

Boundedness and continuity of Gaussian processes

A Gaussian random process (or better, random function) $X = (X_t)_{t \in T}$ indexed by a set T is a family of random variables on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that the law of each finite family $(X_{t_1}, \dots, X_{t_n}), t_1, \dots, t_n \in T$, is Gaussian on \mathbb{R}^n . For simplicity, Gaussian will always mean centered Gaussian. In particular, the law (the distributions of the finite-dimensional marginals) of the process X is uniquely determined by the covariance structure $\mathbb{E}(X_s X_t), s, t \in T$.

Such Gaussian processes are common and appear in numerous contexts, with index sets T of various types (the time interval $T = [0, \infty)$ being of central importance in evolution processes). A natural question is to decide under which conditions such a Gaussian process is almost surely bounded, or continuous provided T is endowed by a topology (such as an interval on the real line). (Due to the rather arbitrary nature of the set T , the question might require to discuss several measurability issues, not really relevant and not addressed here.) These conditions should a priori only involve the covariance structure of the process. This easy-to-state problem arises in various settings, and even the understanding of boundedness on a finite set T , that is the analysis of the random variable $\max_{t \in T} X_t$, can give rise to delicate (numerical) developments.

Going back to early studies by A. Kolmogorov, a key idea in the investigation of this problem, developed in particular by R. Dudley, V. Strassen, V. Sudakov, X. Fernique [5, 6, 22, 23, 24, 8] in the late sixties and early seventies, was to try to connect boundedness and regularity of a Gaussian process $X = (X_t)_{t \in T}$ to the size and geometry of the metric space

(T, d) where d is the L^2 -metric induced by the process itself

$$d(s, t) = \|X_s - X_t\|_{L^2} = (\mathbb{E}(|X_s - X_t|^2))^{1/2}, \quad s, t \in T. \quad (1)$$

This metric is entirely characterized by the covariance structure of the process. It does not necessarily separate points in T , but this is of no importance.

Covering (entropy) numbers of the metric space (T, d) have been a first tool in the investigation of this project, resulting in the famous Dudley-Sudakov entropy bound theorem. It was later improved with the concept of majorizing measures and admissible families of partitions, leading to the Fernique-Talagrand theorem, which provides a complete metric characterization of boundedness and continuity of Gaussian processes, a remarkable statement connecting a probabilistic property to a purely metric one.

This text is mainly taken from [13] and [14]. A brief historical account of the developments of the sixties and seventies has been provided by R. Dudley [7]. The article [6] by the latter, the monograph [25] of V. Sudakov, and the courses [8, 11] by X. Fernique provide a complete account on the Gaussian picture. See also [4, 15, 16]... for more modern expositions. A recent investigation, extended to large families of stochastic processes, is the monograph [27], while the contribution [17] by J. Liu emphasizes a new information-theoretic approach.

Table of contents

1. Kolmogorov's continuity theorem
2. Gaussian process and intrinsic distance
3. The Dudley-Sudakov theorem
4. The Dudley-Fernique theorem
5. Majorizing measures
6. The generic chaining
7. The Fernique-Talagrand theorem
8. Liu's new information-theoretic characterization of boundedness of Gaussian processes

References

1 Kolmogorov's continuity theorem

Kolmogorov's continuity theorem is a basic statement which ensures that a stochastic process (not necessarily Gaussian) $(Z_t)_{t \geq 0}$ indexed on $[0, \infty)$, or any interval of the real line, with real values (for simplicity) admits a version (i.e. a process $(\tilde{Z}_t)_{t \geq 0}$ such that, for every $t \geq 0$, $\tilde{Z}_t = Z_t$ almost surely) with continuous sample paths. The efficiency of the theorem stems from the fact that only conditions on the increments of the process are required.

Theorem 1 (Kolmogorov's continuity theorem). *Let $(Z_t)_{t \geq 0}$ be a stochastic process indexed by $[0, \infty)$. Assume that for every $T > 0$, there exist $\alpha, \beta > 0$ and $C > 0$ such that*

$$\mathbb{E}(|Z_s - Z_t|^\alpha) \leq C |s - t|^{1+\beta}$$

for all $0 \leq s, t \leq T$. Then $(Z_t)_{t \geq 0}$ admits a continuous version.

It may be shown moreover that the version has locally κ -Hölder continuous paths for every $0 < \kappa < \frac{\beta}{\alpha}$.

The proof of this theorem is based on a fundamental *chaining argument*. Basically, and with $T = 1$ for simplicity, given an increasing sequence of subdivisions of $[0, 1]$, for instance $\mathcal{T}_0 = \{0\}$, $\mathcal{T}_n = \{\frac{k}{2^n}; k = 0, 1, \dots, 2^n\}$, $n \geq 0$, and denoting by $s_n(t)$ the nearest element of \mathcal{T}_n from $t \in [0, 1]$, it amounts to the representation

$$Z_t = Z_0 + \sum_{n \geq 1} (Z_{s_n(t)} - Z_{s_{n-1}(t)})$$

and the inequality

$$\sup_{t \in [0, 1]} Z_t \leq Z_0 + \sum_{n \geq 1} \sup_{t \in [0, 1]} (Z_{s_n(t)} - Z_{s_{n-1}(t)})$$

on which conditions on increments may be exploited, the point being that the last supremum is actually a maximum on a finite set. This scheme will be abundantly illustrated in the next sections in the framework of Gaussian processes.

2 Gaussian process and intrinsic distance

Kolmogorov's continuity theorem may easily be applied to Gaussian processes $X = (X_t)_{t \in T}$ indexed by an interval of the real line, provided $\mathbb{E}(|X_s - X_t|^2)$ may be controlled in terms of the distance between s and t . Due to the equivalence of moments of Gaussian random variables, it is indeed enough to consider the L^2 -moment. But actually, the strong integrability properties of Gaussian random variables, illustrated for example by the Gaussian tails

$$\mathbb{P}(|X_s - X_t| \geq u) \leq e^{-\frac{u^2}{2} \mathbb{E}(|X_s - X_t|^2)}, \quad u \geq 0, \quad (2)$$

suggest on the one hand that Kolmogorov's theorem could be improved, in terms of the parameters α and β , within the class of Gaussian processes.

On the other hand, Gaussian processes $X = (X_t)_{t \in T}$ indexed by more general sets T may be considered. As presented in the introduction, their study is connected to the size of the metric space (T, d) for the intrinsic distance (1) (reflected on the parameters α and β in Kolmogorov's theorem).

Within this abstract framework, the main point of the investigation is the question of boundedness of the paths $t \mapsto X_t$, $t \in T$. Once the appropriate bounds on the supremum of X are obtained, the characterization of continuity (whenever T is topological) easily follows. Readers are referred to [8, 14, 15, 27] for technical details in this regard, not developed here.

In addition, due to the integrability properties of norms of Gaussian random vectors or supremum of Gaussian processes (cf. [1]), various cumbersome and unessential measurability questions are avoided by considering the supremum functional

$$F(T) = \sup \left\{ \mathbb{E} \left(\sup_{t \in U} X_t \right); U \text{ finite in } T \right\}.$$

(If $S \subset T$, $F(S)$ is defined in the same way.) Thus, $F(T) < \infty$ if and only if X is almost surely bounded in any reasonable sense. In particular, the main question will reduce to a uniform control of $F(U)$ over the finite subsets U of T .

3 The Dudley-Sudakov theorem

A standard and useful way to measure the size of a metric space (T, d) is provided by *entropy* or *covering numbers*. For every $\varepsilon > 0$, let $N(T, d; \varepsilon)$ denote the minimal number of (open to fix the idea) balls of radius ε for the metric d that are necessary to cover T . The logarithm of these covering numbers $N(T, d; \varepsilon)$ is usually referred to as entropy numbers.

The two main results concerning regularity of Gaussian processes under entropy conditions, due to R. Dudley [5, 6] for the upper-bound and V. Sudakov [24] for the lower-bound (cf. [6, 8]) are summarized in the following statement. The parameter set T is endowed with the intrinsic metric d (1).

Theorem 2 (The Dudley-Sudakov theorem). *There are numerical constants $C_1 > 0$ and $C_2 > 0$ such that for every Gaussian process $X = (X_t)_{t \in T}$,*

$$C_1^{-1} \sup_{\varepsilon > 0} \varepsilon \sqrt{\log N(T, d; \varepsilon)} \leq F(T) \leq C_2 \int_0^\infty \sqrt{\log N(T, d; \varepsilon)} d\varepsilon. \quad (3)$$

Possible numerical values are $C_1 = 6$ and $C_2 = 42$ (see below). The integral on the right-hand side of (3) is often called *Dudley's entropy integral*. As emphasized by R. Dudley

himself [7], while he defined such an integral, the upper-bound in (3) was first established by V. Sudakov in [24, 25]. He also refers to [19]. The convergence of the integral is understood for the small values of ε since it stops at the diameter $D(T) = \sup\{d(s, t); s, t \in T\}$. Actually, if any of the three terms of (3) is finite, then (T, d) is totally bounded, and in particular $D(T) < \infty$.

As mentioned above, it may be shown that the process $X = (X_t)_{t \in T}$ actually admits an almost surely continuous (with respect to d) version when the entropy integral is finite. Conversely, if $X = (X_t)_{t \in T}$ is continuous, it holds true that $\lim_{\varepsilon \rightarrow 0} \varepsilon \sqrt{\log N(T, d; \varepsilon)} = 0$ (cf. [8]). Roughly speaking, continuity at t_0 follows from the bounds (3) applied to the process $X_t - X_{t_0}$ for t in a ball of radius $\eta > 0$.

Theorem 2 may be tested on the Kolmogorov theorem. On the interval $[0, T]$, with $\alpha = 2$ and some $\beta > 0$, the covering numbers $N(T, d; \varepsilon)$ are of the order of $\varepsilon^{-2/(1+\beta)}$, so that the entropy integral is clearly finite, actually for any $\beta > -1$.

The proof of the right-hand side of (3) emphasizes the basic chaining argument.

Proof. It may be assumed that T is finite. Let $q > 1$ (usually an integer), that is thought of as a power of discretization; a posteriori, its value is completely arbitrary. Let n_0 be the largest integer n in \mathbb{Z} such that $N(T, d; q^{-n}) = 1$. For every $n \geq n_0$, consider a family of cardinality $N(T, d; q^{-n}) = N(n)$ of balls of radius q^{-n} covering T . One may therefore construct a partition \mathcal{A}_n of T of cardinality $N(n)$ on the basis of this covering with sets of diameter less than $2q^{-n}$. In each A of \mathcal{A}_n , fix a point of T and denote by T_n the collection of these points. For each t in T , denote by $A_n(t)$ the element of \mathcal{A}_n that contains t . For every t and every n , let then $s_n(t)$ be the element of T_n such that $t \in A_n(s_n(t))$. Note that $d(t, s_n(t)) \leq 2q^{-n}$ for every t and $n \geq n_0$.

The main argument of the proof is the so-called chaining argument going back to A. Kolmogorov in his proof of Theorem 1. For every $t \in T$, write

$$X_t = X_{s_0} + \sum_{n > n_0} (X_{s_n(t)} - X_{s_{n-1}(t)}) \quad (4)$$

where $s_0 = s_{n_0}(t)$ may be chosen independent of $t \in T$ (and the sum is finite). Note that

$$d(s_n(t), s_{n-1}(t)) \leq 2q^{-n} + 2q^{-n+1} = 2(q+1)q^{-n}.$$

Let $c_n = 4(q+1)q^{-n} \sqrt{\log N(n)}$, $n > n_0$. It follows from (4) that

$$\begin{aligned} F(T) &= \mathbb{E} \left(\sup_{t \in T} X_t \right) \leq \sum_{n > n_0} c_n + \mathbb{E} \left(\sup_{t \in T} \sum_{n > n_0} |X_{s_n(t)} - X_{s_{n-1}(t)}| \mathbb{1}_{\{|X_{s_n(t)} - X_{s_{n-1}(t)}| > c_n\}} \right) \\ &\leq \sum_{n > n_0} c_n + \mathbb{E} \left(\sum_{n > n_0} \sum_{(u,v) \in H_n} |X_u - X_v| \mathbb{1}_{\{|X_u - X_v| > c_n\}} \right) \end{aligned}$$

where $H_n = \{(u, v) \in T_n \times T_{n-1}; d(u, v) \leq 2(q+1)q^{-n}\}$. If G is a real centered Gaussian variable with variance less than or equal to σ^2 , for every $c > 0$

$$\mathbb{E}(|G| \mathbb{1}_{\{|G|>c\}}) \leq \sigma e^{-c^2/2\sigma^2}. \quad (5)$$

Hence,

$$\begin{aligned} F(T) &\leq \sum_{n>n_0} c_n + \sum_{n>n_0} \text{Card}(H_n) 2(q+1)q^{-n} \exp(-c_n^2/8(q+1)^2q^{-2n}) \\ &\leq \sum_{n>n_0} 4(q+1)q^{-n} \sqrt{\log N(n)} + \sum_{n>n_0} 2(q+1)q^{-n} \\ &\leq 7(q+1) \sum_{n>n_0} q^{-n} \sqrt{\log N(n)} \end{aligned}$$

where it is used that $\text{Card}(H_n) \leq N(n)^2$. Since

$$\begin{aligned} \int_0^\infty \sqrt{\log N(T, d; \varepsilon)} d\varepsilon &\geq \sum_{n>n_0} \int_{q^{-n-1}}^{q^{-n}} \sqrt{\log N(T, d; \varepsilon)} d\varepsilon \\ &\geq (1-q^{-1}) \sum_{n>n_0} q^{-n} \sqrt{\log N(n)}, \end{aligned}$$

the conclusion follows. If $q = 2$, the value $C_2 = 42$ is acceptable.

The proof of the lower-bound in (3) is an easy consequence of the Gaussian comparison inequalities going back to Slepian's lemma [21]. The convenient tool is the Sudakov-Chevet-Fernique inequality, see [2]. Fix $\varepsilon > 0$ and let $n \leq N(T, d; \varepsilon)$. There exist therefore t_1, \dots, t_n in T such that $d(t_i, t_j) \geq \varepsilon$. Let then g_1, \dots, g_n be independent standard normal random variables. For every $i, j = 1, \dots, n, i \neq j$,

$$\mathbb{E}\left(\left|\frac{\varepsilon}{\sqrt{2}} g_i - \frac{\varepsilon}{\sqrt{2}} g_j\right|^2\right) = \varepsilon^2 \leq d(t_i, t_j)^2 = \mathbb{E}(|X_{t_i} - X_{t_j}|^2).$$

Therefore, by the Sudakov-Chevet-Fernique inequality,

$$F(T) \geq \mathbb{E}\left(\max_{1 \leq i \leq n} X_{t_i}\right) \geq \frac{\varepsilon}{\sqrt{2}} \mathbb{E}\left(\max_{1 \leq i \leq n} g_i\right).$$

Now, it is classical (cf. [3]) that $\mathbb{E}(\max_{1 \leq i \leq n} g_i) \geq c\sqrt{\log n}$ for some numerical $c > 0$ (one may choose c such that $\frac{\sqrt{2}}{c} \leq 6$). Since n is arbitrary less than or equal to $N(T, d; \varepsilon)$, the claim follows. Theorem 2 is therefore fully established. \square

It is worthwhile mentioning that the Dudley entropy bound applies to each (centered) stochastic process $(X_t)_{t \in T}$, indexed by some metric space (T, d) , such that, for some constant $C > 0$,

$$\mathbb{P}(|X_s - X_t| \geq u) \leq C e^{-\frac{u^2}{C} d(s,t)^2}, \quad u \geq 0,$$

for all $s, t \in T$, $u \geq 0$. Such processes are called sub-Gaussian. The argument further applies to other types of tails, by suitably modifying the entropy integral (cf. [19, 14, 27]).

4 The Dudley-Fernique theorem

Although rather tight, it soon appeared historically that the Dudley-Sudakov entropy bounds of Theorem 3 are not precise enough to fully characterize boundedness of a Gaussian process in terms of the complexity of the associate metric space (T, d) . It should be emphasized however that the Dudley entropy integral does describe boundedness and continuity of stationary Gaussian processes (when the intrinsic distance d is translation invariant for a group structure on the parameter set T). The result is due to X. Fernique [8] and states the following.

Theorem 3 (The Dudley-Fernique theorem). *Let $X = (X_t)_{t \in T}$ be a centered Gaussian process, stationary with respect to a translation invariant group action on the parameter set T . Then, up to numerical constants,*

$$\mathbb{E}(\sup_{t \in T} X_t) \sim \int_0^\infty \sqrt{\log N(T, d; \varepsilon)} d\varepsilon.$$

A short and simple proof of this result is exposed in [13]. The equivalence of the theorem has been instrumental in the study of random Fourier series by M. Marcus and G. Pisier [18].

5 Majorizing measures

To try to remedy to the limitations of entropy numbers to characterize boundedness and continuity of Gaussian processes, a sharper upper-bound was emphasized by X. Fernique in the form of majorizing measures [8]. Trying to imagine what can be used instead of the entropy numbers in order to sharpen the conclusions of Theorem 2, it is important to realize that one feature of entropy is that it attributes an equal weight to each piece of the parameter set T (so that stationary instances may then indeed be characterized by this tool). The following definition, taking care of local patterns, appears as a possible sharper substitute. Let $q > 1$ (an integer), and $\mathcal{A} = (\mathcal{A}_n)_{n \in \mathbb{Z}}$ be an increasing sequence of finite partitions of T (i.e. each $A \in \mathcal{A}_{n+1}$ is contained in some $B \in \mathcal{A}_n$) such that the diameter

$D(A)$ of each element A of \mathcal{A}_n is less than or equal to $2q^{-n}$. If $t \in T$, denote by $A_n(t)$ the element of \mathcal{A}_n that contains t . Now, for each partition \mathcal{A}_n , consider a family of non-negative weights $\alpha_n(A)$, $A \in \mathcal{A}_n$, such that $\sum_{A \in \mathcal{A}_n} \alpha_n(A) \leq 1$. Set then

$$\Theta_{\mathcal{A},\alpha}(T, d) = \sup_{t \in T} \sum_n q^{-n} \sqrt{\log \frac{1}{\alpha_n(A_n(t))}}.$$

It is worthwhile mentioning that for $2q^{-n} \geq D(T)$, one can take $\mathcal{A}_n = \{T\}$ and $\alpha_n(T) = 1$.

Now, Fernique's observation is that the proof of Dudley's upper bound may be (almost exactly) repeated so to yield that, for any such family $\{\mathcal{A}, \alpha\}$ of partitions and weights,

$$F(T) \leq C \Theta_{\mathcal{A},\alpha}(T, d) \tag{6}$$

where $C = C(q) > 0$. Hence the single entropy integral is replaced by the functionals $\Theta_{\mathcal{A},\alpha}(T, d)$ varying with the partitions and weights.

The family of weights $\{\mathcal{A}, \alpha\}$ and the functional $\Theta_{\mathcal{A},\alpha}$ may also be synthesized as a single "majorizing measure" probability m on the Borel sets of T of the form

$$\sup_{t \in T} \int_0^\infty \sqrt{\log \frac{1}{m(B(t, \varepsilon))}} d\varepsilon \tag{7}$$

where $B(t, \varepsilon)$ is the ball in T with center t and radius $\varepsilon > 0$. As such, Fernique's result is closely related to the integral bounds in [20, 12]. It is not difficult to check (cf. [27, 13]) that Dudley's entropy integral produces a majorizing measure, and that in a stationary setting, the relevant majorizing measure is the Haar measure on the (pseudo-) group T , and is equivalent to Dudley's integral.

6 The generic chaining

Modern expositions have replaced majorizing measures by admissible partitions, providing a purely metric description. The monograph [27] presents all the details of the relationships between these various objects and definitions.

Given a (finite) metric space (T, d) , consider an increasing sequence of finite partitions $(\mathcal{C}_n)_{n \geq 0}$ of T such that $\text{Card}(\mathcal{C}_0) = 1$ and $\text{Card}(\mathcal{C}_n) \leq 2^{2^n}$, $n \geq 1$. (Recall that by increasing, it is meant that each element of \mathcal{C}_n is contained in a cell of \mathcal{C}_{n-1} .) For each $n \geq 0$, fix a point in each element C of \mathcal{C}_n , and denote by T_n the collection of those points. By construction, $\text{Card}(T_n) \leq 2^{2^n}$ (while $\text{Card}(T_0) = 1$). For a point t in T , denote by $C_n(t)$ the unique element of \mathcal{C}_n that contains t , and by $s_n(t)$ the element of T_n such that $t \in C_n(s_n(t))$ ($s_0(t) = s_0$ may be taken as an arbitrary fixed point in T).

With such a sequence of partitions, the chaining argument may be developed as with entropy numbers or majorizing measures (weights). Indeed, given a (centered) Gaussian process $(X_t)_{t \in T}$ with intrinsic metric $d(s, t) = (\mathbb{E}|X_s - X_t|^2)^{1/2}$, $s, t \in T$, start again with the decomposition (4)

$$X_t = X_{s_0} + \sum_{n \geq 1} (X_{s_n(t)} - X_{s_{n-1}(t)}).$$

Arguing then almost as in the proof of the upper-bound in (3), but with the truncation c_n depending on t ,

$$\begin{aligned} X_t - X_{s_0} &\leq \sum_{n \geq 1} 2^{(n+2)/2} d(s_n(t), s_{n-1}(t)) \\ &\quad + \sum_{n \geq 1} |X_{s_n(t)} - X_{s_{n-1}(t)}| \mathbb{1}_{\{|X_{s_n(t)} - X_{s_{n-1}(t)}| > 2^{(n+2)/2} d(s_n(t), s_{n-1}(t))\}}. \end{aligned}$$

Taking the supremum in $t \in T$ and integrating,

$$\begin{aligned} F(T) &= \mathbb{E} \left(\sup_{t \in T} X_t \right) \\ &\leq \sup_{t \in T} \sum_{n \geq 1} 2^{(n+2)/2} d(s_n(t), s_{n-1}(t)) \\ &\quad + \mathbb{E} \left(\sup_{t \in T} \sum_{n \geq 1} |X_{s_n(t)} - X_{s_{n-1}(t)}| \mathbb{1}_{\{|X_{s_n(t)} - X_{s_{n-1}(t)}| > 2^{(n+2)/2} d(s_n(t), s_{n-1}(t))\}} \right) \\ &\leq \sup_{t \in T} \sum_{n \geq 1} 2^{(n+2)/2} d(s_n(t), s_{n-1}(t)) \\ &\quad + \mathbb{E} \left(\sum_{n \geq 1} \sum_{(u, v) \in T_n \times T_{n-1}} |X_u - X_v| \mathbb{1}_{\{|X_u - X_v| > 2^{(n+2)/2} d(u, v)\}} \right) \\ &\leq \sup_{t \in T} \sum_{n \geq 1} 2^{(n+2)/2} d(s_n(t), s_{n-1}(t)) + D(T) \sum_{n \geq 1} \text{Card}(T_n \times T_{n-1}) e^{-2^{n+1}} \end{aligned}$$

where (5) applied to $(X_u - X_v)/d(u, v)$ with $d(u, v) \leq D(T)$ is used in the last step. Now

$$d(s_n(t), s_{n-1}(t)) \leq d(s_n(t), t) + d(t, s_{n-1}(t)) \leq D(C_n(t)) + D(C_{n-1}(t))$$

while

$$\text{Card}(T_n \times T_{n-1}) \leq 2^{2^n} \cdot 2^{2^{n-1}} \leq 2^{2^{n+1}}.$$

As a conclusion,

$$F(T) \leq 5 \sup_{t \in T} \sum_{n \geq 0} 2^{n/2} D(C_n(t)) + D(T) \leq 6 \sup_{t \in T} \sum_{n \geq 0} 2^{n/2} D(C_n(t)) \quad (8)$$

(since $C_0(t) = T$).

7 The Fernique-Talagrand theorem

Call the sequence $\mathcal{C} = (\mathcal{C}_n)_{n \geq 0}$ described in the previous section a “sequence of admissible partitions”. The preceding analysis (8) expresses that

$$F(T) \leq 6 \gamma_2(T, d) \tag{9}$$

where

$$\gamma_2(T, d) = \inf \sup_{t \in T} \sum_{n \geq 0} 2^{n/2} D(C_n(t)),$$

the infimum being taken over all such sequences of admissible partitions.

Within the notation of Section 4, denote by $\Theta(T, d)$ the infimum of the functional $\Theta_{\mathcal{A}, \alpha}(T, d)$ over all possible choices of partitions $(\mathcal{A}_n)_{n \in \mathbb{Z}}$ and weights $\alpha_n(A)$, $A \in \mathcal{A}_n$, $n \in \mathbb{Z}$. Alternatively, consider the infimum over all probability measures m in (7). These quantities are all equivalent, up to numerical constants, to $\gamma_2(T, d)$ (cf. [27]).

The functional $\gamma_2(T, d)$ is therefore an alternate upper bound on $F(T) = \mathbb{E}(\sup_{t \in T} X_t)$. A main achievement by M. Talagrand in 1987 [26] is that it is actually also a lower bound.

Theorem 4 (The Fernique-Talagrand theorem). *There exists a numerical constant $K > 0$ such that, for any Gaussian process $(X_t)_{t \in T}$,*

$$\frac{1}{K} \gamma_2(T, d) \leq F(T) \leq K \gamma_2(T, d). \tag{10}$$

This major result thus provides a purely metric characterization of almost sure boundedness of Gaussian processes in terms of functionals associated to families of weights, majorizing measures, or sequences of admissible partitions. The monograph [27] is an extensive investigation of families of such functionals, with a wide range of applications far away from just Gaussian processes.

8 Liu’s new information-theoretic characterization of boundedness of Gaussian processes

In the striking contribution [17], J. Liu proposes a new information-theoretic characterization of boundedness of Gaussian processes, with in particular a lifting argument reducing to stationary Gaussian processes for which the classical metric entropy numbers produce a full and simple description by the Dudley-Fernique theorem. The following briefly outlines the main idea.

If $X = (X_t)_{t \in T}$ is a stochastic process on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\sup_{t \in T} \mathbb{E}(|X_t|) < \infty$, and m a probability measure on T , set

$$\mathcal{F}(X, m) = \sup \mathbb{E}(X_Z) \quad (11)$$

where the supremum is taken over all random variables Z on T with distribution m . Since only depending on the law P_X of X on \mathbb{R}^T , it is also denoted $\mathcal{F}(P_X, m)$. This functional was introduced by X. Fernique in the seminal works [9, 10] to try to understand the expected supremum of random processes, in particular Gaussian processes. Fernique's intuition was that the (or a) majorizing measure (as the infimum of the functional (7)) should be closely connected to the distribution of the maximizer of the Gaussian process, that is the distribution of the random variable τ such that $X_{\tau(\omega)}(\omega) = \sup_{t \in T} X_t(\omega)$ (assume T is separated by the metric d). This intuition appears to be correct as developed in Section 4 of [Ta87].

Given T finite, for M an integer ≥ 1 let $\mathcal{P}_M = \{\frac{1}{M} \sum_{k=1}^M \delta_{t_k} ; t_1, \dots, t_M \in T\}$ be the set of empirical probability measures of size M on the points of T . For integers $M, N \geq 1$ and $m \in \mathcal{P}_M$, consider next the set

$$\mathbf{T}_N(m) = \left\{ \mathbf{t} = (t_1, \dots, t_{MN}) \in T^{MN} ; \frac{1}{MN} \sum_{k=1}^{MN} \delta_{t_k} = m \right\}$$

of sequences of T^{MN} in which each $t \in T$ appears exactly $MNm(\{t\})$ times. Liu's fundamental lifting principle ([17, Lemma 5], [28, Proposition 3.1]) expresses that under the preceding notation,

$$\mathcal{F}(X, m) = \lim_{N \rightarrow \infty} \mathbb{E} \left(\sup_{\mathbf{t} \in \mathbf{T}_N(m)} \mathbf{X}_{\mathbf{t}}^{m, N} \right) \quad (12)$$

where

$$\mathbf{X}_{\mathbf{t}}^{m, N} = \frac{1}{MN} \sum_{k=1}^{MN} X_{\mathbf{t}_k}^{(k)}, \quad \mathbf{t} = (t_1, \dots, t_{MN}) \in \mathbf{T}_N(m),$$

where $X^{(1)}, X^{(2)}, \dots$ are independent copies of the process X . By centering of X , the limit in N may be shown to be increasing. The main point of the result is that, for each $N \geq 1$, the process $(\mathbf{X}_{\mathbf{t}}^{m, N})_{\mathbf{t} \in \mathbf{T}_N(m)}$ is stationary with respect to the action of the symmetric group \mathcal{S}_{MN} . Indeed, \mathcal{S}_{MN} acts on $\mathbf{T}_N(m)$ by $\sigma(\mathbf{t}) = (\mathbf{t}_{\sigma(1)}, \dots, \mathbf{t}_{\sigma(MN)})$ for $\sigma \in \mathcal{S}_{MN}$ and $\mathbf{t} = (t_1, \dots, t_{MN}) \in \mathbf{T}_N(m)$. This action is transitive and since $\mathbf{X}_{\sigma(\mathbf{t})}^{m, N} = \frac{1}{MN} \sum_{k=1}^{MN} X_{\mathbf{t}_k}^{(\sigma^{-1}(k))}$, the processes $(\mathbf{X}_{\sigma(\mathbf{t})}^{m, N})_{\mathbf{t} \in \mathbf{T}_N(m)}$ and $(\mathbf{X}_{\mathbf{t}}^{m, N})_{\mathbf{t} \in \mathbf{T}_N(m)}$ have the same distribution for every $\sigma \in \mathcal{S}_{MN}$. The representation (13) thus reduces the computation of Fernique's functional $\mathcal{F}(X, m)$ for an arbitrary random process to the computation of the expected supremum of a stationary process.

Now $\mathbb{E}(\sup_{t \in T} X_t) = \sup_m \mathcal{F}(X, m)$, and, by continuity of the map $m \mapsto \mathcal{F}(X, m)$, the supremum may actually be taken over all probability measures in \mathcal{P}_M , $M \geq 1$. Given then

a centered Gaussian process $X = (X_t)_{t \in T}$ indexed by T (finite), the theorem provides a full characterization of boundedness of X in terms of the entropy integral for the N -extended (Gaussian) processes $(\mathbf{X}_{\mathbf{t}}^{m,N})_{\mathbf{t} \in \mathbf{T}_N(m)}$ indexed by $\mathbf{T}_N(m)$ (as $N \rightarrow \infty$) as formulated by the Dudley-Fernique theorem, that is

$$\mathbb{E} \left(\sup_{\mathbf{t} \in \mathbf{T}_N(m)} \mathbf{X}_{\mathbf{t}}^{m,N} \right) \sim \int_0^\infty \sqrt{\log(\mathbf{T}_N(m), \mathbf{d}; \varepsilon)} d\varepsilon$$

for the intrinsic metric

$$\mathbf{d}(\mathbf{s}, \mathbf{t})^2 = \frac{1}{MN} \sum_{k=1}^{MN} d(\mathbf{s}_k, \mathbf{t}_k)^2, \quad \mathbf{s}, \mathbf{t} \in \mathbf{T}_N(m).$$

Therefore

$$\begin{aligned} \mathbb{E} \left(\sup_{t \in T} X_t \right) &\sim \sup_{m \in \mathcal{P}_M, M \geq 1} \lim_{N \rightarrow \infty} \mathbb{E} \left(\sup_{\mathbf{t} \in \mathbf{T}_N(m)} \mathbf{X}_{\mathbf{t}}^{m,N} \right) \\ &\sim \sup_{m \in \mathcal{P}_M, M \geq 1} \lim_{N \rightarrow \infty} \int_0^\infty \sqrt{\log(\mathbf{T}_N(m), \mathbf{d}; \varepsilon)} d\varepsilon \end{aligned} \quad (13)$$

(up to numerical constants).

In his work [17], J. Liu provides an information-theoretic description of the right-hand side of (13) via coding theory, and from which he establishes its equivalence with the Talagrand $\gamma_2(T, d)$ functional, briefly described below. But (13) already provides a simple direct metric characterization of boundedness of Gaussian process in terms of the classical metric entropy numbers.

Liu's information-theoretic description is as follows (referring to [17] for a more careful exposition). For m a probability measure on a metric space (T, d) and $\varepsilon > 0$, set

$$\Pi_\varepsilon(m) = \left\{ (U, V); \mathcal{L}(U) = \mathcal{L}(V) = m, \mathbb{E}(d(U, V)^2) \leq \varepsilon^2 \right\}.$$

Define then

$$i_m(\varepsilon) = \inf_{(U, V) \in \Pi_\varepsilon(m)} \mathbb{I}(U, V)$$

where $\mathbb{I}(U, V)$ is the mutual information for the law $P_{(U, V)}$ of (U, V) defined as the relative entropy

$$\mathbb{I}(U, V) = \int_T \log \left(\frac{P_{(U, V)}}{P_U \otimes P_V} \right) dP_{(U, V)}$$

of $P_{(U, V)}$ with respect to the product law $P_U \otimes P_V$ of the marginals. J. Liu then establishes by the main theorem above together with basic coding theory that

$$\mathcal{F}(X, m) \sim \int_0^\infty \sqrt{i_m(\varepsilon)} d\varepsilon$$

up to numerical constants. The supremum over m of the right-hand side is then showed to be equivalent to the $\gamma_2(T, d)$ functional, providing an alternate proof of the majorizing measure theorem.

References

- [1] Admissible shift, reproducing kernel Hilbert space, and abstract Wiener space. *The Gaussian Blog*.
- [2] Gaussian comparison inequalities. *The Gaussian Blog*.
- [3] Some basics on Gaussian measures and variables. *The Gaussian Blog*.
- [4] R. Adler. *An introduction to continuity, extrema, and related topics for general Gaussian processes*. Lecture Notes–Monograph Series 12. Institute of Mathematical Statistics (1990).
- [5] R. Dudley. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* 1, 290–330 (1967).
- [6] R. Dudley. Sample functions of the Gaussian process. *Ann. Probab.* 1, 66–103 (1973).
- [7] R. Dudley. V.N. Sudakov’s work on expected suprema of Gaussian processes. *High dimensional probability VII. Progr. Probab.* 71, 37–43. Springer (2016).
- [8] X. Fernique. Régularité des trajectoires des fonctions aléatoires gaussiennes. *École d’Été de Probabilités de St-Flour 1974. Lecture Notes in Math.* 480, 1–96. Springer (1975).
- [9] X. Fernique. Évaluations de processus gaussiens composés. Probability in Banach spaces. Lect. Notes in Math. 526, 67–83. Springer (1976).
- [10] X. Fernique. Évaluations de certaines fonctionnelles associées à des fonctions aléatoires gaussiennes. *Probab. Math. Statist.* 2, 1–29 (1981).
- [11] X. Fernique. *Fonctions aléatoires gaussiennes, vecteurs aléatoires gaussiens*. Université de Montréal, Centre de Recherches Mathématiques (1997).
- [12] A. M. Garsia, E. Rodemich, H. Rumsey Jr.. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Math. J.* 20, 565–578 (1978).
- [13] M. Ledoux. Isoperimetry and Gaussian Analysis. *École d’Été de Probabilités de St-Flour 1994. Lecture Notes in Math.* 1648, 165–294. Springer (1996).
- [14] M. Ledoux, M. Talagrand. *Probability in Banach spaces (Isoperimetry and processes)*. Ergebnisse der Mathematik und ihrer Grenzgebiete 23. Springer (1991).
- [15] M. Lifshits. *Gaussian random functions*. Kluwer (1995).
- [16] M. Lifshits. *Lectures on Gaussian processes*. SpringerBriefs Math. (2012).

- [17] J. Liu. Simple and sharp generalization bounds via lifting (2025).
- [18] M. Marcus, G. Pisier. *Random Fourier series with applications to harmonic analysis*. Ann. Math. Studies, vol. 101. Princeton Univ. Press (1981).
- [19] G. Pisier. Some applications of the metric entropy bound in harmonic analysis. *Banach Spaces, Harmonic Analysis and Probability Theory. Lecture Notes in Mathematics* 995, 123–154. Springer (1983).
- [20] C. Preston. Continuity properties of some Gaussian processes. *Ann. Math. Statist.* 43, 285–292 (1972).
- [21] D. Slepian. The one-sided barrier problem for Gaussian noise. *Bell. System Tech. J.* 41, 463–501 (1962).
- [22] V. N. Sudakov. Gaussian measures, Cauchy measures and ε -entropy. *Soviet Math. Dokl.* 10, 310–313 (1969).
- [23] V. N. Sudakov. Gaussian random processes and measures of solid angles in Hilbert spaces. *Soviet Math. Dokl.* 12, 412–415 (1971).
- [24] V. N. Sudakov. A remark on the criterion of continuity of Gaussian sample functions. *Proceedings of the Second Japan-USSR Symposium on Probability Theory. Lecture Notes in Math.* 330, 444–454. Springer (1973).
- [25] V. N. Sudakov. *Geometric problems in the theory of infinite-dimensional probability distributions (in Russian)*. Trudy Mat. Inst. Steklov 141 (1976). English translation Proc. Steklov Inst. Math., American Mathematical Society (1979).
- [26] M. Talagrand. Regularity of Gaussian processes. *Acta Math.* 159, 99–149 (1987).
- [27] M. Talagrand. *Upper and lower bounds for stochastic processes – decomposition theorems*. Ergebnisse der Mathematik und ihrer Grenzgebiete 60. Springer (2021).
- [28] R. van Handel. On the subGaussian comparison theorem (2025).