[|X_n| > u - \sum_{m=1}^{n-1}|I_m(f_m)|\right] \\
&\leq \mathbf{P}\left[|X_n| > u - \sum_{m=1}^{n-1}|I_m(f_m)| ; \sum_{m=1}^{n-1}|I_m(f_m)| \leq ru\right] + \mathbf{P}\left[\sum_{m=1}^{n-1}|I_m(f_m)| > ru\right] \\
&\leq \mathbf{P}\left[|X_n| > (1-r)u\right] + \sum_{m=1}^{n-1}\mathbf{P}\left[|I_m(f_m)| > \frac{ru}{n-1}\right] \qquad (37)
\end{aligned}$$

The following lemma is a trivial consequence of the results in [3].

Lemma 4.16 *Let $f \in \mathcal{H}^{\otimes m}$. Then there exists a constant $M_m(f)$ such that*

$$\mathbf{P}\left[|I_m(f_m)| > u\right] \leq \exp\left(-\left(\frac{u}{M_m(f)}\right)^{2/m} + \eta_m(u)\,u^{2/m}\right)$$

where $\lim_{u\to 0}\eta_m(u) = 0$.

Armed with this Lemma, and with the inequalities (32) or (34), and choosing r so that $(1-r) > (1+\varepsilon)^{-n/2}$, we may write from (37)

$$\begin{aligned}
\mathbf{P}\left[\left|\sup_I X - \mu\right| > u\right] &\leq \mathbf{P}\left[|X_n| > u/(1+\varepsilon)^{-n/2}\right] + \sum_{m=1}^{n-1}\mathbf{P}\left[|I_m(f_m)| > \frac{ru}{n-1}\right] \\
&\leq 2\exp\left(-\frac{1}{(1+\varepsilon)}\left(\frac{u}{M}\right)^{2/n}\right) \\
&\quad + \sum_{m=1}^{n-1}\exp\left(-\left(\frac{ru}{M_m(f_m)(n-1)}\right)^{2/m} + \eta_m(u)\,u^{2/m}\right) \\
&\leq 2\exp\left(-\frac{1}{(1+\varepsilon)}\left(\frac{u}{M}\right)^{2/n}\right) + (n-1)\exp\left(-\frac{1}{2}\left(\frac{u}{K}\right)^{2/(n-1)}\right) \qquad (38) \\
&\leq 2(1+\varepsilon)\exp\left(-\frac{1}{(1+\varepsilon)}\left(\frac{u}{M}\right)^{2/n}\right), \qquad (39)
\end{aligned}$$

where in line (38), the constant K is $(n-1)\max_{m\in\{1,\cdots n-1\}} M_m(f_m)$ and u is chosen so large that for every $m \leq n-1$, $\eta_m(u) < 2^{-1}(r/K)^{2/m}$, while in line (39), u is chosen so large that the second term in (38) is less than ε times the first. The first statement of the corollary is proved, and the second follows trivially due to the fact that $\varepsilon > 0$ is arbitrary. ■

5 Appendix

5.1 Efficient constant in the Burkholder-Davis-Gundy inequality

Proposition 5.1 *For any square integrable martingale Y, and any $p \geq 2$, we have*

$$\mathbf{E}\left[\sup_{s\in[0,t]} |Y(s)|^p\right] \leq c(p)\,\mathbf{E}\left[|\langle Y\rangle(t)|^{p/2}\right]$$

where the constant $c(p)$ satisfies $c(2) = 1$ and for any $p > 2$,

$$c(p) = \left(\frac{1}{2}\frac{p^{p+1}}{(p-1)^{p-1}}\right)^{p/2} \leq \left(\sqrt{e/2}\right)^p p^p.$$

Proof. One only needs to keep track of the constants in the classical proof of this inequality: starting with Ito's formula (where the function $f(x) = |x|^p$ is of class C^2),

$$\begin{aligned}
\mathbf{E}\,|Y(t)|^p &= \mathbf{E}\left[\int_0^t p\,|Y(s)|^{p-1}\operatorname{sgn}(Y(s))\,dY(s) + \frac{1}{2}\int_0^t p(p-1)\,|Y(s)|^{p-2}\langle Y\rangle(ds)\right] \\
&= \frac{p(p-1)}{2}\mathbf{E}\left[\int_0^t |Y(s)|^{p-2}\langle Y\rangle(ds)\right] \leq \frac{p(p-1)}{2}\mathbf{E}\left[\left(\sup_{s\in[0,t]}|Y(s)|\right)^{p-2}\langle Y\rangle(t)\right] \\
&\leq \frac{p(p-1)}{2}\mathbf{E}\left[\left(\sup_{s\in[0,t]}|Y(s)|\right)^{p}\right]^{(p-2)/p}\mathbf{E}\left[|\langle Y\rangle(t)|^{p/2}\right]^{2/p}.
\end{aligned}$$

The proposition's constant $c(p)$ follows from some elementary calculations and Doob's inequality

$$\mathbf{E}\left[\left(\sup_{s\in[0,t]}|Y(s)|\right)^p\right] \leq (p/(p-1))^p \sup_{s\in[0,t]}\mathbf{E}\left[|Y(s)|^p\right].$$

The second statement in the proposition is equally elementary. ∎

5.2 Proof of Lemma 3.7

Such a Φ as in the statement of the lemma can be replaced by an approximation Φ_m such that Φ_m is of class C^1, such that $\Phi = \Phi_m$ for all points distant by more than $1/m$ of all hyperplanes, and such that $\Phi - \Phi_m$ and $\nabla\Phi_m$ are both bounded uniformly in m by multiples of $|\nabla\Phi|_\infty$: this can be achieved by interpolating Φ and $\nabla\Phi$ from the boundary of the $1/m$-neighborhood T_m of the union T of the hyperplanes using scaled polynomials. For example, in the case we are interested in, let P be a polynomial of degree 4 on $[-1,1]$, which is increasing and convex, such that $P(-1) = P'(-1) = 0$ and $P(1) = P'(1) = 1$. Define the function $\Phi_m = \Phi$ off the set $T_m = \{|x-y| < 1/m\}$, and on that set define $\Phi_m(x,y) = m^{-1}P(m(x-y)) + y$. This sequence Φ_m has the required property, and in fact $|\nabla\Phi_m|_\infty \leq 1$ and $|\Phi - \Phi_m|_\infty \leq 1$. Now since Φ_m converges to Φ pointwise, the dominated convergence theorem implies that $\Phi_m(Z)$ converges to $\Phi(Z)$ in $L^2(\Omega)$. Moreover, we can write using the chain rule (6) for C^1 functions: $\Phi_m(Z) \in \mathbf{D}^{1,2}$ and

$$D_r\Phi_m(Z) = (1 - \mathbf{1}_{T_m}(Z))\,\nabla\Phi(Z)\,D_rZ + \mathbf{1}_{T_m}(Z)\,\nabla\Phi_m(Z)\,D_rZ.$$

Since $\mathbf{1}_{T_m}(Z)$ converges to 0 almost surely, and $D.Z \in L^2(\Omega\times[0,1])$, by the dominated convergence theorem in $L^2(\Omega\times[0,1])$, we have $D.\Phi_m(Z)$ converging to $\nabla\Phi(Z)\,D.Z$ in that space. Now we invoke the fact (see [16]) that the Malliavin derivative operator D is a closed operator from its domain $\mathbf{D}^{1,2}$ into $L^2(\Omega\times[0,1])$, to conclude that $\Phi(Z) \in \mathbf{D}^{1,2}$ and $D.\Phi(Z) = \nabla\Phi(Z)\,D.Z$, as was to be proved.

5.3 Proof of Lemma 4.6

The proofs of this lemma's statements are elementary; we detail some of them. First, we have using Chebyshev's inequality:

$$\begin{aligned}\mathbf{P}\left[|X|>u\right]&=\mathbf{P}\left[\exp(X/\delta)^{2/n}>\exp(u/\delta)^{2/n}\right]\\&\leq\exp\left(-(u/\delta)^{2/n}\right)\mathbf{E}\left[\exp(X/\delta)^{2/n}\right]\\&\leq 2\exp\left(-(u/\delta)^{2/n}\right),\end{aligned}$$

which is the first statement of the lemma. This then implies that

$$\begin{aligned}\mathbf{E}\left[X^2\right]&=\int_0^\infty\mathbf{P}\left[|X|>\sqrt{u}\right]du\\&\leq 2\int_0^\infty\exp\left(-(\sqrt{u}/\delta)^{2/n}\right)du=2\delta^2\int_0^\infty e^{-v^{1/n}}dv=v_n\delta^2\end{aligned}$$

where $v_n=2\int_0^\infty e^{-v^{1/n}}dv$, hence the second statement. The proof of the estimate for $\mathbf{E}\left[\exp\left(\lambda|X|^{1/n}\right)\right]$ is left to the reader.

For the first converse, let $c>1$ be fixed. Using the estimate $\mathbf{P}\left[|X|>u\right]\leq 2\exp\left(-(u/\delta)^{2/n}\right)$, we get

$$\begin{aligned}\mathbf{E}\left(\exp\left[\left(\frac{X}{c\delta}\right)^{2/n}\right]\right)&=\int_0^\infty\mathbf{P}\left[|X|>c\delta(\log r)^{n/2}\right]dr\\&\leq 1+\int_1^\infty 2\exp\left(-c^{2/n}\log r\right)dr\\&=1+2\int_1^\infty r^{-c^{2/n}}dr\\&=1+\frac{2}{c^{2/n}-1}.\end{aligned}$$

Thus we only need to choose $v_n'=c=3^{n/2}$. The proofs of the other converses are left to the reader.

5.4 Proof of Lemma 4.10

The proof uses three simple facts from the theory of Wiener chaoses. For any symmetric function g in $\mathcal{H}^{\otimes m}$, the first fact is simply the definition of $I_m(g)$ as an iterated Itô integral in (3). The second, from Step 2 in Section 2, is the calculation $D_rI_m(g)=mI_m(g(\cdot,r))$. The last, from Lemma 1.2.4 in [16], says that

$$\mathbf{E}\left[I_m(g_m)\,|\mathcal{F}_t\right]=I_m\left(g_m\mathbf{1}_{[0,t]}^{\otimes m}\right).$$

For $X\in\mathbf{D}^{n,2}$, we may now calculate, for $s_n\leq s_{n-1}\leq\cdots\leq s_2\leq s_1\leq 1$,

$$\begin{aligned}D^{(n)}_{s_n,\cdots,s_2,s_1}X&=D^{(n)}_{s_n,\cdots,s_2}\left(\sum_{m=1}^\infty mI_{m-1}\left(f_m(s_1,\cdot)\right)\right)\\&=\sum_{m=2}^\infty m(m-1)D^{(n)}_{s_n,\cdots,s_3}\left(I_{m-2}\left(f_m(s_1,s_2,\cdot)\right)\right)\\&\;\;\vdots\\&=\sum_{m=n}^\infty m(m-1)\cdots(m-n+1)I_{m-n}\left(f_m(s_1,s_2,\cdots,s_n,\cdot)\right)\end{aligned}$$

Thus we obtain

$$\mathbf{E}\left[I_{m-n}\left(f_m\left(s_1,s_2,\cdots,s_n,\cdot\right)\right)|\mathcal{F}_{s_n}\right]=I_{m-n}\left(h_{m,s_1,s_2,\cdots,s_n}\right)$$

where the function h above is defined by

$$h_{m,s_1,s_2,\cdots,s_n}\left(s_{n+1},\cdots,s_m\right)=f_m\left(s_1,\cdots,s_m\right)\prod_{j=n+1}^{m}\mathbf{1}_{s_j\le s_n},$$

which proves that $h_{m,s_1,s_2,\cdots,s_n}$ is symmetric in the variables $s_{n+1},\cdots,s_m$, and thus we can write

$$\mathbf{E}\left[I_{m-n}\left(f_m\left(s_1,s_2,\cdots,s_n,\cdot\right)\right)|\mathcal{F}_{s_n}\right]=(m-n)!\int_0^{s_n}\int_0^{s_{n+1}}\cdots\int_0^{s_{m-1}}f_m\left(s_1,\cdots,s_m\right)dW_{s_m}\cdots dW_{s_{n+1}}.$$

The following calculation now finishes the proof of the lemma:

$$\begin{aligned}
&\int_0^1\int_0^{s_1}\cdots\int_0^{s_{n-1}}\mathbf{E}\left[D^{(n)}_{s_n,\cdots,s_2,s_1}X|\mathcal{F}_{s_n}\right]dW_{s_n}\cdots dW_{s_1}\\
&=\sum_{m=n}^{\infty}\frac{m!}{(m-n)!}\int_0^1\int_0^{s_1}\cdots\int_0^{s_{n-1}}\mathbf{E}\left[I_{m-n}\left(f_m\left(s_1,s_2,\cdots,s_n,\cdot\right)\right)|\mathcal{F}_{s_n}\right]dW_{s_n}\cdots dW_{s_1}\\
&=\sum_{m=n}^{\infty}m!\int_0^1\int_0^{s_1}\cdots\int_0^{s_{n-1}}\left[\int_0^{s_n}\int_0^{s_{n+1}}\cdots\int_0^{s_{m-1}}f_m\left(s_1,\cdots,s_m\right)dW_{s_m}\cdots dW_{s_{n+1}}\right]dW_{s_n}\cdots dW_{s_1}\\
&=\sum_{m=n}^{\infty}I_m\left(f_m\right).
\end{aligned}$$

References

[1] Adler, R. (1990). *An introduction to continuity, extrema, and related topics for general Gaussian processes.* Inst. Math. Stat., Hayward, CA.

[2] Alòs, E.; Mazet, O.; Nualart, D. (2001). Stochatic calculus with respect to Gaussian processes. *Annals of Probability* **29,** 766-801.

[3] Borell, C. (1978). Tail probabilities in Gauss space. In *Vector Space Measures and Applications, Dublin 1977.* Lecture Notes in Math. **644**, 71-82. Springer-Verlag.

[4] Carmona, R.A.;. Molchanov, S.A. (1994). *Parabolic Anderson Model and Intermittency.* Memoirs A.M.S. **418.**

[5] Cheridito, P.; Nualart, D. (2005). Stochastic integral of divergence type with respect to fractional Brownian motion with Hurst parameter H in (0,1/2). To appear in the *Annales de l'Institute Henri Poincaré (B) Probability and Statistics.*

[6] Decreusefond, L.; Üstünel, A.-S. (1997). Stochastic analysis of the fractional Brownian motion. *Potential Analysis*, **10**, 177-214.

[7] Hu, Y.-Z.; Oksendal, B.; Sulèm, A. (2003). Optimal consumption and portfolio in a Black-Scholes market driven by fractional Brownian motion. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **6**, no. 4, 519–536.

[8] Florescu, I.; Viens, F. (2005) Sharp estimation of the almost-sure Lyapunov exponent for the Anderson model in continuous space. In press in *ProbabTheory and Rel. Fields*, 2006, 1-42.

[9] Karatzas, I.; Shreve, S. (1988). *Brownian Motion and Stochastic Calculus.* Springer-Verlag.

[10] Ledoux, M. (1996). *Isoperimetry and Gaussian analysis.* In: Lectures on probability theory and statistics (Saint-Flour, 1994), 165–294, Lecture Notes in Math., **1648**, Springer-V.

[11] Ledoux, M.; Talagrand, M. (1991) *Probability in Banach Spaces.* Springer-Verlag.

[12] Malliavin, P. (2002). *Stochastic Analysis.* Springer-V.

[13] Malliavin, P.; Thalmaier, A. (2005). *Stochastic Calculus of Variations in Mathematical Finance.* Springer-V.

[14] Maslowski, B.; Nualart, D. (2003). Stochastic evolution equations driven by fBm. *Journal of Functional Analysis* **202**, 277-305.

[15] Mocioalca, O.; Viens, F. (2004). Skorohod integration and stochastic calculus beyond the fractional Brownian scale. *Journal of Functional Analysis*, **222** (2), 385-434.

[16] Nualart, D. (1995). *The Malliavin calculus and related topics.* Springer-V., New-York.

[17] Nualart, D.; Viens, F. (2000). Evolution equation of a stochastic semigroup with white-noise drift. Ann. Probab. **28**, no. 1, 36–73.

[18] Rovira, C.; Tindel, S. (2005). *On the Brownian directed polymer in a Gaussian random environment. J. Functional Analysis*, **222**, no. 1, 178–201.

[19] Talagrand, M. (1990). Sample boundedness of stochastic processes under increment conditions. *Ann. Prob.* **18**, no. 1, 1-49.

[20] Tindel, S.; Tudor, C.A.; Viens, F. (2004). Sharp Gaussian regularity on the circle and application to the fractional stochastic heat equation. *Journal of Functional Analysis*, **217** (2), 280-313.

[21] Tudor, C.A.; Viens, F. (2005). *Statistical Aspects of the Fractional Stochastic Calculus.* Preprint.

[22] Üstünel, A.-S. (1995). *An Introduction to Analysis on Wiener Space.* Lecture Notes in Mathematics, **1610**. Springer-Verlag.

[23] Weber, M. (1998). Stochastic processes with values in exponential type Orlicz spaces. *Ann. Prob.* **16,** 1365-1371.