Supremum Concentration Inequality and Modulus of Continuity for Sub-nth Chaos Processes

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

[1]Dept. Statistics and Dept. Mathematics; viens@purdue.edu (corresponding author),
[2] Dept. Mathematics; avizcarr@purdue.edu,

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

May 7, 2007

Abstract

This article provides a detailed analysis of the behavior of suprema and moduli of continuity for a large class of random fields which generalize Gaussian processes, sub-Gaussian processes, and random fields that are in the nth chaos of a Wiener process. An upper bound of Dudley type on the tail of the random field's supremum is derived using a generic chaining argument; it implies similar results for the expected supremum, and for the field's modulus of continuity. We also utilize a sharp and convenient condition using iterated Malliavin derivatives, to arrive at similar conclusions for suprema, via a different proof, which does not require full knowledge of the covariance structure.

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.

[*]Both authors partially supported by NSF grant 0606615

1 Introduction

The regularity properties of random processes have long been studied, going as far back as Kolmogorov's celebrated chaining argument and criterion (see [12, Theorem I.2.1]), and in the 70's and 80's the sharp work of Fernique, Talagrand, and others (see [6], [13]) for the Gaussian case. These latter works drew upon ideas of R. Dudley, who in 1967 (see [4]) gave a sufficient condition for the boundedness of Gaussian processes based on the so called ε-entropy integral. All these results, of upper bound type, are largely also valid in the case of sub-Gaussian processes (see [8]). The question remains of how many such results are true for other classes of processes. While the majorizing measure conditions of Talagrand show that supremum estimates can be achieved for processes with tails decaying no slower than exponentially (in the nomenclature of [8], these are processes with increments in the Orlicz space relative to convex Young functions of the type $\exp(z^q) - 1$, with $q \geq 1$), the authors of this note provided in [17] an extension of Dudley's entropy upper bound to all $q = 2/n$ with n an integer.

They presented their work by defining a new class of processes with the so-called *sub-nth chaos* property. These sub-nth chaos processes are an extension of sub-Gaussian processes. Such a process (random field) X on an arbitrary index set I is essentially required to have increments $X(t) - X(s)$ with tails that decay no slower than $\exp(-c|z|^{2/n}/\delta(s,t))$ where δ is some pseudo-metric on I. This definition was motivated by results including: (i) the fundamental observation [17, Lemma 3.3], that if the Malliavin derivative of a random variable X is almost surely bounded, then X must be sub-Gaussian; and (ii) the discovery in that same article of conditions on the nth Malliavin derivatives extending this observation to the sub-nth chaos case.

Also in [17], using a new concise Malliavin-derivative-based proof, a Borell-Sudakov[1]-type concentration inequality for such processes was proved, which shows that the supremum of a sub-nth chaos process is again a sub-nth chaos random variable with a well-controlled scale. While the proof in [17] completely generalized the standard Borell-Sudakov inequality to the sub-Gaussian and sub-2nd-chaos cases, it ran into inefficiencies in the case of higher order chaos.

In this article we correct these inefficiencies by combining a new use of iterated Malliavin derivatives, via a fractional exponential Poincaré lemma, with the innovative relation between Malliavin derivatives and suprema of processes discovered in [17]. This technique provides a new usage of the versatile Malliavin calculus, whose many applications are discussed in [9]; we summarize the Malliavin derivative's properties that we use herein, making our treatment essentially self-contained.

This article also contains an alternate way of establishing concentration inequalities for suprema of sub-nth chaos processes. Inspired by Ledoux and Talagrand's generic chaining arguments (see [14]), we use chaining to prove a Dudley entropy upper bound which holds in its usual expectation form, as well as in a tail form, for suprema of sub-nth chaos processes. These bounds allow us to prove, without the use of Malliavin derivatives, that concentration for these suprema is of sub-nth-chaos type, with respect to a scale which we estimate precisely, generalizing, up to a universal constant, the classical concentration of Borell-Sudakov-type for Gaussian processes.

The Malliavin-derivative-based and chaining-argument-based proofs of sub-nth-chaos supremum concentration use hypotheses that are evidently morally close, but it is not yet possible to say how close. Our fractional exponential Poincaré inequality appears to indicate that Malliavin-derivative conditions are slightly stronger. They do allow a simpler, much more accessible proof. We also believe that in many Brownian-based applications, checking that an nth Malliavin derivative is almost surely bounded may be significantly easier than checking the moment conditions for the sub-nth chaos property, particularly for solutions of stochastic equations; this is due to the need to use iterations of Itô's formula for the latter, which is algebraically

[1] Borell-Sudakov concentration inequalities are also known as Borell-Sudakov-Tsirel'son inequalities. Because of the appearance in the work of R. Dudley of these results' basic elements for scalar random variables, the inequalities should perhaps be called Dudley-Borell-Sudakov-Tsirel'son. For conciseness, we will continue to use the appellation Borell-Sudakov.

usually more difficult than calculating iterated Malliavin derivatives. In specific problems, it should be easy to identify which condition is most convenient.

As an application of our estimates, we choose a very basic and general one, which does not, in principle, rely on Malliavin derivative calculations: we derive almost-sure moduli of continuity for sub-nth chaos processes, with results that generalize those stated for general one-parameter Gaussian processes such as in [15]. The fractional exponential Poincaré inequality then allows one to use Malliavin derivative estimates to check that the sub-nth chaos conditions are satisfied; in this sense, Malliavin derivatives can be crucial in establishing path regularity.

Throughout this paper, $d\bar{s}$ denotes the Lebesgue measure on $(\mathbf{R}_+)^k$ for any integer k. Summarizing, our main results can be stated (partially) as follows.

- Assume there exists a positive integer n and a non-random pseudo-metric δ on an index set I such that, for a separable random field X on I, either of the following two conditions hold for all pairs $(x,y) \in I^2$:

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

  or, almost surely,

$$\left|D_{\cdot}^{(n)}\left(X(x)-X(y)\right)\right|^2_{L^2(dr)} := \int_{(\mathbf{R}_+)^n}\left|D_{\bar{s}}^{(n)}\left(X(x)-X(y)\right)\right|^2 d\bar{s} \leq \delta^2(x,y).$$

  where $D^{(n)}$ is the nth iterated Malliavin derivative with respect to a Brownian motion defined on the same probability space as X. Then,

$$\mu := \mathbf{E}\left[\sup_{t\in I} X_t\right] \leq C_n \int_0^\infty \left(\log N_\delta(\varepsilon)\right)^{n/2} d\varepsilon$$

  and

$$\mathbf{P}\left[\left|\sup_{t\in I} X_t - \mu\right| > u\right] \leq 2\exp\left(-\frac{1}{2}\left(\frac{u}{C_n D}\right)^{2/n}\right),$$

  where C_n is a constant depending only on n, $N_\delta(\varepsilon)$ is the smallest number of δ-balls of radius ε needed to cover I, and D is the diameter of I under δ.

- Moreover, if I is an interval (or a smooth one-dimensional manifold) and $\delta(x,y) \leq d(|x-y|)$ for some increasing function d, then the function

$$f_d(h) := d(h)(\log(1/h))^{n/2}$$

  is, up to a constant depending on n only, almost surely a uniform modulus of continuity for X on I.

The structure of this article is the following: Section 2 gives the definitions of the sub-nth chaos property, and introduces the context of supremum concentration inequalities. Section 3 gives these inequalities using the generic chaining method. Section 4 uses the iterated Malliavin derivatives method. Section 5 finishes the article with the example of almost-sure moduli of continuity.

2 Sub-nth chaos processes

2.1 Definitions

We first recall the definition of sub-nth chaos random variables and processes as an extension of sub-Gaussian processes to heavier-tailed objects.

Definition 2.1 *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*

$$\mathbf{E}\left[\exp\left(\left(\frac{X}{M}\right)^{2/n}\right)\right] \leq 2. \tag{1}$$

When $n = 1$, such an X is sub-Gaussian relative to the scale $\sqrt{5}M$, see [17].

Remark 2.2 *Note that the following statement implies the sub-nth-Gaussian chaos property (1) and is also implied by it, with different universal multiplicative constant c in each implication: for all $u > 0$*

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

Specifically, (1) implies (2) with $c = 1$, and (2) implies (1) with $c = 1/3$.

Definition 2.3 *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 2.4 *Let δ and X be as in the previous definition. We use the notation $N_\delta(\varepsilon)$, and say that $N_\delta(\varepsilon)$ is a metric entropy for X, if $N_\delta(\varepsilon)$ is the smallest number of balls of radius ε in the pseudo-metric δ needed to cover I.*

2.2 Overview of Supremum Concentration Inequalities

In the next two sections, we prove a tail estimate version of the Dudley upper bound for sub-nth chaos processes, use it to derive estimates on the location and concentration of a sub-nth chaos process's supremum, and obtain similar concentration results via Malliavin derivative conditions.

For reference and comparison, we recall the statements of standard Gaussian estimates of supremum location and concentration. Let Z be a separable centered Gaussian field on an index set I that is almost-surely bounded, with canonical metric δ defined by $\delta^2(x,y) = \mathbf{E}\left[(Z(x) - Z(y))^2\right]$. Let $Y = \sup_I Z$. The so-called Dudley upper bound is $\mathbf{E}[Y] \leq K\int_0^\infty \sqrt{\log N_\delta(\varepsilon)}d\varepsilon$ where K is a universal constant (see [4]) and $N_\delta(\varepsilon)$ is defined as in the previous section. A corresponding lower bound, due to Fernique (see [5]) with a smaller universal constant, is known to hold for homogeneous (shift invariant in law) Gaussian processes on groups. The often-called Borell-Sudakov (or simply Borell) inequality, which may be improperly named as it first appears at the scalar process level in Dudley's work, is $\mathbf{P}[|Y - \mathbf{E}Y| > u] \leq 2\exp\left(-u^2/2\sigma^2\right)$ where $\sigma^2 := \max_{x\in I} Var[Z(x)]$. This result says precisely that the supremum of a Gaussian field is sub-Gaussian with respect to the scale σ.

These Gaussian results (except for Fernique's lower bound) were generalized in [17] to sub-Gaussian fields, and, under some additional assumption on Malliavin derivatives, to sub-2nd-chaos fields. In the next section, by using a general chaining argument, we show that such generalizations do not need stronger assumptions, and work in all chaoses. We also prove a Malliavin-derivatived-based concentration result, using a non-chaining argument, in Section 4. Both concentration results show, analogously to the Gaussian case, that the random field X's supremum is a sub-nth chaos random variable, relative to scales which consistently generalize the σ in the Gaussian case above. This means that concentration results for suprema of sub-nth chaos fields can be obtained using two separate sets of hypotheses: one using standard sub-nth chaos estimates on X as defined in the previous section, and one using Malliavin-derivative boundedness conditions.

It may be said that the concentration results via Malliavin derivatives are more powerful, because they do not require that the process have the sub-nth chaos property: they only require that a process's supremum have a finite expectation, and that each one-dimensional distribution of the process be a sub-nth chaos random variable. This very weak assumption on the process's covariance structure was not known to be sufficient for Borell-Sudakov concentration until it was noticed in [17]. This article proves that it is sufficient for sub-nth-chaoses as well.

At the same time, at the random variable level, because of Proposition 4.2 (consequence of a fractional exponential Poincaré inequality proved in the Appendix), the Malliavin derivative conditions seem stronger than mere sub-nth-chaos conditions, although this is not yet clear that the gap between the two is of any significance. It is clear, though, that depending on the situation, one or the other set of hypotheses may be most advantageous.

3 Generic chaining argument method

Our first result is a tail estimate similar to the Dudley inequality, which generalizes the latter to sub-nth-chaos processes.

Theorem 3.1 *For each fixed positive integer n, there exist universal constants C_n and 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 for each $t_0 \in I$ and $s \geq 0$,*

$$\mathbf{P}\left[\sup_{t\in I}|X_t - X_{t_0}| > C_n M + sC'_n D\right] \leq 2e^{-s^{2/n}/2}$$

where $M = \int_0^\infty (\log N_\delta(\varepsilon))^{n/2} d\varepsilon$ and $D = Diam_\delta(I)$ is the diameter of I under δ.

In addition, if D and M are finite, then X is almost-surely bounded.

The above constants can be taken to be equal to

$$C_n = 2^{2n+1}(q+1)/(1-q^{-1}) f_n(q),$$
$$C'_n = \frac{1}{2}(q+1)/\left(1-q^{-1}\right),$$

where q is any fixed value > 1 and $f_n(q) = \sum_{\ell=1}^\infty q^{-\ell+1}\ell^{n/2}$. One may optimize q to obtain a value of C'_n that does not depend on n: with $q = 1+\sqrt{2}$, $C'_n = \sqrt{2} + 3/2$. Depending on the respective values of M and D, it may be preferable to minimize C_n instead, in which case C'_n will depend on n, or to perform some other optimization which takes s into consideration.

One may remove the assumption that X is separable, by changing the definition of the probability in Theorem 3.1: generically define

$$\mathbf{P}\left[\sup_{t\in I}|X_t| > C\right] = \sup\left\{\mathbf{P}\left[\sup_{t\in F}|X_t| > C\right] : F \subset I; F \text{ finite.}\right\}.$$

The same remark holds for other theorems below which require separability, replacing expectations of suprema by suprema of expectations of suprema over finite sets. We will not comment further on this point.

Proof of Theorem 3.1. Since X is separable, by the monotone convergence theorem, we may and do assume that I is finite. Without loss of generality, we can assume that $X_{t_0} = 0$; if this is not true, consider the random field $Y_t = X_t - X_{t_0}$.

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 I. One may therefore construct a partition $\mathcal{A}_\ell$ of I of cardinality N_ℓ on the basis of this covering with sets of diameter less than $2q^{-\ell}$. Denote by I_ℓ the collection of points which are centers of each ball A of $\mathcal{A}_\ell$. For each t in I, 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 I_ℓ such that $t \in A_\ell(s_\ell(t))$. Note that $\delta(t, s_\ell(t)) \le q^{-\ell}$ for every t and $\ell \ge \ell_0$. Also note that

$$\delta(s_\ell(t), s_{\ell-1}(t)) \le q^{-\ell} + q^{-\ell+1} = (q+1)q^{-\ell}. \tag{3}$$

Hence, by the second inequality in Lemma 4.6 of [17], 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 above sum, we thus get that almost surely for every t,

$$X_t = X_{t_0} + \sum_{\ell>\ell_0}\left(X_{s_\ell(t)} - X_{s_{\ell-1}(t)}\right) = \sum_{\ell>\ell_0}\left(X_{s_\ell(t)} - X_{s_{\ell-1}(t)}\right) \tag{4}$$

where $s_{\ell_0}(t) := t_0$. This decomposition is the basis of the so-called chaining argument.

Let $\{c_\ell\}_{\ell>\ell_0}$ be a sequence of positive numbers which will be chosen later. Note that if $\forall t \in I$ and $\forall \ell > \ell_0, |X_{s_\ell(t)} - X_{s_{\ell-1}(t)}| \le c_\ell$ then from (4), $\forall t \in I, |X_t| \le \sum_{\ell>\ell_0} c_\ell$. Then

$$\begin{aligned}
\mathbf{P}\left(\sup_{t\in I}|X_t| > \sum_{\ell>\ell_0} c_\ell\right) &= \mathbf{P}\left(\exists t \in I : |X_t| > \sum_{\ell>\ell_0} c_\ell\right)\\
&\le \mathbf{P}(\exists t \in I, \exists \ell > \ell_0 : |X_{s_\ell(t)} - X_{s_{\ell-1}(t)}| > c_\ell)\\
&\le \mathbf{P}(\exists \ell > \ell_0, \exists t \in I_\ell, \exists t' \in I_{\ell-1} : |X_t - X_{t'}| > c_\ell)\\
&\le \sum_{\ell>\ell_0} 2N_\ell^2 \exp\left(-\left(\frac{c_\ell}{\delta(t,t')}\right)^{2/n}\right) &(5)\\
&\le 2\sum_{\ell>\ell_0} N_\ell^2 \exp\left(-\left(\frac{c_\ell}{(q+1)q^{-\ell}}\right)^{2/n}\right) &(6)\\
&= 2\sum_{\ell>\ell_0} \exp\left(2\log N_\ell - b_\ell^{2/n}\right) &(7)
\end{aligned}$$

where in line (5) we used (2), and in line (6) we used (3), and in line (7) $c_\ell := (q+1)q^{-\ell} b_\ell$.

Set

$$b_\ell = 8^{n/2}[(\log N_\ell)^{n/2} + (\ell-\ell_0)^{n/2}] + s. \tag{8}$$

Since for any $u, v, \alpha \ge 0$, $(u+v)^\alpha$ exceeds $(u^\alpha + v^\alpha)/2$, we have that

$$\begin{aligned}
b_\ell^{2/n} &\ge 4[(\log N_\ell)^{n/2} + (\ell-\ell_0)^{n/2}]^{2/n} + s^{2/n}/2\\
&\ge 2[\log N_\ell + (\ell-\ell_0)] + s^{2/n}/2. &(9)
\end{aligned}$$

Thus,

$$\begin{aligned}
2\sum_{\ell>\ell_0} \exp\left(2\log N_\ell - b_\ell^{2/n}\right) &\le 2\sum_{\ell>\ell_0} \exp\left(2\log N_\ell - 2\log N_\ell - 2(\ell-\ell_0) - s^{2/n}/2\right)\\
&\le 2e^{-s^{2/n}/2}\sum_{\ell>\ell_0} \exp(-2(\ell-\ell_0))\\
&\le 2e^{-s^{2/n}/2}.
\end{aligned}$$

To finish the proof, we note that $N_\ell \geq 2$ for $\ell > \ell_0$ and that the series $f_n(q) := \sum_{\ell>0} q^{-\ell+1}\ell^{n/2}$ converges, hence

$$\begin{aligned}\sum_{\ell>\ell_0} c_\ell &= (q+1)8^{n/2}\sum_{\ell>\ell_0} q^{-\ell}[(\log N_\ell)^{n/2} + (\ell-\ell_0)^{n/2}] + s\,(q+1)\sum_{\ell>\ell_0} q^{-\ell} \\ &= (q+1)8^{n/2}\left(\sum_{\ell>\ell_0} q^{-\ell}(\log N_\ell)^{n/2} + q^{-\ell_0-1}f_n(q)\right) + s\,(q+1)\sum_{\ell>\ell_0} q^{-\ell} \\ &= (q+1)8^{n/2}q^{-\ell_0-1}\left(\sum_{\ell>\ell_0} q^{-\ell+\ell_0+1}(\log N_\ell)^{n/2} + f_n(q)\right) + s\frac{q+1}{1-q^{-1}}q^{-\ell_0-1}.\end{aligned}$$

Now with $A := \sum_{\ell>\ell_0} q^{-\ell+\ell_0+1}(\log N_\ell)^{n/2} \geq \left(\frac{1}{2}\right)^{n/2}$ and $B := f_n(q) \geq 1$, we use the relation $A + B \leq 2\cdot 2^{n/2}AB$, obtaining

$$\begin{aligned}\sum_{\ell>\ell_0} c_\ell &\leq 2(q+1)8^{n/2}2^{n/2}f_n(q)\sum_{\ell>\ell_0} q^{-\ell}(\log N_\ell)^{n/2} + s\frac{q+1}{1-q^{-1}}q^{-\ell_0-1} \\ &\leq 2\frac{q+1}{1-q^{-1}}4^n f_n(q)\sum_{\ell>\ell_0}\int_{q^{-\ell-1}}^{q^{-\ell}}(\log N_\delta(\varepsilon))^{n/2}d\varepsilon + s\frac{q+1}{1-q^{-1}}q^{-\ell_0-1} \\ &\leq C_n M + 2C_n' s q^{-\ell_0-1}\end{aligned}$$

where $C_n = 2^{2n+1}(q+1)/(1-q^{-1})f_n(q)$, and $C_n' = \frac{1}{2}(q+1)/\left(1-q^{-1}\right)$, and we used the definition of M and the fact that $N_\delta(\varepsilon)$ is decreasing in ε.

We now notice that by definition of ℓ_0, there exists a point $t^* = s_{\ell_0}(t) \in I$ such that

$$q^{-\ell_0-1} \leq \max\{\delta(t,t^*) : t \in I\} \leq q^{-\ell_0}.$$

Therefore, we can bound $q^{-\ell_0-1}$ above by $\frac{D}{2}$ where $D = Diam_\delta(I)$, the diameter of I under δ, and conclude that

$$\mathbf{P}\left(\sup_{t\in I}|X_t - X_{t_0}| > C_n M + sC_n' D\right) \leq \mathbf{P}\left(\sup_{t\in I}|X_t - X_{t_0}| > \sum_{\ell>\ell_0} c_\ell\right) \leq 2e^{-s^{2/n}/2},$$

which ends the proof of the theorem, except for the statement of its last sentence.

To prove that last statement, now assume that X is not almost surely bounded, so that in particular I is infinite. Thus we have

$$\mathbf{P}[\forall N \in \mathbf{N}, \exists t \in I : |X_t| > N] = p > 0.$$

This implies

$$\mathbf{P}\left[\forall s > 0, \sup_t |X_t| > C_n M + sC_n' D\right] \geq p$$

which contradicts the result of the previous paragraph unless D is infinite or M is infinite. ■

The tail estimate of the previous theorem shows that the supremum of a sub-nth chaos process is a sub-nth chaos random variable, although the scale parameter in this property is not quite clear, since there is no explicit mention of the mean of $\sup X$. As a first step to make this scale clearer, we notice the following, which also allows a sharpening of the last statement of Theorem 3.1.

Corollary 3.2 *With n, X, I, δ, and M as in Theorem 3.1, there exist a universal constant C_n depending only on n, such that for each $t_0 \in I$ and $s \geq 1$,*

$$\mathbf{P}\left(\sup_{t\in I}|X_t - X_{t_0}| > sC_n M\right) \leq 2e^{-s^{2/n}/2}.$$

In particular, if M is finite, then X is almost-surely bounded.

The constant C_n is the same as in Theorem 3.1: it can be taken as $\min_{q>1} 2^{2n+1}(q+1)/\left(1-q^{-1}\right)f_n(q)$.

Proof. We reclaim the calculations in the proof of Theorem 3.1. The definition of c_ℓ remains the same in form, but instead of (8), we use

$$b_\ell = 8^{n/2}s[(\log N_\ell)^{n/2} + (\ell-\ell_0)^{n/2}].$$

Making use of the inequality $2AB \geq A+B$ for $A, B \geq 1$, we still obtain (9), and thus still also

$$\mathbf{P}\left(\sup_{t\in I}|X_t| > \sum_{\ell>\ell_0} c_\ell\right) \leq 2e^{-s^{2/n}/2}.$$

The calculation method for $\sum_{\ell>\ell_0} c_\ell$ is again identical, with the factor s appearing multiplicatively instead of additively:

$$\sum_{\ell>\ell_0} c_\ell \leq 2s\frac{q+1}{1-q^{-1}}4^n f_n(q)\, M = sC_nM$$

where $C_n = 2^{2n+1}(q+1)/(1-q^{-1})f_n(q)$. ■

An upper bound on the expected supremum, whose form is most reminiscent of what is usually called "Dudley's" inequality, was already proved in [17], using a distinct chaining argument from what we do above, but follows now trivially from the previous corollary.

Corollary 3.3 *Let n, X, I, C_n be as in Theorem 3.1 and assume X is centered. Then*

$$\mathbf{E}\sup_{t\in I} X_t \leq C_n'' \int_0^\infty (\log N_\delta(\varepsilon))^{n/2} d\varepsilon$$

where $C_n'' = C_n\left(1 + 2\int_1^\infty e^{-s^{2/n}/2}ds\right)$ depends only on n.

Proof. Using the same notation as in Theorem 3.1, and the fact that X is centered,

$$\begin{aligned}
\mathbf{E}\sup_{t\in I} X_t &= \mathbf{E}\sup_{t\in I}(X_t - X_{t_0}) \\
&\leq \mathbf{E}\sup_{t\in I}|X_t - X_{t_0}| \\
&= \int_0^\infty \mathbf{P}\left(\sup_{t\in I}|X_t - X_{t_0}| > u\right)du \\
&\leq C_nM + \int_{C_nM}^\infty \mathbf{P}\left(\sup_{t\in I}|X_t - X_{t_0}| > u\right)du \\
&= C_nM + C_nM\int_1^\infty \mathbf{P}\left(\sup_{t\in I}|X_t - X_{t_0}| > sC_nM\right)ds \\
&\leq C_nM + 2C_nM\int_1^\infty e^{-s^{2/n}/2}ds \\
&= C_n''M
\end{aligned}$$

■

We are now in a position to state and prove a full concentration inequality, of Dudley-Borell-Sudakov type, for $\sup X$. Throughout this article, we assume that the expected maximum in the statements of the Borell-Sudakov-type inequality (Theorems 3.4, 4.1, Corollary 3.5) is finite.

Theorem 3.4 *Let n, X, I, C_n, C_n', M, and D be as in Theorem 3.1, and C_n'' be as in Corollary 3.3. Let $\mu = \mathbf{E}\sup_{t\in I} X_t$. Then for all $u \geq 2(C_n + C_n'')M$,*

$$\mathbf{P}\left[\left|\sup_{t\in I} X_t - \mu\right| > u\right] \leq 2\exp\left(-\frac{1}{2}\left(\frac{u}{2C_n'D}\right)^{2/n}\right).$$

Note that the statement of the theorem can also be made to hold for all $u \in (\varepsilon, 2(C_n + C_n'')M]$ for any fixed $\varepsilon > 0$, as long as one is willing to change the constant C_n' above, allowing it to depend also on ε. It is preferable to use the form given above, however, in order to be able to use the sharper constant D in the tail estimate. Thus, up to a constant depending only on n, D appears as a scale with respect to which $\sup_{t\in I} X_t - \mu$ is a sub-nth chaos random variable.

Proof. By Theorem 3.1, for any $u \geq \mu$, defining s via $C_n M + sC_n' D = u - \mu$, we have

$$\begin{aligned}\mathbf{P}\left[\left|\sup_{t\in I} X_t - \mu\right| > u\right] &\leq \mathbf{P}\left[\sup_{t\in I} |X_t| + \mu > u\right] \\ &= \mathbf{P}\left[\sup_{t\in I} |X_t| > C_n M + sC_n' D\right] \\ &\leq 2e^{-s^{2/n}/2}.\end{aligned}$$

Now define $r = u/(KD)$ where K will be chosen below. If we impose $s \geq r$, since $rKD - \mu = C_n M + sC_n' D$, it follows that $rKD \geq C_n M + rC_n' D + \mu$. Thus choosing $K = 2C_n'$ we get

$$r \geq \frac{C_n M + \mu}{(K - C_n')D} = \frac{C_n M + \mu}{C_n' D}.$$

Hence, translating this into a condition on u, we have that if

$$u \geq \frac{C_n M + \mu}{C_n' D} KD = 2(C_n M + \mu)$$

so that indeed $s \geq r$. We immediately get

$$\mathbf{P}\left[\left|\sup_{t\in I} X_t - \mu\right| > u\right] \leq 2e^{-s^{2/n}/2} \leq 2e^{-r^{2/n}/2} = 2\exp\left(-\frac{1}{2}\left(\frac{u}{2C_n' D}\right)^{2/n}\right).$$

Since by Corollary 3.3, $\mu \leq C_n'' M$, the lower bound on u above holds as soon as $u \geq 2(C_n + C_n'')M$, which ends the proof of the theorem. ■

If instead we use the result from Corollary 3.2, following a similar argument above with $sC_n M = u - \mu, r = u/(KM), K = 2C_n$, we also get the following.

Corollary 3.5 *Let n, X, I, C_n, and M be as in Corollary 3.2, and C_n'' be as in Corollary 3.3. Let $\mu = \mathbf{E}\sup_{t\in I} X_t$. Then for all $u \geq 2C_n'' M$,*

$$\mathbf{P}\left[\left|\sup_{t\in I} X_t - \mu\right| > u\right] \leq 2\exp\left(-\frac{1}{2}\left(\frac{u}{2C_n M}\right)^{2/n}\right).$$

In this result, M seems to be a sub-nth chaos scale. The reader might wonder which of Theorem 3.4 and Corollary 3.5 is sharpest. The answer is trivial if it is possible to find a relation between D and M. Although this is not clear in a general sub-nth chaos case, we may expect that typically D should be smaller than M up to a constant, the reason being that D has to do with expectations of increments, with a supremum outside the expectation, whereas M is an upper bound on an expectation with a supremum inside. This is very clear in the centered Gaussian case, as the following calculation shows. In calculating D, by writing $X_t - X_s = (X_t - X_{t_0}) - (X_s - X_{t_0})$ we can assume there exists t_0 such that $X_{t_0} = 0$. Thence

$$\begin{aligned}D^2 &= \sup_{s,t\in I} \mathbf{E}\left[(X_t - X_s)^2\right] \leq \sup_{s,t\in I} \mathbf{E}\left[2X_t^2 + 2X_s^2\right] = 4\sup_{t\in I} \mathbf{E}\left[X_t^2\right] \\ &= 2\pi\left(\sup_{t\in I} \mathbf{E}\left[|X_t|\right]\right)^2 \leq 2\pi\left(\mathbf{E}\left[\sup_{t\in I} |X_t|\right]\right)^2 \\ &\leq 2\pi\left(\mathbf{E}\,|X_{t_0}| + 2\mathbf{E}\left[\sup_{t\in I} X_t\right]\right)^2 = 8\pi\mathbf{E}^2\left[\sup_{t\in I} X_t\right] \leq 8\pi\left(C_1'' M\right)^2,\end{aligned}$$

by Corollary 3.3, so that $D \leq \sqrt{8\pi} C_1'' M$.

4 Iterated Malliavin derivatives method

The use of boundedness of (iterated) Malliavin derivatives as conditions for the nth chaos property was first introduced in [17].

For purposes of comparison, we cite the basic result in the sub-Gaussian scale. Let X be an almost-surely bounded and centered random field on an index set I, and assume that for every $x \in I$, $X(x)$ is a member of the Malliavin-Sobolev space $\mathbf{D}^{1,2}$ (see definition below). Assume for each $x \in I$ there exists a non-random value $\sigma(x)$ such that

$$|D.X(x)|^2_{L^2(dr)} := \int_0^\infty |D_r X(x)|^2 \, dr \leq \sigma^2(x).$$

Then, as in the classical Dudley-Borell-Sudakov concentration inequality, with $\sigma^2 = \sup_{x\in I} \sigma^2(x)$, we have

$$\mathbf{P}\left[\left|\sup_{x\in I} X(x) - \mathbf{E}\sup_{x\in I} X(x)\right| > u\right] \leq 2\exp\left(-\frac{u^2}{2\sigma^2}\right). \tag{10}$$

The idea here is that a random variable with a Malliavin derivative that is almost-surely bounded in $L^2(dr)$ by a non-random constant has the sub-Gaussian property with respect to the scale of that non-random constant, and at the process level, we get sub-Gaussian concentration of a process's supremum relative to the smallest possible scale, namely the max of the process's one-dimensional distributions' scales. In this section, we generalize the above result to the sub-nth chaos level, using n-thly iterated Malliavin derivatives. This idea was already used on [17], although the results therein were not optimally stated or proved, an issue we correct here. The following review of Malliavin derivatives' properties will be useful to the reader.

Let W be a Brownian motion in the filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbf{P}\right)$, where $\mathcal{F}$ is the sigma-field generated by the process W. For any centered Gaussian random variable $X \in L^2(\Omega, \mathcal{F}, \mathbf{P}))$, we have that $X = \int_0^\infty f(s)\, dW(s)$ for some non-random $f \in L^2(\mathbf{R}_+)$. The Malliavin derivative of X at time $s \in \mathbf{R}_+$ is defined a.e. as $D_s X := f(s)$. For any function Φ on $\mathbf{R}$ that is continuously differentiable a.e. and such that $Y := \Phi(X) \in L^2(\Omega, \mathcal{F}, \mathbf{P})$, the Malliavin derivative of Y is defined via the chain rule as $D_s Y = \Phi'(X) D_s X = \Phi'(X) f(s)$, $\mathbf{P} \times dr$-a.e. This chain rule also applies to functions of a Gaussian vector in $\mathcal{F}$ in the obvious way. The Malliavin derivative operator is extended as a closed operator to the set $\mathbf{D}^{1,2}$ of random variables whose Malliavin derivatives are in $L^2(\Omega \times \mathbf{R}_+)$. See details of the extension in [11]. The nth Malliavin derivative operator $D^{(n)}$ is defined by iterating D n-fold, and thus depends on n time parameters, while still acting on a single r.v. Thus $\mathbf{D}^{n,2}$ is the set of all random variables X in $\mathcal{F}$ such that $D^{(n)}X \in L^2(\Omega \times (\mathbf{R}_+)^n)$. Lastly, a word on chaos expansions (see [11] or [17] for details). Every random variable in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ can be expanded as $X = \mathbf{E}[X] + \sum_{k=1}^\infty I_k(f_k)$ where I_k is the iterated Wiener-Ito integral with respect to W, and f_k is a symmetric function in $L^2\left((\mathbf{R}_+)^k\right)$. Elementary work with Hermite polynomials shows that $D_s I_k(f_k) = k I_{k-1}(f_k(\cdot, x))$. In particular, the Malliavin derivative of a non-random quantity is 0, the Malliavin derivative of a Gaussian r.v. is non-random, the Malliavin derivative of a 2nd-chaos random variable is Gaussian, etc..., and the nth Malliavin derivative of $I_n(f_n)$ is $n! f_n$, which implies that any condition on the nth Malliavin derivative of X ignores its terms in the chaoses of orders $0, 1, \cdots, n-1$.

We have the following supremum concentration inequality.

Theorem 4.1 *Assume a separable almost-surely bounded random field X on an index set I is such that for all $x \in I$, $X(x) \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ and for some integer n, for all $x \in I$, $D^{(n)}X(x)$ is almost-surely bounded in $L^2(dr)$ by a non-random constant. More specifically define*

$$\sigma^2(x) := ess\ sup \left\{ \int_{(\mathbf{R}_+)^n} \left| D^{(n)}_{r_1,\cdots,r_n} X(x) \right|^2 dr_1 \cdots dr_n \right\} \tag{11}$$

and assume $\sigma := \sup_{x\in I}\sigma(x) < \infty$. *Also assume that pairs of values of* X *are a.s. distinct. Then there is a constant* C_n *depending only on* n *such that for any* $\varepsilon > 0$, *for* u *large enough,*

$$\mathbf{P}\left[\left|\sup_{x\in I} X(x) - \mathbf{E}\sup_{x\in I} X(x)\right| > u\right] \le 2(1+\varepsilon)\exp\left(-\frac{1}{2(1+\varepsilon)}\left(\frac{u}{C_n\sigma}\right)^{2/n}\right). \tag{12}$$

In fact, $\mathbf{P}$*-almost surely,*

$$\int_{(\mathbf{R}_+)^n}\left|D^{(n)}_{r_1,\cdots,r_n}\sup_{x\in I} X(x)\right|^2 dr_1\cdots dr_n \le \sigma^2. \tag{13}$$

In order to prove this theorem, we need the following consequence of a fractional exponential Poincaré inequality, whose proof is relegated to the Appendix.

Proposition 4.2 *Let* Y *be a centered random variable in* $L^2(\Omega)$ *of the form* $Y = \sum_{k=n}^{\infty} I_k(f_k)$, *satisfying almost surely*

$$\int_{(\mathbf{R}_+)^n} d\bar{s}\left|D^{(n)}_{\bar{s}}Y\right|^2 \le M^2$$

for some non-random value M. *Then there exists a constant* C_n *depending only on* n *such that* Y *is a sub-nth chaos r.v. relative to the scale* C_nM, *i.e.*

$$\mathbf{E}\left[\exp\left(\left|\frac{Y}{C_nM}\right|^{2/n}\right)\right] \le 2.$$

The constant C_n *can be taken as* $C_n = (\log K(n)/\log 2)^{n/2}$ *where* $K(n)$ *is defined and estimated in (27), (28).*

Proof of Theorem 4.1. *Step 1: Setup.* We can assume without loss of generality that our index set is finite: $I = I_N = \{1, 2, \cdots N\}$. Indeed, by separability, $\sup_{x\in I} X(x)$ can be evaluated by replacing I by a countable subset. Recall that $\mu := \mathbf{E}\sup_{x\in I} X(x)$ is finite, and thus the random variable $\sup_{x\in I} X(x) - \mu$ is well-defined, and is the almost sure limit of the same r.v. with I replaced by a sequence I_N of increasing sets of size N converging to I; the dominated convergence theorem yields the conclusion of the theorem since its estimates do not depend on N.

Step 2: Proof of (13). 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 $\max_I X \in \mathbf{D}^{n,2}$, an approximation technique can be used (see the proof of [17, Theorem 3.6]): one shows that $\mathbf{1}_{X_{m+1}>Sm}$ can be approximated in $\mathbf{D}^{1,2}$ by a smooth function of $X_{m+1}-S_m$ 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 $k \le n$, the kth-order Malliavin derivative of $\mathbf{1}_{X_{m+1}>Sm}$ is 0 in $L^2\left(\Omega\times(R_+)^k\right)$ as well.

This justifies the following computation, where equalities hold in $L^2(\Omega\times(R_+)^n)$:

$$\begin{aligned} D^{(n)}_{s_n,\cdots,s_2,s_1}S_{m+1} &= D^{(n-1)}_{s_n,\cdots,s_2}\left(D_{s_1}X_{m+1}\mathbf{1}_{X_{m+1}>Sm} + D_{s_1}S_m\mathbf{1}_{X_{m+1}<Sm}\right) \\ &= D^{(n-2)}_{s_n,\cdots,s_3}\left(\left[D_{s_2}D_{s_1}X_{m+1}\right]\mathbf{1}_{X_{m+1}>Sm} + \left[D_{s_2}D_{s_1}S_m\right]\mathbf{1}_{X_{m+1}<Sm}\right) \\ &\vdots \\ &= \left[D^{(n)}_{s_n,\cdots,s_2,s_1}X_{m+1}\right]\mathbf{1}_{X_{m+1}>Sm} + \left[D^{(n)}_{s_n,\cdots,s_2,s_1}S_m\right]\mathbf{1}_{X_{m+1}<Sm}. \end{aligned} \tag{14}$$

Now let $\sigma^{*2}_m = \max\left\{\sigma^2(1);\cdots;\sigma^2(m)\right\}$. Proceeding by induction, assume that $\left\|D^{(n)}_\cdot S_m\right\|_2^2 \le \sigma^{*2}_m$ almost surely, which is satisfied for $m = 1$ by our hypothesis on $D^{(n)}_\cdot X(1)$ since it is assumed to be bounded in

$L^2\left(\left(\mathbf{R}_+\right)^n\right)$ by $\sigma\left(1\right)=\sigma_1^*$. Our hypothesis and equality (14) implies that almost surely

$$\begin{aligned}
&\int_{(\mathbf{R}_+)^n}\left|D^{(n)}_{s_n,\cdots,s_2,s_1}S_{m+1}\right|^2 ds_n\cdots ds_2ds_1\\
&=\mathbf{1}_{X_{m+1}>Sm}\int_{(\mathbf{R}_+)^n}\left|D^{(n)}_{s_n,\cdots,s_2,s_1}X_{m+1}\right|^2 ds_n\cdots ds_2ds_1\\
&+\mathbf{1}_{X_{m+1}<Sm}\int_{(\mathbf{R}_+)^n}\left|D^{(n)}_{s_n,\cdots,s_2,s_1}S_m\right|^2 ds_n\cdots ds_2ds_1\\
&\leq\sigma^2\left(m+1\right)\mathbf{1}_{X_{m+1}>Sm}+\sigma_m^{*2}\mathbf{1}_{X_{m+1}<Sm}\\
&\leq\sigma_{m+1}^{*2},
\end{aligned}$$

induction implies (13) when $m=N$.

Step 3. Translation into a concentration inequality. We can write the chaos decomposition of $\sup_{x\in I}X\left(x\right)$ as $\sum_{k=0}^{\infty}I_k\left(f_k\right)$. Since the nth Malliavin derivative kills off the first n terms, we direct our attention to the random variable

$$Y=\sum_{k=n}^{\infty}I_k\left(f_k\right).$$

Statement (13) along with Proposition 4.2 , with $M=\sigma$, and Chebyshev's inequality, implies the conclusion (12) of the theorem, with $\varepsilon=0$ but with $\sup_{x\in I}X\left(x\right)-\mathbf{E}\sup_{x\in I}X\left(x\right)=\sum_{k=1}^{\infty}I_k\left(f_k\right)$ replaced by Y. The technique used in [17] to recuperate all the terms in the chaos expansion of $\sup_{x\in I}X\left(x\right)$, at the cost of adding an $\varepsilon>0$ in the statement (12), can now be invoked, to conclude the proof of the theorem (see the proof of Corollary 4.14, and especially the result of Lemma 4.15 on the tails of purely kth chaos random variables, in [17], which is a consequence of results sometimes attributed to Ch. Borell, found for instance in [7], see also [2]; all details are omitted). ■

One may wish to compare the Borell-Sudakov inequalities from Theorem 3.4, and from Theorem 4.1, since they provide sub-nth chaos concentrations with respect to two distinct scales D and σ.

Before discussing the differences between these scales, let us note that the hypotheses of the two theorems are close, but not actually comparable. Theorem 4.1 requires less than Theorem 3.4 in terms of joint distribution, since Theorem 3.4 needs X to be a sub-nth chaos process, while Theorem 4.1 only needs each random variable $X\left(x\right)$ to have the sub-nth chaos property. On the other hand, Theorem 3.4 needs only the basic sub-nth chaos property as in our original definitions, while Theorem 4.1 requires a bit more, since condition (11) implies this property via Proposition 4.2, whose converse is not known.

Having said this, let us now show that D and σ are not necessarily comparable, and explain in what cases they are. In order to make any meaningful comparisons, it is necessary to believe that while the boundedness of the nth Malliavin derivative in Proposition 4.2 is perhaps not necessary for the sub-nth chaos property, it is still presumably extremely close. We will not argue in favor or against this belief here, but assuming one has it, one immediately concludes that if X is a sub-nth chaos field on I, the "nearly" best choice for its scale metric δ is

$$\begin{aligned}
\delta\left(s,t\right)&=\text{ess sup}\left\{\int_{(\mathbf{R}_+)^n}\left|D^{(n)}_{r_1,\cdots,r_n}\left(X\left(t\right)-X\left(s\right)\right)\right|^2 dr_1\cdots dr_n\right\}^{1/2}\\
&=\text{ess sup}\left|D^{(n)}\left(X\left(t\right)-X\left(s\right)\right)\right|_{L^2\left(\mathbf{R}_+^n\right)}.
\end{aligned}$$

We can first prove $D \leq 2\sigma$ as follows:

$$\begin{aligned} D &= \sup_{s,t \in I} \delta(s,t) \\ &\leq \sup_{s,t \in I} \text{ess sup} \left(\left| D^{(n)}(X(t)) \right|_{L^2(\mathbf{R}_+^n)} + \left| D^{(n)}(X(s)) \right|_{L^2(\mathbf{R}_+^n)} \right) \\ &= 2 \sup_{t \in I} \text{ess sup} \left| D^{(n)}(X(t)) \right|_{L^2(\mathbf{R}_+^n)} \\ &= 2\sigma. \end{aligned}$$

An opposite inequality, cannot hold in general, but we can easily prove $\sigma \leq D$ assuming for instance that for some t_0, $X(t_0) = 0$. Indeed, we can then write

$$\begin{aligned} \sigma &= \sup_{t \in I} \text{ess sup} \left| D^{(n)}(X(t) - X(t_0)) \right|_{L^2(\mathbf{R}_+^n)} \\ &\leq \sup_{s,t \in I} \text{ess sup} \left| D^{(n)}(X(t) - X(s)) \right|_{L^2(\mathbf{R}_+^n)} \\ &= \sup_{s,t \in I} \delta(s,t) = D. \end{aligned}$$

To finish this discussion, let us give an example where σ can be much larger than D. Consider the last example for X, and define $Y(t) = X(t) + Z$ where Z is an nth-chaos random variable independent of X, with an nth chaos norm $\left| D^{(n)} Z \right|_{L^2(\mathbf{R}_+^n)} = z$, a given non-random constant. Since X and Z are assumed to be independent, their Malliavin derivatives are supported on disjoint subsets of $\mathbf{R}_+^n$. Thus we have for Y that

$$\begin{aligned} \sigma_Y^2 &= \sup_{t \in I} \text{ess sup} \left| D^{(n)}(X(t)) + D^{(n)} Z \right|_{L^2(\mathbf{R}_+^n)}^2 \\ &= \sup_{t \in I} \text{ess sup} \left| D^{(n)}(X(t)) \right|_{L^2(\mathbf{R}_+^n)}^2 + \left| D^{(n)} Z \right|_{L^2(\mathbf{R}_+^n)}^2 \\ &= \sigma_X^2 + z^2. \end{aligned}$$

On the other hand, we obviously have that X and Y share the same diameter D since they share the same scale metric δ due to having the same increments. Thus by choosing z arbitrarily large, D can be made arbitrarily small compared to σ_Y. Hence Theorem 4.1 can sometimes be much less sharp than Theorem 3.4. To avoid such a situation, one may redefine σ in Theorem 4.1 by considering only the oscillation function of X. We will not delve deeper into this issue except to say that no statement of the Borel-Sudakov inequality we have seen in the literature has noted that σ may not be sharp; even in the Gaussian case ($n = 1$), where $\sigma^2 = \max_{t \in I} Var[X(t)]$, this lack of sharpness may occur as described above.

5 Modulus of Continuity

Definition 5.1 *Let f be a continuous increasing function in $\mathbf{R}_+$ such that* $\lim_{0+} f = 0$. *Let $\{Y(t) : t \in I\}$ be a random field on an index set I endowed with a metric ρ. We say that f is an almost sure uniform modulus of continuity for Y on (I, ρ) if there exists an almost-surely positive random variable α_0 such that*

$$\alpha < \alpha_0 \Rightarrow \sup_{s,t \in I : \rho(s,t) < \alpha} |Y(t) - Y(s)| \leq f(\alpha).$$

Corollary 3.3 has an immediate consequence for continuity (see [1]). Consider the random field $Y_{u,v} = X_v - X_u$. Let δ be the canonical metric for X on T defined as $\delta(u,v) = (\mathbf{E}(X(v) - X(u))^2)^{1/2}$ and let δ_Y

be the canonical metric for Y on $T \times T$. Then

$$\delta_Y((u,v),(u',v')) = \left[\mathbf{E}\left((X_v - X_u) - (X_{v'} - X_{u'})\right)^2\right]^{1/2}$$
$$\leq 2\max\left(\delta(u,v), \delta(u',v')\right).$$

This implies that $N_{\delta_Y}(\varepsilon) \leq N_\delta(\varepsilon/2)$. So, Corollary 3.3 implies

$$\mathbf{E}\left[\sup_{(u,v)\in T\times T,\ \delta(u,v)\leq\eta} |X_v - X_u|\right] = \mathbf{E}\left[\sup_{(u,v)\in T\times T,\ \delta(u,v)\leq\eta} |Y_{u,v}|\right]$$
$$\leq C''_n \int_0^{2\eta} (\log N_\delta(\varepsilon/2))^{n/2} d\varepsilon$$
$$= 2C''_n \int_0^{\eta} (\log N_\delta(\varepsilon))^{n/2} d\varepsilon.$$

Thus, we have proved the following result.

Proposition 5.2 *Let X be as in Corollary 3.3. Then*

$$\mathbf{E}\left[\sup_{\delta(u,v)\leq\eta} |X_v - X_u|\right] \leq K_n \int_0^{\eta} (\log N_\delta(\varepsilon))^{n/2} d\varepsilon$$

where $K_n = 2C''_n$ with C''_n the constant depending only on n defined in Corollary 3.3.

A well-known phenomenon in the theory of Gaussian regularity is that a modulus of continuity of X relative to δ is also given almost surely as the right hand side of the previous inequality. The same effect is shown here for sub-nth chaos processes, based on a chaining argument similar to that used to prove Theorem 3.1. In our attempt to derive the next result as a mere consequence of Theorem 3.1, we found that the number of modifications needing to be made to the proof of the latter in order to get the next theorem to work, warranted a whole new proof. It is given below.

Theorem 5.3 *Let X be as in Theorem 3.1. There exists a random number $\eta_0 > 0$ almost surely, such that*

$$\sup_{\delta(u,v)\leq\eta} |X_v - X_u| \leq k_n \int_0^{\eta} (\log N_\delta(\varepsilon))^{n/2} d\varepsilon + 4k_n g(\eta),$$

for all $\eta < \eta_0$, where k_n can be taken as $2^{n+1}(q+1)$ with $q > 1$, where $g(\eta) := \eta \log^{n/2}(\log(1/\eta))$.

Before proving this theorem, we discuss its scope and a corollary in Euclidean space.

Remark 5.4 *There seems to be a slight inefficiency in the above result, due to the presence of the term $g(\eta) = \eta \log^{n/2}\log(1/\eta)$: the theorem cannot differentiate between sub-nth chaos random fields which are a.s Lipshitz continuous relative to δ, and those who have $g(\eta)$ as an a.s. δ-modulus of continuity.*

Therefore this theorem, as stated, does not provide a sharp result for trivial examples such as $X(t) = \sum_{i=1}^{M} f_i(t) G_i$ where $(G_i)_{i=1}^{M}$ is a finite-dimensional sub-nth chaos random vector, and the f_i's are smooth vector fields. We leave it to the reader to check that the proof below can be modified to allow for a sharp result in this case, but of course the result is trivial in that situation.

There may be other less trivial examples of nearly Lipshitz random fields for which the theorem is not sharp, but we suspect these are always pathological, i.e. are highly inhomogeneous. For the vast majority of random fields, including all those encountered in the literature, the entropy integral in the above theorem will dominate $g(\eta)$. The next result shows that in typical examples in Euclidean space, the extra term is not needed.

The corollary below on moduli of continuity can also be established using the so-called Garsia-Rodemich-Rumsey real variable lemma, however, we prefer the current presentation based fully on probabilistic chaining arguments, to avoid seeking techniques outside of probability theory.

Corollary 5.5 *Let X be a separable sub-nth chaos field relative to the pseudo-metric δ on $E \times E$, where E is a subset of the d-dimensional Euclidean space, with Euclidean norm denoted by $|\cdot|$. Assume there is an increasing univariate function on $\mathbf{R}_+$, also denoted by δ, such that the right-hand derivative of δ at 0 exists (possibly equal to $+\infty$), satisfying $\delta(s,t) \leq \delta(|s-t|)$ and the condition*

$$\lim_{r \to 0^+} \delta(r) \left(\log \frac{1}{r}\right)^{n/2} = 0. \tag{15}$$

Then, up to a non-random constant c, $f_\delta(r) := \delta(r)\left(\log r^{-1}\right)^{n/2}$ is an almost-sure uniform modulus of continuity for X on any compact subset of E.

The constant c can be taken as any constant exceeding $4k_n d^{n/2}$ where k_n is as in Theorem 5.3, and therefore does not depend on the distribution of X.

Proof of Theorem 5.3. Following the same construction as in the proof of Theorem 3.1, we get that almost surely for every $u, v \in T$ and $\ell' > \ell_0$,

$$X_v - X_u = X_{s_{\ell'}(v)} - X_{s_{\ell'}(u)} + \sum_{\ell > \ell'} \left(X_{s_\ell(v)} - X_{s_{\ell-1}(v)}\right) - \sum_{\ell > \ell'} \left(X_{s_\ell(u)} - X_{s_{\ell-1}(u)}\right). \tag{16}$$

From the first inequality in Lemma 4.6 of [17], we have that for $c > 0$,

$$\mathbf{P}\left(|X_v - X_u| > c\delta(u,v)\right) \leq 2\exp\left(-c^{2/n}\right).$$

Applying the above inequality, we have

$$\begin{aligned}
&\mathbf{P}\left(\exists\, u, v \in T : \left|X_{s_\ell(v)} - X_{s_\ell(u)}\right| > \delta(s_\ell(u), s_\ell(v))\left(\log\left(\ell^2 N_\ell^2\right)\right)^{n/2}\right) \\
&\leq 2N_\ell^2 \exp\left(-\log\left(\ell^2 N_\ell^2\right)\right) \leq 2\ell^{-2}
\end{aligned}$$

where $N_\ell = N_\delta(q^{-\ell})$. Since this is a summable series, by the Borel-Cantelli lemma, there exists a random almost-surely finite integer $L_{0,1} > \ell_0$ such that with probability one,

$$\ell \geq L_{0,1} \Rightarrow \left|X_{s_\ell(v)} - X_{s_\ell(u)}\right| \leq \delta(s_\ell(u), s_\ell(v))\left(\log\left(\ell^2 N_\ell^2\right)\right)^{n/2}$$

for all $u, v \in T$. Similarly, we also get that

$$\ell \geq L_{0,2} \Rightarrow \left|X_{s_\ell(u)} - X_{s_{\ell-1}(u)}\right| \leq \delta(s_\ell(u), s_{\ell-1}(u))\left(\log\left(\ell^2 N_\ell N_{\ell-1}\right)\right)^{n/2}$$

for all $u \in T$, for a possibly larger random integer $L_{0,2}$, so that we denote $L_0 = \max(L_{0,1}, L_{0,2})$.

Putting these into (16) in which we replace ℓ' by some $L \geq L_0$, we get that

$$\begin{aligned}
|X_v - X_u| \leq\ & \delta(s_L(u), s_L(v))\left(\log\left(L^2 N_L^2\right)\right)^{n/2} + \sum_{\ell > L} \delta(s_\ell(u), s_{\ell-1}(u))\left(\log\left(\ell^2 N_\ell N_{\ell-1}\right)\right)^{n/2} \\
&+ \sum_{\ell > L} \delta(s_\ell(v), s_{\ell-1}(v))\left(\log\left(\ell^2 N_\ell N_{\ell-1}\right)\right)^{n/2}.
\end{aligned}$$

Now define $\eta_0 := q^{-L_0}$ and for any $\eta \leq \eta_0$, let L be the unique integer such that $q^{-L-1} < \eta \leq q^{-L}$. Hence for any pair (u,v) such that $\delta(u,v) \leq \eta$,

$$\begin{aligned}
\delta(s_L(u), s_L(v)) &\leq \delta(s_L(u), u) + \delta(u,v) + \delta(v, s_L(v)) \\
&\leq q^{-L} + \eta + q^{-L} \\
&\leq 3q^{-L}.
\end{aligned}$$

Using this, (3), and the fact that $N_{\ell-1} \leq N_\ell$ for any $\ell > \ell_0$, we get that for the above pair (u, v),

$$\begin{aligned}
|X_v - X_u| &\leq 3q^{-L} \left(\log\left(L^2 N_L^2\right)\right)^{n/2} + 2\sum_{\ell > L}(q+1)q^{-\ell}\left(\log\left(\ell^2 N_\ell N_{\ell-1}\right)\right)^{n/2} \\
&\leq 2\,(q+1)\sum_{\ell \geq L} q^{-\ell}\left(\log\left(\ell^2 N_\ell^2\right)\right)^{n/2} \\
&\leq k_n\left[F(L) + I(L)\right] \qquad (17)
\end{aligned}$$

where

$$k_n := 2^{n+1}(q+1),$$

and

$$F(L) := \sum_{\ell \geq L} q^{-\ell} \log^{n/2} \ell,$$

and also

$$I(L) := \sum_{\ell \geq L} q^{-\ell} (\log N_\ell)^{n/2}.$$

Note that for $\varepsilon \in (q^{-\ell-1}, q^{-\ell}]$, $N_\ell \leq N_\delta(\varepsilon)$ and $\ell \leq \log(1/\varepsilon)/\log q$. Therefore, approximating the series $I(L)$ and $F(L)$ by Riemann integrals, we get

$$I(L) \leq \int_0^\eta \log^{n/2}(N_\delta(\varepsilon))\, d\varepsilon \qquad (18)$$

and

$$F(L) \leq \int_0^\eta \log^{n/2}\left(\frac{\log(1/\varepsilon)}{\log q}\right) d\varepsilon \leq 2\int_0^\eta \log^{n/2}(\log(1/\varepsilon))\, d\varepsilon \qquad (19)$$

where in the last inequality, we used the fact that $\eta_0 = q^{-L_0}$ so that $\log(\log(1/\varepsilon)) \geq \log(\log q)$. Putting (17), (19), and (18) together, we finally get that for any $\eta \leq \eta_0$, for any pair (u, v) such that $\delta(u, v) \leq \eta$,

$$\begin{aligned}
|X_v - X_u| &\leq k_n \int_0^\eta \log^{n/2}(N_\delta(\varepsilon))\, d\varepsilon + 2k_n \int_0^\eta \log^{n/2}(\log(1/\varepsilon))\, d\varepsilon \\
&\leq k_n \int_0^\eta \log^{n/2}(N_\delta(\varepsilon))\, d\varepsilon + 4k_n \eta \log^{n/2}(\log(1/\eta)),
\end{aligned}$$

which finishes the proof of Theorem 5.3. ■

Proof of Corollary 5.5. Since we only need to evaluate a modulus of continuity on compact subsets of E up to undetermined non-random constants, it is sufficient to assume that X is defined on a subset of the box $x_0 + [-M, M]^d$ where M can be as small as needed and the center x_0 of this box is arbitrary; one can then obviously tile any compact subset of E with such boxes defined by M fixed and a collection of centers x_0. Thus, without loss of generality, we assume $E = [-M, M]^d$ where we may choose $M > 0$ as small as desired. In view of the homogeneous upper bound $\delta(s, t) \leq \delta(|s - t|)$ in the corollary's assumptions, we may indeed use $x_0 = 0$ without loss of generality.

From Theorem 5.3, there exists an almost-surely positive random number η_0 and a universal constant k_n depending only on n such that

$$\sup_{\delta(u,v) \leq \eta} |X_v - X_u| \leq k_n \int_0^\eta (\log N_\delta(\varepsilon))^{n/2} d\varepsilon + 4k_n g(\delta(h)),$$

for all $\eta < \eta_0$, where $N_\delta(\varepsilon)$ is the smallest number of δ-balls of radius ε needed to cover E. Denote by δ^{-1} the inverse function of the univariate δ. Since the set $\{|u - v| \leq \delta^{-1}(\eta)\}$ is smaller than the set $\{\delta(u, v) \leq \eta\}$, denoting $h = \delta^{-1}(\eta)$, the inequality above implies

$$\sup_{|u-v| \leq h} |X_v - X_u| \leq k_n \int_0^{\delta(h)} (\log N_\delta(\varepsilon))^{n/2} d\varepsilon + 4k_n g(\delta(h)), \qquad (20)$$

for all $h < \delta^{-1}(\eta_0)$.

The set $E = [-M, M]^d$ may be covered with balls of δ-radius ε by using instead Euclidean balls of radius $r := \delta^{-1}(\varepsilon)$ centered at the points of $(r\mathbf{Z})^d \cap E$. This is done with $(2M/r)^d$ Euclidean balls if $2M/r$ is an integer, which proves

$$N_\delta(\varepsilon) \le \left(\frac{2M}{\delta^{-1}(\varepsilon)} + 1\right)^d.$$

Now choose $M = 1/4$. By choosing η small enough, we can make $h = \delta^{-1}(\eta)$ as small as we wish, since $\lim_0 \delta^{-1} = 0$; thus, also using the fact that δ^{-1} is increasing, choose η so that $\delta^{-1}(\varepsilon) \le \delta^{-1}(\eta) \le 1/2$. This guarantees that

$$N_\delta(\varepsilon) \le \frac{1}{(\delta^{-1}(\varepsilon))^d}.$$

From (20), this proves that for some almost surely positive h_0, for $h < h_0$,

$$\sup_{|u-v|\le h} |X_v - X_u| \le k_n d^{n/2} \int_0^{\delta(h)} \left(\log \frac{1}{\delta^{-1}(\varepsilon)}\right)^{n/2} d\varepsilon + 4k_n g(\delta(h)). \tag{21}$$

We may now use a change of measure in the last integral above in the Riemann-Stieltjes sense, to transform the almost-sure uniform modulus of continuity for X we have just obtained in (21):

$$I_\delta(h) := \int_0^{\delta(h)} \left(\log \frac{1}{\delta^{-1}(\varepsilon)}\right)^{n/2} d\varepsilon = \int_0^h \left(\log \frac{1}{r}\right)^{n/2} d\delta(r).$$

By an integration by parts, we may now write

$$I_\delta(h) = \delta(h)\left(\log \frac{1}{h}\right)^{n/2} + n \int_0^h \frac{1}{2r}\left(\log \frac{1}{r}\right)^{n/2-1} \delta(r)dr.$$

The last term above is positive; moreover, by using the estimate $\delta(r) \le \delta(h)$ therein, we get that this term is bounded above by $\delta(h)\left(\log h^{-1}\right)^{n/2}$. Hence

$$I_\delta(h) \le 2\delta(h)\left(\log \frac{1}{h}\right)^{n/2} = 2f_\delta(h). \tag{22}$$

Combining (21) and (22), with $c' = 4k_n d^{n/2}$, we obtain

$$\sup_{|u-v|\le h} |X_v - X_u| \le c'\left(f_\delta(h) + g(\delta(h))\right). \tag{23}$$

To complete the proof of the Corollary, it is sufficient to show that $g(\delta(h)) = o(f_\delta(h))$. If we assume that $\delta(h)/h$ is bounded below by a positive constant, this follows from a straightforward calculation. This assumption is satisfied for all usual examples of regularity scales because there δ is concave, and therefore $\delta(h)/h$ increases to $\delta'(0)$ as $h \to 0$. Moreover it is not a restriction in any case. One can show that any random field for which $\delta(t, t+h)/h$ tends to 0 as $h \to 0$ uniformly for all t in an open set of the Euclidean space is actually constant on that set. If a random field satisfying the hypotheses of the corollary also had $\lim_{h\to 0} \delta(h)/h = \delta'(0) = 0$, then we would have, uniformly in t, $\lim_{h\to 0} \delta(t, t+h)/h = 0$, and the conclusion of the Corollary would hold trivially for this constant random field [even though it would not be sharp]. Let now $c > c'$. By letting h be so small that $g(\delta(h)) < ((c-c')/c')f_\delta(h)$, (23) finishes the proof of the Corollary. ■

The end of the above proof points to an inefficiency in the result, which also appears in all works in the literature which deal with almost-sure uniform moduli of continuity. Consider the example $X(t) = tW(t)$ where W is a standard Brownian motion. In this case, for all $s, t \in [0, 1]$, we may write that this Gaussian process has a canonical metric bounded as $\delta(s,t) \leq K|t-s|^{1/2}$ where K is a universal constant. Thus $f_\delta(h) = h^{1/2}\log^{1/2}(1/h)$ is an almost sure uniform modulus of continuity for X up to a universal constant, which is of course a sharp result, since lower bounds of the same order can also be obtained in this very simple case when looking at the entire interval $[0,1]$. However, this hides the fact that at $t=0$, X is almost-surely differentiable (indeed, its derivative is 0 almost surely!). Because of the extreme inhomogeneity of X near the origin, no statement on uniform moduli of continuity can accurately describe X's behavior there.

The result of Corollary 5.5 is nevertheless quite sharp, in the following sense. First note that the integral $I_\delta(h)$ in (21) is also bounded below by $f_\delta(h)$. Thus there is no hope to better the result of the corollary analytically. More specifically, one can show that if $\delta(u,v) = \delta(|u-v|)$, a *homogeneous* case, then $f_\delta(h)$ is in fact, an upper and lower bound for the entropy integral in (20), with constants 2 and 1, respectively; therefore in this case, $f_\delta(h)$ appears indeed as a sharp estimate of the modulus of continuity of X. Referring to the homogeneous Gaussian case for comparison, it was proved in [15] that f_δ is a uniform modulus of continuity if and only if the canonical metric of X is commensurate with δ. Our corollary above is a first step in proving the same result for sub-nth processes. To prove such a strong result for these processes, however, one would need a strong lower bound assumption on the tails of the processes' increments. Even in the sub-Gaussian case, such a problem is open.

We close this article with a note on the scope of Condition (15). In the Gaussian case, it is well understood that this condition is sufficient for almost-sure boundedness and continuity, while in the corresponding homogeneous case with scalar parameter, when the condition is not satisfied, one can prove that the process is almost surely unbounded on any open interval; this is a consequence of the so-called Fernique lower bound (e.g. opposite inequality from Corollary 3.3 with $n=1$). When $n>1$, it is simple enough to conjecture that such sharpness should also hold; this would be easier to prove than the program outlined in the previous paragraph, but it is still not a straightforward endeavor, since even in the sub-Gaussian case, formulating a lower bound hypothesis sufficient to obtain an analogue Fernique's lower bound appears to be non-trivial. We may simply state at this stage that the following regularity scales all yield continuous sub-nth chaos processes: Condition (15) is satisfied by the following scales (c is a generic constant which depends on n and the other regularity parameters H, β, γ given below):

1. $\delta(r) = r^H$, $H \in (0,1)$ (fractional Brownian scale): $f_\delta(r) \leq cr^H \log^{n/2}(1/r)$;
2. $\delta(r) = \left(\log \frac{1}{r}\right)^{-\beta}$, $\beta > n/2$ (logarithmic Brownian scale): $f_\delta(r) \leq c \log^{-\beta+n/2}(1/r)$;
3. $\delta(r) = r\log^\gamma \frac{1}{r}$, $\gamma \geq 0$ (nearly Lipshitz case): $f_\delta(r) \leq cr\log^{\gamma+n/2}(1/r)$.

For more information on the stochastic analysis of Gaussian random fields with regularity scales 1 and 2, see [3] and [10], respectively.

6 Appendix: the fractional exponential Poincaré inequality; proof of Proposition 4.2.

Here we prove Proposition 4.2. In the proof below, we denote by $\|\cdot\|$ the norm of $L^2\left((\mathbf{R}_+)^k, d\bar{s}\right)$ for any integer k. We will also make use of the notation $\|\|\|\cdot\|\|\|$ for iterated L^2 norms.

Step 1: Setup. We only need to show that

$$\mathbf{E}\left[\exp\left(|Y/M|^{2/n}\right)\right] \leq K(n)$$

where the constant $K(n)$ depends only on n. Indeed, if $K(n) \leq 2$, we may take $C_n = 1$, and otherwise we must take $C_n > 1$. We can thus use Jensen's inequality to write

$$\mathbf{E}\left[\exp\left(\left|\frac{Y}{C_n M}\right|^{2/n}\right)\right] = \mathbf{E}\left[\left(\exp\left(|Y/M|^{2/n}\right)\right)^{(C_n)^{-2/n}}\right]$$
$$\leq K(n)^{(C_n)^{-2/n}}$$

so that it is sufficient to take $C_n = (\log K(n) / \log 2)^{n/2}$.

Step 2: Poincaré inequality for fractional exponential moments. We will use a coupling inequality of Üstünel for convex functions, namely Theorem 9.2.2 in [16]. Since the function $\exp\left(|x|^{2/n}\right)$ is convex only for $|x|^{2/n} \geq (n-2)/2$, we write brutally

$$\mathbf{E}\left[\exp\left(|Y/M|^{2/n}\right)\right] \leq \mathbf{E}\left[\exp\left(|Y/M|^{2/n} \vee a_n\right)\right]$$

where $a_n := (n/2 - 1)$. Now we claim that Theorem 9.2.2 in [16], allows us to prove the following.

Lemma 6.1 (Fractional Exponential Poincaré Inequality) *For any centered random variable $X \in L^2(\Omega)$, and any value $\alpha \in [0,1]$,*

$$\begin{aligned}\mathbf{E}_X\left[\exp\left(|X|^{\alpha} \vee a_{2/\alpha}\right)\right] &\leq \mathbf{E}_X \otimes \mathbf{E}_Z\left[\exp\left(\left|Z\frac{\pi}{2}\|DX\|\right|^{\alpha} \vee a_{2/\alpha}\right)\right] \\ &= \mathbf{E}_Z\left[\mathbf{E}_X\left[\exp\left(\left|Z\frac{\pi}{2}\|DX\|\right|^{\alpha} \vee a_{2/\alpha}\right)\right]\right] \end{aligned} \tag{24}$$

where under the measure $P_X \otimes P_Z$, Z and X are independent, and Z is standard normal.

Similarly, for a centered random field X on an index set I with $X(t) \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ for all $t \in I$, and with $\|\cdot\|_{\mathcal{F}}$ a norm on a function space $\mathcal{F}$ such that $X(\cdot) \in \mathcal{F}$ almost-surely,

$$\begin{aligned}\mathbf{E}_X\left[\exp\left(\|X\|_{\mathcal{F}}^{\alpha} \vee a_{2/\alpha}\right)\right] &\leq \mathbf{E}_X \otimes \mathbf{E}_Z\left[\exp\left(\left\|Z\frac{\pi}{2}\|DX\|\right\|_{\mathcal{F}}^{\alpha} \vee a_{2/\alpha}\right)\right] \\ &= \mathbf{E}_Z\left[\mathbf{E}_X\left[\exp\left(\left\|Z\frac{\pi}{2}\|DX\|\right\|_{\mathcal{F}}^{\alpha} \vee a_{2/\alpha}\right)\right]\right], \end{aligned} \tag{25}$$

Proof. To prove (24), simply note that Üstünel's notation $I_1(\nabla\varphi(w))(z) = \int_0^1 \frac{d}{dt}\nabla\varphi(w,t)\,dz_t$ is none other than $\int_0^1 (D_t\varphi)(\omega)\,dz_t$ where $\varphi = \varphi(\omega)$ is a random variable w.r.t ω, and z is a Brownian motion on a space not related to ω. Therefore, with respect to the randomness of the Brownian motion z (and hence, with ω fixed), the random variable $I_1(\nabla\varphi(w))(z)$ has the same distribution as a Gaussian r.v. with variance $\int_0^1 |D_t\varphi|^2(\omega)\,dt$. Now we repeat precisely the proof of Theorem 9.2.3 in [16] with $U(x) = \exp\left(|x|^{\alpha} \vee a_{2/\alpha}\right)$. Let $\tilde{X} := X(\tilde{\omega})$ be an independent copy of $X = X(\omega)$ and we write X as $X = X(\omega) - \mathbf{E}_{\tilde{X}}X(\tilde{\omega})$. Using Jensen's inequality w.r.t. $\mathbf{E}_{\tilde{X}}$ for the convex function $x \mapsto \exp\left(|X(\omega) - x|^{\alpha} \vee a_{2/\alpha}\right)$

$$\begin{aligned}\mathbf{E}_X\left[\exp\left(|X|^{\alpha} \vee a_{2/\alpha}\right)\right] &= \mathbf{E}_X\left[\exp\left(|X(\omega) - \mathbf{E}_{\tilde{X}}X(\tilde{\omega})|^{\alpha} \vee a_{2/\alpha}\right)\right] \\ &\leq \mathbf{E}_X\left[\mathbf{E}_{\tilde{X}}\left[\exp\left(|X(\omega) - X(\tilde{\omega})|^{\alpha} \vee a_{2/\alpha}\right)\right]\right].\end{aligned}$$

Next we invoke Theorem 9.2.2 in [16] with the $U(x)$ and the remark above on the law of $I_1(\nabla\varphi(w))(z)$ for w fixed, yielding that the last expression above is

$$\begin{aligned}\mathbf{E}_X\left[\mathbf{E}_{\tilde{X}}\left[\exp\left(|X(\omega) - X(\tilde{\omega})|^{\alpha} \vee a_{2/\alpha}\right)\right]\right] &= \mathbf{E}_X\left[\mathbf{E}_{\tilde{X}}\left[U(X(\omega) - X(\tilde{\omega}))\right]\right] \\ &\leq \mathbf{E}_X\left[\mathbf{E}_Z\left[U\left(\frac{\pi}{2}I_1(\nabla X(\omega))(z)\right)\right]\right] \\ &= \mathbf{E}_X \otimes \mathbf{E}_Z\left[\exp\left(\left|Z\frac{\pi}{2}\|DX\|\right|^{\alpha} \vee a_{2/\alpha}\right)\right]\end{aligned}$$

where Z is standard normal and which proves (24). To prove (25), we claim that Lemma 9.2.1 in [16] also holds for processes and fields.

Lemma (Intermediate Lemma) *Let $X = \{X_s : s \in I\}$ be a Gaussian field on an index set I with values in $\mathbf{R}^d$ and let $U(X) = \exp(\|X\|^\alpha \vee a_{2/\alpha})$, where $\|\cdot\|$ is a norm on a function space $\mathcal{F}$ such that $X \in \mathcal{F}$ a.s. For any C^1-function $V : \mathbf{R}^d \to \mathbf{R}$, with V' its gradient, we have the following inequality:*

$$\mathbf{E}[U(V(X_\cdot) - V(Y_\cdot))] \leq \mathbf{E}\left[U\left(\frac{\pi}{2}(V'(X_\cdot), Y_\cdot)_{\mathbf{R}^d}\right)\right],$$

where Y is an independent copy of X and $\mathbf{E}$ is the expectation with respect to the product measure.

The proof of this intermediate lemma follows the same proof in [16]. Indeed let $X_s(\theta) = X_s \sin\theta + Y_s \cos\theta$. Then

$$\begin{aligned} V(X_s) - V(Y_s) &= \int_0^{\pi/2} \frac{d}{d\theta} V(X_s(\theta)) d\theta \\ &= \int_0^{\pi/2} (V'(X_s(\theta)), X_s'(\theta))_{\mathbf{R}^d} \, d\theta \\ &= \frac{\pi}{2} \int_0^{\pi/2} (V'(X_s(\theta)), X_s'(\theta))_{\mathbf{R}^d} \, d\tilde{\theta} \end{aligned}$$

where $d\tilde{\theta} = \frac{d\theta}{\pi/2}$. We have

$$\begin{aligned} U(V(X_\cdot) - V(Y_\cdot)) &= \exp\left(\left\|\int_0^{\pi/2} \frac{\pi}{2}(V'(X_\cdot(\theta)), X_\cdot'(\theta))_{\mathbf{R}^d} \, d\tilde{\theta}\right\|^\alpha \vee a_{2/\alpha}\right) \\ &\leq \exp\left(\left(\int_0^{\pi/2} \left\|\frac{\pi}{2}(V'(X_\cdot(\theta)), X_\cdot'(\theta))_{\mathbf{R}^d}\right\| d\tilde{\theta}\right)^\alpha \vee a_{2/\alpha}\right) \\ &\leq \int_0^{\pi/2} \exp\left(\left\|\frac{\pi}{2}(V'(X_\cdot(\theta)), X_\cdot'(\theta))_{\mathbf{R}^d}\right\|^\alpha \vee a_{2/\alpha}\right) d\tilde{\theta} \qquad (26) \\ &= \int_0^{\pi/2} U\left(\frac{\pi}{2}(V'(X_\cdot(\theta)), X_\cdot'(\theta))_{\mathbf{R}^d}\right) d\tilde{\theta} \end{aligned}$$

where in line (26), we used that the function $\|X\| \mapsto \exp\left(\|X\|^\alpha \vee a_{2/\alpha}\right)$ is convex. Moreover $X_s(\theta)$ and $X_s'(\theta)$ are two independent Gaussian processes with the same law as X_s. Hence

$$\begin{aligned} \mathbf{E}[U(V(X_\cdot) - V(Y_\cdot))] &\leq \int_0^{\pi/2} \mathbf{E}\left[U\left(\frac{\pi}{2}(V'(X_\cdot), Y_\cdot)_{\mathbf{R}^d}\right)\right] d\tilde{\theta} \\ &= \mathbf{E}\left[U\left(\frac{\pi}{2}(V'(X_\cdot), Y_\cdot)_{\mathbf{R}^d}\right)\right]. \end{aligned}$$

This proves the Intermediate Lemma.■

It immediately implies that Theorem 9.2.2 in [16] also holds for processes and fields. A similar argument as in the proof for (24) now proves (25) from this extension of Theorem 9.2.2 in [16]. This finishes the proof of the Fractional Exponential Poincaré Inequality Lemma 6.1. ■

Step 3: Iteration. Now we simply apply inequality (25) to inequality (24) with $\alpha = 2/n$ and iterate. Assuming $X = \sum_{k=n}^{\infty} I_k(f_k)$, the iterated Malliavin derivatives of X up to order $n-1$ will have mean 0, justifying the repeated use of (25) below. In the first pair of iterations we have

$$\begin{aligned}
&\mathbf{E}_X\left[\exp\left(|X|^{2/n}\vee a_n\right)\right]\\
&\leq \mathbf{E}_{Z_1}\left[\mathbf{E}_X\left[\exp\left(\left|Z_1\frac{\pi}{2}\right|^{2/n}\|DX\|^{2/n}\vee a_n\right)\right]\right]\\
&\leq \mathbf{E}_{Z_1}\left[\mathbf{E}_{Z_2}\left[\mathbf{E}_X\left[\exp\left(\left(\left|Z_2Z_1\left(\frac{\pi}{2}\right)^2\right|^{2/n}\|\,\|DDX\|\,\|_{\mathcal{F}}^{2/n}\right)\vee a_n\right)\right]\right]\right]\\
&= \mathbf{E}_{Z_1}\left[\mathbf{E}_{Z_2}\left[\mathbf{E}_X\left[\exp\left(\left(\left|Z_2Z_1\left(\frac{\pi}{2}\right)^2\right|^{2/n}\left\|D^{(2)}X\right\|^{2/n}\right)\vee a_n\right)\right]\right]\right],
\end{aligned}$$

where Z_1 and Z_2 are independent standard normals, and we used the fact that, in the notation of Lemma 6.1, since $\mathcal{F}$ is $L^2([0,1])$, we may write $L^2([0,1];\mathcal{F}) = L^2\left([0,1]^2\right)$. Hence, iterating this procedure we get, using a measure $\mathbf{P}_Z$ under which $Z_1, Z_2, \cdots, Z_n$ are IID standard normals,

$$\mathbf{E}_X\left[\exp\left(|X|^{2/n}\vee a_n\right)\right]\leq \mathbf{E}_Z\left[\mathbf{E}_X\left[\exp\left(\left(|Z_n\cdots Z_2Z_1|^{2/n}\frac{\pi^2}{4}\left\|D^{(n)}X\right\|^{2/n}\right)\vee a_n\right)\right]\right]$$

Now, replacing X by Y/M and noting that our hypothesis on Y says $\left\|D^{(n)}X\right\|\leq 1$,

$$\begin{aligned}
&\mathbf{E}_X\left[\exp\left(|Y/M|^{2/n}\vee a_n\right)\right]\\
&\leq \mathbf{E}_Z\left[\exp\left(\left(\frac{\pi^2}{4}|Z_n\cdots Z_2Z_1|^{2/n}\right)\vee(n/2-1)\right)\right] =: K(n)
\end{aligned}\tag{27}$$

The latter being a universal constant depending only on n, the theorem is proved. In order to make the definition of the constant $K(n)$ more explicit, we may use the relation between algebraic and geometric means to get

$$K(n)\leq\left(\mathbf{E}_{Z_1}\left[\exp\left(\frac{\pi^2}{4n}Z_1^2\vee(n/2-1)\right)\right]\right)^n.\tag{28}$$

The proof of Proposition 4.2 is complete.■

References

[1] Adler, R.J.; Taylor, J. (2007). *Random Fields and Geometry.* Springer Monographs in Mathematics.

[2] 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.

[3] Cheridito, P.; Nualart, D. (2005). Stochastic integral of divergence type with respect to fractional Brownian motion with Hurst parameter $H\in(0,1/2)$. *Ann. Inst. H. Poincaré Probab. Statist.* **41**, no. 6, 1049-1081.

[4] Dudley, R.M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *Journal of Functional Analysis* **1**, 290-330.

[5] Fernique, X. (1971). Régularité des processus gaussienes. Invent. Math. **12**, 304-320.

[6] Fernique, X. (1975). *Régularité des trajectoires des fonctions aléatoires gaussiennes.* In Ecole d'Eté de Probabilités de Saint-Flour IV-1974, Lecture Notes in Math. **480**, 1-96. Springer.

[7] 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-Verlag.

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

[9] Malliavin, P. (2002). *Stochastic Analysis.* Springer-Verlag.

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

[11] Nualart, D. (2006). *The Malliavin calculus and related topics, 2nd Ed.* Springer-Verlag.

[12] D. Revuz and M. Yor (1999). *Continuous Martingales and Brownian Motion. Third edition.* Springer-Verlag.

[13] Talagrand, M. (1987). Regularity of Gaussian processes. *Acta Math.* **159**, no. 1-2, 99–149.

[14] Talagrand, M. (2005). *The Generic Chaining.* Springer-Verlag.

[15] 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.

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

[17] Vizcarra, A.; Viens, F. (2007). Some applications of the Malliavin calculus to sub-Gaussian and non-sub-Gaussian random fields. *Seminar on Stochastic Analysis, Random Fields and Applications V. Progress in Probability.* Birkhäuser Verlag.