

# Some superconcentration inequalities for extrema of stationary Gaussian processes

Kevin Tanguy  
University of Toulouse, France

*Kevin Tanguy is with the Institute of Mathematics of Toulouse (CNRS UMR 5219).  
Université Paul Sabatier 31062 Toulouse, France.*

---

## Abstract

This note is concerned with concentration inequalities for extrema of stationary Gaussian processes. It provides non-asymptotic tail inequalities which fully reflect the fluctuation rate, and as such improve upon standard Gaussian concentration. The arguments rely on the hypercontractive approach developed by Chatterjee for superconcentration variance bounds. Some statistical illustrations complete the exposition.

---

## 1. Introduction

### 1.1. Convergence of extremes

As an introduction, we recall some classical facts about weak convergence of extrema of stationary Gaussian sequences and processes.

#### 1.1.1. Stationary Gaussian sequences

Let  $(X_i)_{i \geq 0}$  be a centered stationary Gaussian sequence such that  $\mathbb{E}[X_i^2] = 1$ ,  $i \geq 0$ , with covariance  $\text{Cov}(X_i, X_j) = \phi(|i - j|)$ ,  $i, j \geq 0$ , where  $\phi : \mathbb{N} \rightarrow \mathbb{R}$ . An extensive study has been developed in this setting towards the asymptotic behaviour of the maximum  $M_n = \max_{i=1, \dots, n} X_i$  (Berman, Mittal, Pickands... cf. e.g. [13]). A sample result is the following theorem (see [13]).

**Theorem 1.** *Let  $(X_i)_{i \geq 0}$  be a stationary Gaussian sequence with covariance function  $\phi$  such that  $\phi(n) \log n \rightarrow 0$  as  $n \rightarrow \infty$ . Then*

$$a_n(M_n - b_n) \xrightarrow[n \rightarrow \infty]{} G$$

in distribution where  $a_n = (2 \log n)^{1/2}$ ,

$$b_n = (2 \log n)^{1/2} - \frac{1}{2}(2 \log n)^{-1/2}(\log \log n + \log 4\pi)$$

and the random variable  $G$  has a Gumbel distribution (for all  $t \in \mathbb{R}$ ,  $\mathbb{P}(G \leq t) = e^{-e^{-t}}$ ).

## 1.1.2. Stationary Gaussian processes

Let  $(X_t)_{t \geq 0}$  be a centered stationary Gaussian process such that  $\mathbb{E}[X_t^2] = 1$ ,  $t \geq 0$ , with covariance function  $\text{Cov}(X_s, X_t) = \phi(|t - s|)$ , for  $t, s \geq 0$ , where  $\phi : [0, +\infty) \rightarrow \mathbb{R}$ . Consider two behaviors, as  $t \rightarrow 0$ , of the covariance function  $\phi$ :

$$\phi(t) = 1 - \frac{\lambda_2 t^2}{2} + o(t^2) \quad (1.1)$$

$$\phi(t) = 1 - C|t|^\alpha + o(|t|^\alpha) \quad (1.2)$$

The first case ensures that  $X_t$  is differentiable and  $\lambda_2 = -\phi''(0)$  is a spectral moment, whereas the second case concerns non-differentiable (but continuous) processes (such as the Ornstein-Uhlenbeck process). For more details about this topic, see [13]. For any  $T > 0$ , set  $M_T = \sup_{t \in [0, T]} X_t$ .

**Theorem 2.** *Let  $(X_t)_{t \geq 0}$  be a stationary Gaussian process such that  $\phi(t) \log t \rightarrow 0$  as  $t \rightarrow \infty$ . Then*

$$a_T(M_T - b_T) \xrightarrow{T \rightarrow \infty} G$$

in distribution where  $a_T = (2 \log T)^{1/2}$ ,  $b_T$  depends on the hypothesis (1.1) or (1.2) and  $G$  has a Gumbel distribution.

The aim of this note is to quantify the preceding asymptotic statements into sharp concentration inequalities fully reflecting the fluctuation rate of the maximum (respectively the supremum). Such variance bounds with the correct scale were first obtained in this context by Chatterjee in [7] (section 9.6), as part of the superconcentration phenomenon. The results presented here strengthen the variance bounds into exponential tail inequalities. One main result is the following statement.

**Theorem 3.** *Let  $(X_i)_{i \geq 0}$  be a centered stationary Gaussian sequence with covariance function  $\phi$ . Assume that  $\phi$  is non-increasing and satisfies  $\phi(1) < 1$ . Then, there exists  $\alpha = \alpha(\phi) \in (0, 1)$  and  $c = c(\phi, \alpha) > 0$  such that for all  $n \geq 2$ ,*

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| > t) \leq 6e^{-ct/\sqrt{\max(\phi(n^\alpha), 1/\log n)}}, \quad t \geq 0. \quad (1.3)$$

*Remark.* Under the hypothesis of Theorem 1,  $\max(\phi(n^\alpha), 1/\log n) = 1/\log n$  for  $n$  large enough, which is exactly the fluctuation rate. Observe furthermore that integrating (1.3), against  $t \rightarrow 2t$ , recovers the variance bounds of [7]. One last observation, under the hypothesis of Theorem 1,  $\phi(n) \rightarrow 0$  as  $n \rightarrow \infty$ . This implies that  $\sup_{n \geq 1} |\phi(n)| < 1$  (according to [13] p.86), so the hypothesis  $\phi(1) < 1$  can be removed.

It is important to compare Theorem 3 to classical Gaussian concentration (see e.g. [14]) which typically produces

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 2e^{-t^2/2}, \quad t \geq 0. \quad (1.4)$$

While of Gaussian tail, such bounds do not reflect the fluctuations of the extremum  $M_n$  of Theorem 1. Moreover, with respect to this Gaussian bound, Theorem 3 actually provides the correct tail behavior of the maximum  $M_n$  in

the form of a superconcentration inequality in accordance with the fluctuation result (see Corollary 1.9 in [9]) and the limiting Gumbel distribution, even if our method doesn't give any lower bound, (since  $\mathbb{P}(G > x) \sim e^{-x}$  as  $x \rightarrow \infty$ ). Let us also emphasize that Theorem 3 covers the classical independent case, when all the  $X_i$ 's are independent, by taking  $\phi = 0$  and provides the following concentration inequality which will be useful in the last section,

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct\sqrt{\log n}}, \quad t \geq 0. \quad (1.5)$$

Observe in addition that Theorem 3 expresses a concentration property of the maximum around its mean whereas, in the regime of the convergence of extremes of Theorem 1, the centerings are produced by explicit values  $b_n$ . Actually, up to numerical constants, the same inequalities hold true around the constant  $b_n$  instead of the mean. To this task, it is enough to prove that  $\sup_n \mathbb{E}[|a_n(M_n - b_n)|] < \infty$ . Set  $Z_n = a_n(M_n - b_n)$ . Letting  $M'_n$  be an independent copy of  $M_n$ , set similarly  $Z'_n = a_n(M'_n - b_n)$ . Now, integrating (1.3),  $\sup_n \mathbb{E}[|a_n(M_n - \mathbb{E}[M_n])|] < \infty$ . Hence  $\sup_n \mathbb{E}[|Z_n - Z'_n|] < \infty$  from which it easily follows that  $\sup_n \mathbb{E}[|Z_n|] < \infty$ .

The next statement is the analogue of Theorem 3 for stationary Gaussian processes, suitably quantifying Theorem 2. It is presented in the more general context of a centered stationary Gaussian field  $(X_t)_{t \in \mathbb{R}^d}$  such that  $\mathbb{E}[X_t^2] = 1$ ,  $t \in \mathbb{R}^d$ , with covariance function  $\phi$ . According to [7], very little seems actually to be known on the asymptotic fluctuations of the supremum of stationary Gaussian processes indexed by  $\mathbb{R}^d$  when the dimension  $d$  is greater than two. Some recent specific results are available for Gaussian fields with logarithmic correlation (see e.g. [1, 11, 16, 10] and [12] for an overview on the subject). However, extending similarly the variance bounds of [7], we obtain a concentration inequality for the supremum of a Gaussian field over a subset  $A$  of  $\mathbb{R}^d$ . The case  $d = 1$  thus covers Theorem 2.

**Theorem 4.** *Let  $(X_t)_{t \in \mathbb{R}^d}$  be a stationary Gaussian Euclidean field with covariance function  $\phi$ . Assume that  $t \mapsto \phi(t)$  is non-increasing and  $\phi(1) < 1$ . If  $A$  is a subset of  $\mathbb{R}^d$ , denote by  $N(A) = N(A, 1)$  the minimal number of balls with radius 1 needed to cover  $A$ . Set  $M(A) = \sup_{s \in A} X_s$ . Then, there exist  $C = C(\phi, d) > 0$  and  $c = c(\phi, d) > 0$  only depending on  $\phi$  and  $d$  such that, for all  $A \subset \mathbb{R}^d$  with  $N(A) > 1$ ,*

$$\mathbb{P}(|M_A - \mathbb{E}[M_A]| \geq t) \leq 6e^{-ct/\sqrt{K_{N(A)}}}, \quad t \geq 0,$$

where

$$K_{N(A)} = \max(\phi(N(A)^C), 1/\log N(A)).$$

*Remark.* When  $d = 1$  and  $A = [0, T]$  for some  $T > 0$ ,  $N(A) = T$  and under the hypotheses of Theorem 2,  $K_{N(A)} = 1/\log T$  for  $T$  large enough.

The proofs of Theorem 3 and 4 are based on the hypercontractive approach developed by Chatterjee [7] towards variance bounds. The task will be to adapt the argument to reach exponential tail inequalities at the correct fluctuation regime. This will be achieved via the corresponding variance bound at the level of Laplace transforms.

The paper is organized as follows. In the next section, we present the semigroup tools used to prove the main results. The third section is devoted to the proof of Theorems 3 and 4. In Section 4, we present some illustrations to further Gaussian models. In the final section, we present an illustration in statistical testing.

## 2. Framework and tools

### 2.1. Notation

For any  $n \geq 1$ , denote by  $X = (X_1, \dots, X_n)$  a centered Gaussian vector, with maximum  $M_n = \max_{i=1, \dots, n} X_i$ . Set  $I = \operatorname{argmax}_{i=1, \dots, n} X_i$ .

The main technical part of this work is provided by the following extension of Chatterjee's approach [7] adapted to exponential concentration bounds. The very basis of the argument is the hypercontractivity of the Ornstein-Uhlenbeck semigroup together with the exponential version of Poincaré inequality (cf. [5, 4, 7]).

**Theorem 5.** *Let  $X = (X_1, \dots, X_n)$  be a centered Gaussian vector with covariance matrix  $\Gamma$ . Assume that for some  $r_0 \geq 0$ , there exists a non trivial covering  $\mathcal{C}(r_0)$  of  $\{1, \dots, n\}$  verifying the following properties:*

- *for all  $i, j \in \{1, \dots, n\}$  such that  $\Gamma_{ij} \geq r_0$ , there exists  $D \in \mathcal{C}(r_0)$  such that  $i, j \in D$ ;*
- *there exists  $C \geq 1$  such that, a.s.,  $\sum_{D \in \mathcal{C}(r_0)} 1_{\{I \in D\}} \leq C$ .*

Let  $\rho(r_0) = \max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D)$ .

Then, for every  $\theta \in \mathbb{R}$ ,

$$\operatorname{Var} \left( e^{\theta M_n / 2} \right) \leq C \frac{\theta^2}{4} \left( r_0 + \frac{1}{\log(1/\rho(r_0))} \right) \mathbb{E} \left[ e^{\theta M_n} \right]. \quad (2.1)$$

In particular,

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct/\sqrt{K_{r_0}}}, \quad t \geq 0,$$

where  $K_{r_0} = \max \left( r_0, \frac{1}{\log(1/\rho(r_0))} \right)$  and  $c > 0$  is a universal constant.

*Remark.* By monotone convergence, we can obtain the same result for a supremum instead of a maximum. This fact will be useful to obtain an application of Theorem 5 for a Gaussian Euclidean field.

*Proof.* As announced, the scheme of proof follows [7]. The starting point is the representation formula

$$\operatorname{Var}(f) = \int_0^\infty e^{-t} \mathbb{E} [\nabla f \cdot P_t \nabla f] dt \quad (2.2)$$

for the variance of a (smooth) function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  along the Ornstein-Uhlenbeck semigroup  $(P_t)_{t \geq 0}$ . General references on the Ornstein-Uhlenbeck

semigroup and more general Markov semigroups, as well as such semigroup interpolation formulas are [5, 4].

Following [7], given a centered Gaussian vector  $X$  with covariance matrix  $\Gamma = (\Gamma_{ij})_{1 \leq i, j \leq n}$ , apply (2.2) to (a smooth approximation of)  $f = e^{\theta M/2}$  where  $\theta \in \mathbb{R}$ ,  $M(x) = \max_{i=1, \dots, n} (Bx)_i$  and  $\Gamma = B^t B$ . It yields that, for all  $\theta \in \mathbb{R}$ ,

$$\text{Var} \left( e^{\theta M_n/2} \right) = \frac{\theta^2}{4} \int_0^\infty e^{-t} \mathbb{E} \left[ \sum_{i, j=1}^n \Gamma_{ij} 1_{\{X \in A_i\}} e^{\theta M_n/2} 1_{\{X^t \in A_j\}} e^{\theta M_n^t/2} \right] dt.$$

Here  $\{X \in A_i\} = \{I = i\}$ ,  $i = 1, \dots, n$ , and  $X^t = e^{-t}X + \sqrt{1 - e^{-2t}}Y$ , where  $Y$  is an independent copy of  $X$ , with corresponding maximum  $M_n^t$ . This process shares the same hypercontractivity property as the Ornstein-Uhlenbeck process (cf. [7]). Indeed, assuming without loss of generality that  $B$  is invertible (otherwise restrict to a subspace of  $\mathbb{R}^n$ ), define the semigroup  $(Q_t)_{t \geq 0}$  by

$$Q_t f(x) = \mathbb{E} \left[ f \left( e^{-t}x + \sqrt{1 - e^{-2t}}Y \right) \right], \quad x \in \mathbb{R}^n$$

for  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  smooth enough. Let  $Z$  be a standard Gaussian vector on  $\mathbb{R}^n$  so that  $Y = BZ$  in law. Thus, by defining  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  as  $g(x) = f(Bx)$ , we have, for all  $x \in \mathbb{R}^n$ ,

$$\begin{aligned} Q_t f(x) &= \mathbb{E} \left[ f \left( e^{-t}x + \sqrt{1 - e^{-2t}}Y \right) \right] \\ &= \mathbb{E} \left[ f \left( e^{-t}BB^{-1}x + \sqrt{1 - e^{-2t}}BZ \right) \right] \\ &= \mathbb{E} \left[ g \left( e^{-t}B^{-1}x + \sqrt{1 - e^{-2t}}Z \right) \right] = P_t g(B^{-1}x) \end{aligned}$$

Therefore, by the hypercontractive property of  $(P_t)_{t \geq 0}$  with the same exponents  $p$  and  $q$  as the Ornstein-Uhlenbeck semigroup,

$$\mathbb{E} [|Q_t f(X)|^q]^{1/q} = \|P_t g\|_q \leq \|g\|_p = \mathbb{E} [|f(X)|^p]^{1/p}, \quad (2.3)$$

where  $\|\cdot\|_r$ , for any  $r \geq 1$ , stands for the  $L^r$ -norm with respect to the standard Gaussian measure on  $\mathbb{R}^n$ .

Now, for every  $t \geq 0$ , with  $I^t = \text{argmax}_{i=1, \dots, n} X_i^t$ ,

$$\begin{aligned} \mathcal{I} &= \mathbb{E} \left[ \sum_{i, j=1}^n \Gamma_{ij} 1_{\{X \in A_i\}} e^{\theta M_n/2} 1_{\{X^t \in A_j\}} e^{\theta M_n^t/2} \right] \\ &= \mathbb{E} \left[ e^{\theta(M_n + M_n^t)/2} \sum_{i, j=1}^n \Gamma_{ij} 1_{\{I=i\}} 1_{\{I^t=j\}} \right] \\ &= \mathbb{E} \left[ e^{\theta(M_n + M_n^t)/2} \Gamma_{I I^t} \right] \\ &= \sum_{k \geq 0} \mathbb{E} \left[ e^{\theta(M_n + M_n^t)/2} \Gamma_{I I^t} 1_{\{2^{-k-1} \leq \Gamma_{I I^t} \leq 2^{-k}\}} \right]. \end{aligned}$$

To ease the notation, denote  $\Gamma_{I I^t}$  by  $\Gamma$  and set  $F^t = M_n + M_n^t$ . Let  $k_0 =$

$\min\{k \geq 0; r_0 \leq 2^{-k-1}\}$ . Cutting the preceding sum into two parts, we obtain

$$\begin{aligned} \mathcal{I} &\leq \sum_{k=0}^{k_0} 2^{-k} \mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{\Gamma \geq r_0\}} \right] + \sum_{k=k_0+1}^{\infty} 2^{-k} \mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{2^{-k-1} \leq \Gamma \leq 2^{-k}\}} \right] \\ &\leq 2 \sum_{D \in \mathcal{C}(r_0)} \mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{I, I^t \in D\}} \right] + \sum_{k \geq 0} r_0 \mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{2^{-k-1} \leq \Gamma \leq 2^{-k}\}} \right] \\ &= 2 \sum_{D \in \mathcal{C}(r_0)} \mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{I, I^t \in D\}} \right] + r_0 \mathbb{E} \left[ e^{\theta F^t/2} \right]. \end{aligned}$$

By the Cauchy-Schwarz inequality and Gaussian rotational invariance,

$$\mathbb{E} \left[ e^{\theta F^t/2} \right] \leq \mathbb{E} \left[ e^{\theta M_n} \right].$$

On the other hand, by Hölder's inequality with  $\frac{1}{p} + \frac{1}{q} = 1$ ,

$$\begin{aligned} \mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{I, I^t \in D\}} \right] &= \mathbb{E} \left[ e^{\theta M_n/2} \mathbf{1}_{\{I \in D\}} e^{\theta M_n^t/2} \mathbf{1}_{\{I^t \in D\}} \right] \\ &\leq \mathbb{E} \left[ e^{\theta p M_n/2} \mathbf{1}_{\{I \in D\}} \right]^{1/p} \mathbb{E} \left[ e^{\theta q M_n^t/2} \mathbf{1}_{\{I^t \in D\}} \right]^{1/q}. \end{aligned}$$

Next, by the hypercontractivity property (2.3),

$$\mathbb{E} \left[ e^{\theta q M_n^t/2} \mathbf{1}_{\{I^t \in D\}} \right]^{1/q} \leq \mathbb{E} \left[ e^{\theta p M_n/2} \mathbf{1}_{\{I \in D\}} \right]^{1/p}$$

provided that  $e^{2t} = \frac{q-1}{p-1}$ , that is  $p = 1 + e^{-t} < 2$ . As a consequence,

$$\mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{I, I^t \in D\}} \right] \leq \mathbb{E} \left[ e^{\theta p M_n/2} \mathbf{1}_{\{I \in D\}} \right]^{2/p}$$

and by a further use of Hölder's inequality on the right hand side of the preceding inequality,

$$\mathbb{E} \left[ e^{\theta F^t/2} \mathbf{1}_{\{I, I^t \in D\}} \right] \leq \mathbb{P}(I \in D)^{\frac{2-p}{p}} \mathbb{E} \left[ e^{\theta M_n} \mathbf{1}_{\{I \in D\}} \right].$$

Combining the preceding with the assumptions, we get that

$$\mathcal{I} \leq \left( r_0 + C \rho(r_0)^{\frac{2-p}{p}} \right) \mathbb{E} \left[ e^{\theta M_n} \right].$$

Finally,

$$\text{Var} \left( e^{\theta M_n/2} \right) \leq C \frac{\theta^2}{4} \left( r_0 + \int_0^\infty e^{-t} \rho(r_0)^{\tanh(t/2)} dt \right) \mathbb{E} \left[ e^{\theta M_n} \right].$$

The announced inequality (2.1) then follows since

$$\int_0^\infty e^{-t} (\rho(r_0))^{\tanh(t/2)} dt \leq \frac{1}{\log(1/\rho(r_0))}.$$

To end the proof of Theorem 5 and produce exponential tail bounds, we combine (2.1) with the following lemma (see [14] p.50).

**Lemma 6.** Let  $Z$  be a random variable and  $K > 0$ . Assume that for every  $|\theta| \leq 2/\sqrt{K}$ ,

$$\text{Var}\left(e^{\theta Z/2}\right) \leq K \frac{\theta^2}{4} \mathbb{E}\left[e^{\theta Z}\right]. \quad (2.4)$$

Then,

$$\mathbb{P}(|Z - \mathbb{E}[Z]| > t) \leq 6e^{-ct/\sqrt{K}},$$

for all  $t \geq 0$ , with  $c > 0$  a numerical constant.

*Remark.* Such a lemma has been used in prior works [6, 8] in order to obtain exponential concentration tail bounds for first passage percolation models.  $\square$

### 3. Proofs of the main results

#### 3.1. Proof of Theorem 3

The task is to show that the hypothesis of Theorem 5 are satisfied. For  $0 < \alpha < 1$ , choose  $r_0 = \phi(\lfloor n^\alpha \rfloor)$  and  $\mathcal{C}(r_0) = \{D_1, D_2, \dots\}$  where  $D_1 = \{1, \dots, 2\lfloor n^\alpha \rfloor\}$ ,  $D_2 = \{\lfloor n^\alpha \rfloor, \dots, 3\lfloor n^\alpha \rfloor\}$  and so on, where  $\lfloor \cdot \rfloor$  denote the integer part of a real number.

It is easily seen that this covering satisfies the hypothesis of Theorem 5. Indeed, take  $i, j \in \{1, \dots, n\}$  such that  $i < j$ . By stationarity,  $\text{Cov}(X_{i+1}, X_j) = \phi(j - i)$ . So if  $\phi(j - i) \geq r_0 = \phi(\lfloor n^\alpha \rfloor)$ , since  $\phi$  is non-increasing, we must have  $j - i \leq \lfloor n^\alpha \rfloor$ . By construction, any index  $i$  is at most in three different elements  $D$  of  $\mathcal{C}(r_0)$ . So  $i$  belongs to some  $D_{i_1}$ ,  $D_{i_2}$  and  $D_{i_3}$  with  $i_1 < i_2 < i_3$  and  $j$  belong to  $D_{j_1}$ ,  $D_{j_2}$  and  $D_{j_3}$  with  $j_1 < j_2 < j_3$ . Since  $j - i \leq n^\alpha$  and the length of any  $D$  in  $\mathcal{C}(r_0)$  is  $2\lfloor n^\alpha \rfloor$ , there exists  $s \in \{1, 2, 3\}$  such that  $D_{i_s} = D_{j_s}$  (draw a picture) and we can choose  $D = D_{i_s} \in \mathcal{C}(r_0)$ .

Clearly  $\sum_{D \in \mathcal{C}(r_0)} \mathbb{1}_{\{I \in D\}} \leq C$ , where  $C$  is a universal constant (say  $C = 3$ ). To end the proof, we have to bound  $\max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D)$ . Let  $D \in \mathcal{C}(r_0)$ , and to ease the notation, consider  $D = \{1, \dots, 2\lfloor n^\alpha \rfloor\}$  (the important feature is that  $D$  contains  $2\lfloor n^\alpha \rfloor$  elements). Then

$$\mathbb{P}(I \in D) = \sum_{i=1}^{2\lfloor n^\alpha \rfloor} \mathbb{P}(I = i).$$

By standard Gaussian concentration (see [14] or the appendix of [7]), for all  $i \in D$ ,

$$\begin{aligned} \mathbb{P}(I = i) &= \mathbb{P}(X_i = M_n) \\ &\leq \mathbb{P}(X_i \geq t) + \mathbb{P}(M_n \leq t) \\ &\leq 2e^{-\mathbb{E}[M_n]^2/2}. \end{aligned}$$

The final step of the argument is to achieve a lower bound on  $\mathbb{E}[M_n]$ . To this task, we make use of Sudakov's minoration (see [3] or [15]). Letting  $i, j \in D$ ,  $i \neq j$ ,

$$\mathbb{E}[(X_i - X_j)^2] = 2 - 2\phi(|j - i|) \geq 2(1 - \phi(1)) = \delta$$

since  $\phi(|j-i|) \leq \phi(1)$ . By the assumption on  $\phi(1)$ ,  $\delta > 0$ , and thus by Sudakov's minoration there exists  $c > 0$  (independent of  $n$ ) such that  $\mathbb{E}[M_n] \geq c\delta\sqrt{\log n}$ . As a consequence of the preceding,

$$\mathbb{P}(I = i) \leq \frac{2}{n^{(c\delta)^2/2}} = \frac{2}{n^\epsilon}.$$

Hence

$$\mathbb{P}(I \in D) \leq \frac{4}{n^{\epsilon-\alpha}} = \frac{4}{n^\eta}$$

where  $\alpha$  is such that  $\eta = \epsilon - \alpha > 0$ , namely  $\alpha < (c\delta)^2/2 < 1$ .

The hypotheses of Theorem 5 are therefore fulfilled with  $\rho(r_0) \leq 4/n^\eta$  concluding the proof of Theorem 3.

### 3.2. Proof of Theorem 4

We again follow Chatterjee's proof of variance bounds for Gaussian Euclidean fields (see Theorem 9.12 in [7]). For any Borel set  $A \subset \mathbb{R}^d$ , set  $m(A) = \mathbb{E}[M(A)] = \mathbb{E}[\sup_{s \in A} X_s]$ . Under the hypothesis of the theorem, it is proved in [7] that for any  $A \subset \mathbb{R}^d$  such that  $N(A) > 1$ ,

$$c_1(\phi, d)\sqrt{\log N(A)} \leq m(A) \leq c_2(\phi, d)\sqrt{\log N(A)}$$

where  $c_1 = c_1(\phi, d)$  and  $c_2 = c_2(\phi, d)$  are positive constants that depend only on the covariance  $\phi$  and the dimension  $d$  (and not on the subset  $A$ ). Set then

$$s_0 = N(A)^{\frac{1}{8}(c_1/c_2)^2},$$

and assume that  $N(A)$  is large enough so that  $s_0 > 2$ . Let  $r_0 = \phi(s_0)$ . Take a maximal  $s_0$ -net of  $A$  (i.e. a set of points that are mutually separated from each other by a distance strictly greater than  $s_0$  and is maximal with respect to this property), and let  $\mathcal{C}(r_0)$  be the set of  $2s_0$ -ball around the points in the net. In fact,  $\mathcal{C}(r_0)$  is a covering of  $A$  satisfying the condition of Theorem 5. With this construction, it is easily seen that  $\sum_{D \in \mathcal{C}(r_0)} \mathbb{1}_{\{I \in D\}} \leq C$  for some  $C \geq 1$  and

$$\max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D) \leq \frac{1}{N(A)^{C(\phi, d)}}.$$

The conclusion then directly follows from Theorem 5 with  $\rho(r_0) \leq 1/N(A)^{C(\phi, d)}$ .

## 4. Other Gaussian models

In [7], Chatterjee exhibits different models in order to illustrate the so-called superconcentration phenomenon. For each of them, he managed to reach a better variance bound than the one produced by standard Gaussian concentration. We present for some of these models an exponential version.

**Proposition 7.** *For any  $n \geq 2$ , let  $X = (X_1, \dots, X_n)$  be a centered Gaussian vector with  $\mathbb{E}[X_i]^2 = 1$  and  $\mathbb{E}[X_i X_j] \leq \epsilon$  for all  $i, j = 1, \dots, n$  for some  $\epsilon > 0$ . Then*

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct/\sqrt{K_n}}, \quad t \geq 0,$$

where  $K_n = \max(\epsilon, 1/\log n)$ .



*Proof.* Take  $r_0 > \epsilon$  and define  $\mathcal{C}(r_0) = \{\{1\}, \dots, \{N\}\}$ . It is easy to see that the covering  $\mathcal{C}(r_0)$  enters the setting of Theorem 5. Furthermore,  $\sum_{D \in \mathcal{C}(r_0)} 1_{\{I \in D\}} = 1$  and  $\max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D) \leq 1/n^\eta$  for some  $\eta > 0$ . The conclusion follows.  $\square$

As another instance, consider the Gaussian field on  $\{-1, 1\}^n$ ,  $n \geq 1$ , defined in the following way. Let  $X_1, \dots, X_n$  be independent identically distributed standard normals, and define  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  as

$$f(\sigma) = \sum_{i=1}^n X_i \sigma_i.$$

**Corollary 8.** *There exists  $S = (\sigma^1, \dots, \sigma^N) \in \{-1, 1\}^n$  for some  $N$ , such that, for all  $\sigma \neq \sigma' \in S$ ,  $|\sigma \cdot \sigma'| \leq Cn^{2/3}$  and*

$$\mathbb{P}\left(\left|\max_{\sigma \in S} f(\sigma) - \mathbb{E}\left[\max_{\sigma \in S} f(\sigma)\right]\right| \geq t\right) \leq 6e^{-ct/n^{2/3}}, \quad t \geq 0.$$

*Proof.* Chatterjee proved in [7] the existence of  $S \in \{-1, 1\}^n$  on which  $|\sigma \cdot \sigma'| \leq n^{2/3}$ . Since,  $\text{Cov}(f(\sigma), f(\sigma')) = \sigma \cdot \sigma'$ , it suffices to apply Proposition 7 to  $f|_S$  and  $\epsilon = n^{2/3}$ .  $\square$

## 5. Statistical testing

In this last section, we illustrate the preceding results on a statistical testing problem analyzed in [2]. There, the authors used standard Gaussian concentration to obtain an acceptance region for their test. Using some of the material developed here, the conclusion may be reinforced in some instances in the form of a superconcentration bound.

Following the framework and notation of [2], we observe an  $n$  dimensional Gaussian vector  $X = (X_1, \dots, X_n)$  and raise the following hypothesis problem:

- Under the null hypothesis  $H_0$ , the components of  $X$  are independent identically distributed standard normal random variables. Denote then by  $\mathbb{P}_0$  and  $\mathbb{E}_0$  respectively the underlying probability measure and expectation under  $H_0$ .
- To describe the alternative hypothesis  $H_1$ , consider a class  $\mathcal{C} = \{S_1, \dots, S_N\}$  of  $N$  sets of indices such that  $S_k \subset \{1, \dots, n\}$  for all  $k = 1, \dots, N$ . Under  $H_1$ , there exists an  $S \in \mathcal{C}$  such that

$$X_i \text{ has distribution } \begin{cases} \mathcal{N}(0, 1) & \text{if } i \in S^c, \\ \mathcal{N}(\mu, 1) & \text{if } i \in S, \end{cases}$$

where  $\mu > 0$  is a positive parameter. The components of  $X$  are independent under  $H_1$  as well. For any  $S \in \mathcal{C}$ , denote then by  $\mathbb{P}_S$  and  $\mathbb{E}_S$  respectively the underlying probability measure and expectation under  $H_1$ . We will also assume that every  $S \in \mathcal{C}$  has the same cardinality  $|S| = K$ .

We recall that a test is a binary-valued function  $f : \mathbb{R}^n \rightarrow \{0, 1\}$ . If  $f(X) = 0$ , the test accepts the null hypothesis, otherwise  $H_0$  is rejected. As in [2], consider the risk of a test  $f$  measured by

$$R(f) = \mathbb{P}_0(f(X) = 1) + \frac{1}{N} \sum_{S \in \mathcal{C}} \mathbb{P}_S(f(X) = 0).$$

The authors of [2] are interested in determining, or at least estimating, the value of  $\mu$  under which the risk can be made small. Among others results, they used a test based on maxima, called the scan-test, for which they showed that

$$f(X) = 1 \iff 2 \max_{S \in \mathcal{C}} X_S \geq \mu K + \mathbb{E}_0 \left[ \max_{S \in \mathcal{C}} X_S \right]$$

where  $X_S = \sum_{i \in S} X_i$  for some  $S \in \mathcal{C}$ . They obtain the following result.

**Proposition 9.** *The risk of the maximum test  $f$  satisfies  $R(f) \leq \delta$  whenever*

$$\mu \geq \frac{1}{K} \mathbb{E}_0 \left[ \max_{S \in \mathcal{C}} X_S \right] + 2 \sqrt{\frac{2}{K} \log \frac{2}{\delta}}.$$

Together with (1.5), this bound may be improved in accordance with the correct magnitude of the maximum of a standard Gaussian vector.

**Proposition 10.** *For any  $n \geq 2$ , the risk of the maximum test  $f$  satisfies  $R(f) \leq \delta$  whenever*

$$\mu \geq \frac{1}{K} \mathbb{E}_0 \left[ \max_{S \in \mathcal{C}} X_S \right] + \log \frac{6}{\delta} \times \frac{2}{c \sqrt{K \log N}}$$

where  $c > 0$  is a universal constant.

*Proof.* We follow the proof of [2] to get that both, for every  $t \geq 0$ ,

$$\mathbb{P}_0 \left( \max_{S \in \mathcal{C}} X_S \geq \mathbb{E}_0 \left[ \max_{S \in \mathcal{C}} X_S \right] + t \right) \leq 3e^{-ct\sqrt{\log(N)/K}}$$

and, for any  $S \in \mathcal{C}$ ,

$$\begin{aligned} \mathbb{P}_S \left( \max_{S' \in \mathcal{C}} X_{S'} \leq \mu K - t \right) &\leq \mathbb{P}_S \left( \max_{S' \in \mathcal{C}} X_{S'} \leq \mathbb{E}_{S'} \left[ \max_{S' \in \mathcal{C}} X_{S'} \right] - t \right) \\ &\leq 3e^{-ct\sqrt{\log(N)/K}} \end{aligned}$$

since, under  $H_1$ , for a fixed  $S \in \mathcal{C}$ ,  $\mu K = \mathbb{E}_S [X_S] \leq \mathbb{E}_S [\max_{S' \in \mathcal{C}} X_{S'}]$  and  $X_S$  has the same law as any  $X_{S'}$ . Set then  $t$  such that

$$2t = \mu K - \mathbb{E} \left[ \max_{S \in \mathcal{C}} X_S \right]$$

which is positive according to the hypothesis. With the previous inequalities, we obtain a bound on the risk

$$R(f) \leq 6e^{-ct\sqrt{\log(N)/K}}$$

which may then be turned into  $R(f) \leq \delta$  as in the statement.  $\square$

The new bound of Proposition 10 is as good as the one of Proposition 9 when  $\delta \simeq 1/N^\alpha$  for some  $\alpha > 0$ . It is better than Proposition 9 when  $\delta \simeq 1/\log^\alpha(N)$  with  $\alpha > 0$ . Nevertheless, Proposition 9 is better than Proposition 10 when  $\delta \simeq e^{-N^\alpha}$ , for  $\alpha > 0$ . However, it might not be relevant to consider such small risk.

*Aknowledgment.* I thank my Ph.D advisor M. Ledoux for introducing this problem to me and for fruitful discussions.

- [1] J. A. Acosta. Tightness of the recentered maximum of log-correlated gaussian field. *Electronic Journal of Probability, Vol. 19, Article 90*, 2014.
- [2] L. Addario-Berry, N. Broutin, L. Devroye, and G. Lugosi. On combinatorial testing problems. *Ann. Statist. 38 (2010), no. 5*, 2010.
- [3] R. J. Adler. *An introduction to continuity, extrema and related topics for general Gaussian processes*. Institut of Mathematical Statistics Lectures Notes - Monograph Series, 12, 1990.
- [4] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*. Grundlehren der Mathematischen Wissenschaften, 348, 2014.
- [5] D. Bakry, M. Ledoux, and L. Saloff-Coste. *Markov semigroups at Saint-Flour*. Probability at Saint-Flour. Springer, Heidelberg, 2012.
- [6] M. Benaïm and R. Rossignol. Exponential concentration for first passage percolation through modified poincaré inequalities. *Ann. Inst. Henri Poincaré Probab. Stat. 44 (2008), no. 3*, 2008.
- [7] S. Chatterjee. *Superconcentration and related topics*. Springer, 2014.
- [8] M. Damron, J. Hanson, and P. Sosoe. Subdiffusive concentration in first-passage percolation. *Electronic Journal of Probability*, 2014.
- [9] J. Ding, R. Eldan, and A. Zhai. On multiple peaks and moderate deviations for supremum of gaussian field. *Preprint, <http://arxiv.org/abs/1311.5592v2>*, 2013.
- [10] J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated gaussian fields. *Preprint, <http://arxiv.org/abs/1503.04588>*, 2015.
- [11] B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Critical gaussian multiplicative chaos : Convergence of the derivative martingale. *Ann. Prob. 42, no. 5,,* 2014.
- [12] B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Log-correlated gaussian fields : an overview. *Preprint, <http://arxiv.org/abs/1407.5605>*, 2014.

- [13] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics., 1983.
- [14] M. Ledoux. *The concentration of measure phenomenon*. Mathematical Surveys and Monographs, 89., 2001.
- [15] M. Ledoux and M. Talagrand. *Probability in Banach spaces. Isoperimetry and processes*. Springer-Verlage, Berlin, 2011.
- [16] T. Madaule. Maximum of a log-correlated gaussian field. *Arxiv (preprint)*, 2014.