The discrepancy between min-max statistics of Gaussian and Gaussian-subordinated matrices

Giovanni Peccati [*] Nicola Turchi [*†]

December 5, 2022

Abstract

We compute quantitative bounds for measuring the discrepancy between the distribution of two min-max statistics involving either one pair of Gaussian random matrices, or one Gaussian and one Gaussian-subordinated random matrix. In the fully Gaussian setup, our approach allows us to recover quantitative versions of well-known inequalities by Gordon (1985, 1987, 1992), thus generalizing the quantitative version of the Sudakov-Fernique inequality deduced in Chatterjee (2005). On the other hand, the Gaussian-subordinated case yields generalizations of estimates by Chernozhukov et al. (2015) and Koike (2019). As an application, we establish fourth moment bounds for matrices of multiple stochastic Wiener-Itô integrals, that we illustrate with an example having a statistical flavour.

Keywords: Min-max Statistics; Random Matrices; Gaussian Vectors; Gaussian Fields; Gaussian Analysis; Probabilistic Approximations; Malliavin Calculus.

AMS 2020 Classification: 60F05; 60G15; 60G70; 60H05; 60H07.

1 Introduction

1.1 Overview of our contributions

Within the theory of Gaussian processes an important role is played by inequalities of the *Sudakov-Fernique type*. These results consist in comparisons between extremal value statistics of two distinct Gaussian objects, for example, the maxima of two Gaussian random vectors with different variances. The classical Sudakov-Fernique inequality states that, if $(X_1, \dots, X_n)$ and $(Y_1, \dots, Y_n)$ are centered Gaussian random vectors such that $\mathbf{E}\big((X_i - X_j)^2\big) \le \mathbf{E}\big((Y_i - Y_j)^2\big)$ for all pairs of indices, then

$$\mathbf{E}\Big(\max_{i\in\{1,\dots,n\}} X_i\Big) \le \mathbf{E}\Big(\max_{i\in\{1,\dots,n\}} Y_i\Big). \tag{1}$$

The inequality (1) first appeared in the works of Sudakov [21, 22] and Fernique [8], and a proof is also due to Alexander [2]. Vitale [23] was able to remove the zero-mean assumption with the

[*] *Department of Mathematics, University of Luxembourg.*
[†] *Department of Mathematics and its Applications, University of Milano-Bicocca.*

weaker condition that $\mathbf{E}\big((X_1,\dots,X_n)\big)=\mathbf{E}\big((Y_1,\dots,Y_n)\big)$. Similar inequalities have been studied by Gordon [9, 10] and Kahane [12] in the more general setting of higher-dimensional tensors $(X_{i_1,i_2,\dots,i_d})$, where the maximum of the classic Sudakov-Fernique inequality is replaced by quantities of the type $\min_{i_1}\max_{i_2}\cdots X_{i_1,i_2,\dots,i_d}$. In [11] Gordon also studied the comparison between the sums of the first k ordered statistics of two Gaussian random vectors. See e.g. [1, 18], and the references therein, for a sample of applications of estimates directly related to (1) — ranging from the geometry of Gaussian fields, to stochastic differential equations and statistical mechanics.

Whereas the aforementioned results are mostly qualitative, in the reference [3] one can find a quantitative counterpart to (1), using integration by parts formulas (see also [1, Section 2.3]). More precisely, in [3] it is established that, if the two Gaussian vectors X and Y have the same mean, then

$$\Big|\mathbf{E}\Big(\max_{i\in\{1,\dots,n\}} X_i\Big)-\mathbf{E}\Big(\max_{i\in\{1,\dots,n\}} Y_i\Big)\Big|\le\sqrt{\max_{i,j}\big|\mathbf{E}\big((X_i-X_j)^2\big)-\mathbf{E}\big((Y_i-Y_j)^2\big)\big|\log n}. \tag{2}$$

In the first part of the present work, we extend the study of quantitative bounds of the type (2) to the setting considered by Gordon [9, 10, 11] of min-max statistics of Gaussian random matrices. To motivate the reader, we report below one of our principal contributions on the matter — see Section 2 for a full statement and for its proof.

Theorem. *Let (X_{i_1,i_2}) and (Y_{j_1,j_2}) be two $n\times m$ Gaussian random matrices with the same expectation. Then*

$$\Big|\mathbf{E}\big(\min_{i_1}\max_{i_2} X_{i_1,i_2}\big)-\mathbf{E}\big(\min_{i_1}\max_{i_2} Y_{i_1,i_2}\big)\Big|\le\sqrt{\max_{i_1,i_2,j_1,j_2}\big|\mathbf{E}\big((X_{i_1,i_2}-X_{j_1,j_2})^2\big)-\mathbf{E}\big((Y_{i_1,i_2}-Y_{j_1,j_2})^2\big)\big|}$$
$$\times\Big[\sqrt{\Big(1-\frac{1}{n}\Big)\Big(2-\frac{1}{m}\Big)\log n}+\sqrt{\Big(1-\frac{1}{m}\Big)\log m}\Big].$$

We will see that our techniques also allow one to recover as special cases virtually all comparison statements for min-max statistics (and their generalizations, like sums of order statistics) proved in [9, 10, 11]. In particular, an interesting application of our findings is the comparison between the order statistics of two Gaussian random vectors, once they are regarded as min-max of particular Gaussian random matrices. For instance, we can show that

$$\big|\mathbf{E}(X_{(n-1)})-\mathbf{E}(Y_{(n-1)})\big|\le(\sqrt{2}+1)\sqrt{\max_{i,j}\big|\mathbf{E}\big((X_i-X_j)^2\big)-\mathbf{E}\big((Y_i-Y_j)^2\big)\big|\log n},$$

where the index $(n-1)$ indicates the second maximum coordinate of a n-dimensional vector. See Corollary 1 below.

It is apparent that bounds such as the ones described above, involving only first moments of extremal statistics, cannot completely describe the similarity between the distributions of the involved quantities. To overcome this shortcoming, Chernozhukov, Chetverikov, and Kato have established in references [4, 5] (which crucial installments of the so-called *CCK theory*) bounds on the *Kolmogorov distance* between the laws of the maxima of two Gaussian random vectors, so as to give a more precise description of their closeness. In order to achieve their results, the authors employ some novel anti-concentration inequalities for the maximum statistic of a Gaussian random process. These contributions have been recently extended by Koike in [13] — to which we refer the reader for a more comprehensive overview of the CCK theory — where bounds are established on the discrepancy between the maxima of a Gaussian random vector and a smooth Gaussian-subordinated random element.

In Section 3 (see, in particular, Theorem 2) we generalize some of the results from [4, 5, 13] to the aforementioned setting of min-max statistics of random matrices: in particular, we derive a bound for the Kolmogorov distance between the laws of the min-max statistics of two random matrices, one of which is Gaussian. In order to do so, we need to recover some new anti-concentration inequalities suitable for our purposes; see for instance Proposition 9 and Lemma 10 below.

One important by-product of our findings are estimates involving matrices of *multiple Wiener-Itô integrals* (see e.g. [20, Chapter 2]), to which we will devote Section 4. As an example of application of such estimates, suppose that every entry (i_1, i_2) of an $n \times m$ matrix is given by the following random quadratic form

$$F_{i_1,i_2} = \sum_{u,v=1}^{d} A_{i_1,i_2}(u,v)\xi_u\xi_v - \mathbf{E}\Big(\sum_{u,v=1}^{d} A_{i_1,i_2}(u,v)\xi_u\xi_v\Big),$$

where $A_{i_1,i_2}(\cdot,\cdot)$ is a real-valued symmetric matrix for all (i_1, i_2) and $(\xi_1, \ldots, \xi_d)$ is a d-dimensional Gaussian random vector. If (X_{i_1,i_2}) is a $n \times m$ centered Gaussian random matrix with the same covariance structure as (F_{i_1,i_2}), then one has that

$$d_{\mathrm{Kol}}(\min_{i_1}\max_{i_2} F_{i_1,i_2}, \min_{i_1}\max_{i_2} X_{i_1,i_2}) \leq C \max_{i_1,i_2}\big(\mathbf{E}(F^4_{i_1,i_2}) - 3\mathbf{E}(F^2_{i_1,i_2})^2\big)^{1/6} n^{2/3}(\log m)^{1/3}(\log nm)^{2/3}.$$

where $C > 0$ is an absolute constant and $d_{\mathrm{Kol}}(U, V)$ stands for the Kolmogorov distance between the distribution of the random variables U, V (see [20, Appendix C]). An illustration of these findings — inspired by the statistical theory developed in [13] — is presented in Section 4.2.

1.2 Notation

For $m \in \mathbb{N}$, we write $[m]$ to indicate the sets of integers $\{1, \ldots, m\}$. For $k \in [m]$, $p \in [k]$, $\{a_1, \ldots, a_p\} \subseteq [m]$ and $\{b_1, \ldots, b_q\} \subseteq [m] \setminus \{a_1, \ldots, a_p\}$, we define the sets

$$\mathfrak{L}_k := \{L \subseteq 2^{[m]} : |L| = k\},$$
$$\mathfrak{L}_k^{a_1\ldots a_p\,\hat{b}_1\ldots\hat{b}_q} := \{L \subseteq 2^{[m]} : |L| = k, \{a_1, \ldots, a_p\} \subseteq L \text{ and } \{b_1, \ldots, b_q\} \subseteq 2^{[m]} \setminus L\}.$$

Note that $|\mathfrak{L}_k| = \binom{m}{k}$ and $|\mathfrak{L}_k^{a_1\ldots a_p\,\hat{b}_1\ldots\hat{b}_q}| = \binom{m-p-q}{k-p-q}$.

For $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$, let $z_k \in \{z_1, \ldots, z_d\}$ be the k-th ordered statistics of its components, i.e.

$$\min_{i\in[d]} z_i = z_{(1)} \le \ldots \le z_{(k)} \le \ldots z_{(d)} = \max_{i\in[d]} z_i.$$

If $z = (z_{i_1,i_2})_{(i_1,i_2)\in[n]\times[m]} \in \mathbb{R}^{n\times m}$, we write its i_1-th row as $z_{i_1,\cdot} = (z_{i_1,1},\ldots,z_{i_1,m}) \in \mathbb{R}^m$. In particular $z_{i_1,(k)}$ indicates the k-th ordered statistics of the vector $z_{i_1,\cdot}$ and $z_{\cdot,(k)}$ stands for the vector $(z_{1,(k)},\ldots,z_{n,(k)}) \in \mathbb{R}^n$. Throughout the paper, we will refer to the quantity

$$\min\max z := \min_{i_1\in[n]} z_{i_1,(m)} = \min_{i_1\in[n]} \max_{i_2\in[m]} z_{i_1,i_2}$$

as the *min-max statistic* of the matrix z. We will always work on a fixed probability space $(\Omega,\mathcal{F},\mathbf{P})$ and write $\mathbf{E}$ for the expectation with respect to $\mathbf{P}$.

2 Comparison of min-max statistics for two Gaussian random matrices

2.1 Main estimates

The forthcoming statement is one of the main contributions of the present work, containing as special cases several results evoked in the Introduction; in particular, the inequalities (1)–(2) correspond to the case $n = k = 1$, $m \ge 1$ of our result; Theorem 1.4 in [9] corresponds to the case $n, m \ge 1$ and $k = 1$; Theorem 1.3 in [11] corresponds to the choice $n = 1$, $m \ge 1$ and $k \le m-1$ — see the subsequent discussion.

Theorem 1. *Let $X = (X_{i_1,i_2})_{(i_1,i_2)\in[n]\times[m]}$ and $Y = (Y_{i_1,i_2})_{(i_1,i_2)\in[n]\times[m]}$ be two Gaussian random matrices with $\mathbf{E}(X_{i_1,i_2}) = \mathbf{E}(Y_{i_1,i_2})$ for every $(i_1,i_2) \in [n]\times[m]$. Define $\gamma^X_{i_1,i_2;j_1,j_2} := \mathbf{E}((X_{i_1,i_2} - X_{j_1,j_2})^2)$, $\gamma^Y_{i_1,i_2;j_1,j_2} := \mathbf{E}((Y_{i_1,i_2} - Y_{j_1,j_2})^2)$ and let*

$$\gamma := \max_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}} \left|\gamma^X_{i_1,i_2;j_1,j_2} - \gamma^Y_{i_1,i_2;j_1,j_2}\right|.$$

Then, for all $k \in [m]$,

$$\begin{aligned}\Big|\mathbf{E}\Big(\min_{i_1\in[n]} \sum_{h=m-k+1}^{m} X_{i_1,(h)}\Big) - \mathbf{E}\Big(\min_{i_1\in[n]} \sum_{h=m-k+1}^{m} Y_{i_1,(h)}\Big)\Big| \\ \le \sqrt{\gamma k}\cdot\Big[\sqrt{k\Big(1-\frac{1}{n}\Big)\Big(2-\frac{1}{m}\Big)\log n} + \sqrt{\Big(1-\frac{k}{m}\Big)\log\binom{m}{k}}\Big].\end{aligned} \tag{3}$$

Moreover, if, for every $(i_2,j_2) \in [n]\times[m]$,

$$\begin{cases}\gamma^X_{i_1,i_2;j_1,j_2} \le \gamma^Y_{i_1,i_2;j_1,j_2} & \textit{if } i_1 = j_1\\ \gamma^X_{i_1,i_2;j_1,j_2} \ge \gamma^Y_{i_1,i_2;j_1,j_2} & \textit{if } i_1 \ne j_1,\end{cases}$$

then

$$\mathbf{E}\Big(\min_{i_1\in[n]}\sum_{h=m-k+1}^{m} X_{i_1,(h)}\Big) \le \mathbf{E}\Big(\min_{i_1\in[n]}\sum_{h=m-k+1}^{m} Y_{i_1,(h)}\Big).$$

Remark. In order to substantiate the claims preceding the statement of Theorem 1, we put forward the following two special cases: (i) when $n = 1$, then X and Y are m-dimensional Gaussian vectors, and the quantities inside the expectations on the left-hand side of (3) are the sums of the order statistics of orders $m-k+1$ up to m of X and Y (in particular, when $k = 1$ we recover the maxima); (ii) when $k = 1$ and no restrictions are put on n, m then the random variables on the left-hand side of (3) are the min-max statistics of X and Y.

Remark. There is no conceptual obstacle in extending Theorem 1 to the more general case of a d-dimensional Gaussian tensor $(X_{i_1,\dots,i_d})_{(i_1,\dots,i_d)\in[n_1]\times\cdots\times[n_d]}$ and the investigation of the quantity (say d is even without loss of generality and $k_i \le n_i$ for all $i \le d$)

$$\sum_{h_1=1}^{k_1}\sum_{h_2=n_2-k_2+1}^{n_2}\cdots\sum_{h_{d-1}=1}^{k_{d-1}}\sum_{h_d=n_d-k_d+1}^{n_d}(\cdots(X_{\cdot,\dots,\cdot,(h_d)})_{\cdot,\dots,\cdot,(h_{d-1})})\cdots)_{(h_1)},$$

but we decided not to perform it explicitly, in order to keep the length of the paper within bounds.

One remarkable consequence of Theorem 1 is that it yields comparison criteria for the expected values of order statistics associated with Gaussian random vectors.

Corollary 1. *Let $W = (W_i)_{i\in[d]}$ and $Z = (Z_i)_{i\in[d]}$ be two Gaussian random vectors with $\mathbf{E}(W_i) = \mathbf{E}(Z_i)$ for every $i \in [d]$, and let $\gamma = \max_{(i,j)\in[d]^2}|\mathbf{E}\big((W_i - W_j)^2\big) - \mathbf{E}\big((Z_i - Z_j)^2\big)|$. Then, for any $h \in [d]$,*

$$\big|\mathbf{E}(W_{(h)}) - \mathbf{E}(Z_{(h)})\big| \le \sqrt{\gamma}\Big(\sqrt{2\log\binom{d}{h}} + \sqrt{\log h}\Big) \le \sqrt{\gamma}\Big(\sqrt{2h(1+\log(d/h))} + \sqrt{\log h}\Big).$$

Proof. The key idea is that $W_{(h)}$ (respectively, $Z_{(h)}$) is the min-max statistic (see Section 1.2) of a matrix X with $\binom{d}{h}$ rows, where each row corresponds to a distinct subset of W (respectively, Z) with cardinality h (the order of the elements of the subset within a single row is immaterial). To see this, observe that the rows of the matrix X described above are such that: (i) there exists at least one row containing $W_{(h)}$ as a maximal element, and (ii) every other row contains one element that is larger or equal to $W_{(h)}$. Using now Theorem 1 with $n = \binom{d}{h}$, $m = h$ and $k = 1$ yields the first bound. The second bound follows easily from the first noting that $\binom{d}{h} \le (\frac{ed}{h})^h$. $\square$

Remark. An alternate class of local comparison theorems for (vectors of) order statistics of Gaussian matrices can be found in [7] – see the discussion following Theorem 2 below for further details.

Remark. We now show that, when m is fixed, the bound of Theorem 1 is sharp in the order of n and k. First, let $Y \equiv 0$ and X be a matrix with m columns which are the copy of a same n-dimensional standard Gaussian vector X'. Then $\gamma = 2$ for every n and

$$\sum_{h=m-k+1}^{m} X_{i_1,(h)} = kX'_{i_1}.$$

It is known from extreme value theory that, as n diverges, the expectation of $\min_{i_1} X'_{i_1}$ is of order $\sqrt{\log n}$ (up to constants). Then, the expectation of $\sum_{h=m-k+1}^{m} X_{i_1,(h)}$ is of order $k\sqrt{\log n}$, matching the order of the bound (3) in this specific case.

Analogously, when n is fixed, then the bound is sharp in the order of m and k in the regime where $k, m-k \ll m$. To see this, let $Y \equiv 0$ and X be a matrix with n rows which are the copy of a same m-dimensional (transposed) Gaussian vector X'. This time suppose without loss of generality that m is a multiple of k, $m = \tilde{m}k$, $\tilde{m} \in \mathbb{N}$ and that X' is the collection of k copies of the same standard $\tilde{m}$-dimensional Gaussian vector $\tilde{X}$. Note that in this case

$$\sum_{h=m-k+1}^{m} X_{i_1,(h)} = k \max_{j\in[\tilde{m}]} \tilde{X}_j$$

for all $i_1 \in [n]$. However, when $\tilde{m}$ diverges, the expected value of $\max_{j\in[\tilde{m}]} \tilde{X}_j$ is of order $\sqrt{\log \tilde{m}} = \sqrt{\log(m/k)}$, and, accordingly, the expectation of $k \max_{j\in[\tilde{m}]} \tilde{X}_j$ is of order

$$k\sqrt{\log(m/k)} = \sqrt{k}\sqrt{\log(m/k)^k} \approx \sqrt{k}\sqrt{\log\binom{m}{k}},$$

where the last approximation holds in the aforementioned regime of k with respect to m. Again, we recover asymptotically the bound (3).

The next section contains six technical results that are pivotal in the proof of Theorem 1.

2.2 Six ancillary lemmas

Lemma 1. *For $\beta, \delta > 0$ and $k \in [m]$, define the function $f_k^{\beta,\delta} \colon \mathbb{R}^{n\times m} \to \mathbb{R}$ by*

$$f_k^{\beta,\delta}(x) := -\frac{1}{\beta\delta} \log \sum_{\ell_1=1}^{n} \Big(\sum_{L\in\mathfrak{L}_k} \exp\Big(\beta \sum_{\ell_2\in L} x_{\ell_1,\ell_2}\Big)\Big)^{-\delta}.$$

For every $x \in \mathbb{R}^{n\times m}$, one has that

$$f_k^{\beta,\delta}(x) - \frac{1}{\beta}\log\binom{m}{k} \le \Big(\sum_{h=m-k+1}^{m} x_{\cdot,(h)}\Big)_{(1)} \le f_k^{\beta,\delta}(x) + \frac{1}{\beta\delta}\log n.$$

Proof. Let $d \in \mathbb{N}$ and $z \in \mathbb{R}^d$. When $\beta > 0$ the following inequality holds

$$\frac{1}{d}\sum_{i=1}^{d} e^{\beta z_i} \le e^{\beta z_{(d)}} \le \sum_{i=1}^{d} e^{\beta z_i},$$

in particular

$$\frac{1}{\beta}\log\sum_{i=1}^{d} e^{\beta z_i} - \frac{\log d}{\beta} \le z_{(d)} \le \frac{1}{\beta}\log\sum_{i=1}^{d} e^{\beta z_i}. \tag{4}$$

Similarly for the minimum instead, it holds that, for every $\beta' > 0$,

$$-\frac{1}{\beta'}\log\sum_{i=1}^{d} e^{-\beta' z_i} \le z_{(1)} \le -\frac{1}{\beta'}\log\sum_{i=1}^{d} e^{-\beta' z_i} + \frac{\log d}{\beta'}. \tag{5}$$

For each $\ell_1 \in \{1, \dots, n\}$, consider the vector $z = \left(\sum_{\ell_2\in L} x_{\ell_1,\ell_2}\right)_{L\in\mathfrak{L}_k} \in \mathbb{R}^{\binom{m}{k}}$ and apply (4) to it. Notice that $z_{(\binom{m}{k})} = \sum_{h=m-k+1}^{m} x_{\ell_1,(h)}$. We get

$$\frac{1}{\beta}\log\sum_{L\in\mathfrak{L}_k}\exp\Big(\beta\sum_{\ell_2\in L} x_{\ell_1,\ell_2}\Big) - \frac{1}{\beta}\log\binom{m}{k} \le \sum_{h=m-k+1}^{m} x_{\ell_1,(h)} \le \frac{1}{\beta}\log\sum_{L\in\mathfrak{L}_k}\exp\Big(\beta\sum_{\ell_2\in L} x_{\ell_1,\ell_2}\Big).$$

Now we want to isolate the minimum of the vector $\Big(\sum_{h=m-k+1}^{m} x_{\cdot,(h)}\Big) \in \mathbb{R}^n$, and we use (5) to do so. Notice that, for $\beta' = \beta\delta$, we get

$$-\frac{1}{\beta'}\log\sum_{i=1}^{n}\exp\Big(-\beta'\cdot\frac{1}{\beta}\log\sum_{L\in\mathfrak{L}_k}\exp\Big(\beta\sum_{\ell_2\in L} x_{\ell_1,\ell_2}\Big)\Big) = f_k^{\beta,\delta}(x)$$

and that

$$\begin{aligned}-\frac{1}{\beta'}\log\sum_{i=1}^{n}\exp\Big(-\beta'\cdot\Big(\frac{1}{\beta}\log\sum_{L\in\mathfrak{L}_k}\exp\Big(\beta\sum_{\ell_2\in L} x_{\ell_1,\ell_2}\Big) - \frac{1}{\beta}\log\binom{m}{k}\Big)\Big)\\ = -\frac{1}{\beta'}\log\Big[\binom{m}{k}^{\delta}\sum_{i=1}^{n}\exp\Big(-\beta'\cdot\Big(\frac{1}{\beta}\log\sum_{L\in\mathfrak{L}_k}\exp\Big(\beta\sum_{\ell_2\in L} x_{\ell_1,\ell_2}\Big)\Big)\Big)\Big] = f_k^{\beta,\delta}(x) - \frac{1}{\beta}\log\binom{m}{k},\end{aligned}$$

which concludes the proof by monotonicity. □

Remark. In [15], the authors have studied the distribution of order statistics using an approximating function that can be seen as a particular case of $f_k^{\beta,\delta}$, as used in the proof of Corollary 1. In this sense, Lemma 1 can be seen as a generalization of [16, Lemma 4(i)].

For $h \in \{0, \dots, m\}$, let $A \subseteq [m]$ with $|A| = h$. If $h > 0$ we write $A = \{a_1, \dots, a_h\}$ with $a_1 < \dots < a_h$. It is convenient at this point to define the functions, for any $i_1 \in [n]$, $p_{i_1}^A, q_{i_1} \colon \mathbb{R}^{n\times m} \to [0, +\infty)$ as

$$p_{i_1}^A(x) = p_{i_1}^{a_1,\dots,a_h}(x) := \frac{\sum_{L\in\mathfrak{L}_k^{a_1\cdots a_h}}\exp\big(\beta\sum_{\ell\in L} x_{i_1,\ell}\big)}{\sum_{L\in\mathfrak{L}_k}\exp\big(\beta\sum_{\ell\in L} x_{i_1,\ell}\big)},$$

$$q_{i_1}(x) \coloneqq \sum_{\ell_1=1}^{n} \left(\frac{\sum_{L\in\mathfrak{L}_k} \exp\big(\beta \sum_{\ell_2\in L} x_{i_1,\ell_2}\big)}{\sum_{L\in\mathfrak{L}_k} \exp\big(\beta \sum_{\ell_2\in L} x_{\ell_1,\ell_2}\big)} \right)^{\delta}.$$

where we use the notational conventions $p_{i_1}^{\emptyset} = 1$ and $\sum_{L\in\emptyset} = 0$.

Lemma 2. *Let $b_1 < \ldots < b_h$ and $B = \{a_{b_1}, \ldots, a_{b_h}\} \subseteq A$. It holds*

$$\sum_{a_{b_1},\ldots,a_{b_h}=1}^{m} p_{i_1}^{A} = k^h p_{i_1}^{A\setminus B},$$
$$\sum_{i_1=1}^{n} \frac{1}{q_{i_1}} = 1.$$

In particular, for $A = B$ we infer that

$$\sum_{a_1,\ldots,a_h=1}^{m} p_{i_1}^{a_1,\ldots,a_h} = k^h.$$

Since $p_{i_1}^{a_1,a_1} = p_{i_1}^{a_1}$, this implies also that

$$\sum_{\substack{a_2=1\\ a_2\neq a_1}}^{m} p_{i_1}^{a_1,a_2} = (k-1) p_{i_1}^{a_1}.$$

Proof. We only prove the first equation for $h = 1$, the general case is then similarly proved by iteration. For any $x \in \mathbb{R}^{n\times m}$, one has that

$$\begin{aligned}
\sum_{a=1}^{m} \sum_{L\in\mathfrak{L}_k^a} \exp\Big(\beta \sum_{\ell\in L} x_{i_1,\ell}\Big) &= \sum_{a=1}^{m} \sum_{L\in\mathfrak{L}_k} \mathbf{1}_{\{a\in L\}} \exp\Big(\beta \sum_{\ell\in L} x_{i_1,\ell}\Big)\\
&= \sum_{L\in\mathfrak{L}_k} \exp\Big(\beta \sum_{\ell\in L} x_{i_1,\ell}\Big) \sum_{a=1}^{m} \mathbf{1}_{\{a\in L\}}\\
&= k \sum_{L\in\mathfrak{L}_k} \exp\Big(\beta \sum_{\ell\in L} x_{i_1,\ell}\Big),
\end{aligned}$$

that gives the first claim. The second identity of the statement is a straightforward consequence of the definition of q_{i_1}. □

Lemma 3. *For all $(i_1, i_2) \in [n] \times [m]$,*

$$\sum_{(j_1,j_2)\in[n]\times[m]} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2} \partial x_{j_1,j_2}} = 0.$$

Proof. A direct computation shows that

$$\frac{\partial f_k^{\beta,\delta}}{\partial x_{i_1,i_2}} = \frac{p_{i_1}^{i_2}}{q_{i_1}}$$

and that

$$\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}} = \beta\Big[\delta\frac{p_{i_1}^{i_2}p_{j_1}^{j_2}}{q_{i_1}q_{j_1}} + \frac{\mathbf{1}_{\{i_1=j_1\}}}{q_{i_1}}\big(-(1+\delta)p_{i_1}^{i_2}p_{j_1}^{j_2} + p_{i_1}^{i_2,j_2}\big)\Big]. \tag{6}$$

We will evaluate the contribution of the three summands of (6) separately and show that they cancel out to 0. Notice that β/q_{i_1} is a common multiplicative factor to all three so we can ignore it. We use both properties stated in Lemma 2 to deduce that

$$\sum_{(j_1,j_2)\in[n]\times[m]} \delta\,\frac{p_{i_1}^{i_2}p_{j_1}^{j_2}}{q_{j_1}} = \delta k p_{i_1}^{i_2},$$
$$\sum_{(j_1,j_2)\in[n]\times[m]} -(1+\delta)p_{i_1}^{i_2}p_{i_1}^{j_2} = -(1+\delta)kp_{i_1}^{i_2},$$
$$\sum_{(j_1,j_2)\in[n]\times[m]} p_{i_1}^{i_2,j_2} = kp_{i_1}^{i_2},$$

which concludes the proof. ∎

Lemma 4. *With the same notations and assumptions as above, one has that*

$$\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}\cdot(\sigma^Y_{i_1,i_2;j_1,j_2} - \sigma^X_{i_1,i_2;j_1,j_2})$$
$$= -\frac{1}{2}\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]\\(i_1,i_2)\neq(j_1,j_2)}} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}\cdot(\gamma^Y_{i_1,i_2;j_1,j_2} - \gamma^X_{i_1,i_2;j_1,j_2}).$$

Proof. First, note that $\gamma^X_{i_1,i_2;j_1,j_2} = \sigma^X_{i_1,i_2;i_1,i_2} - 2\sigma^X_{j_1,j_2;j_1,j_2} + \sigma^X_{i_1,i_2;i_1,i_2}$, analogously for Y. We can split the first sum of the statement as

$$\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}} \cdot (\sigma^Y_{i_1,i_2;j_1,j_2} - \sigma^X_{i_1,i_2;j_1,j_2})$$

$$= \frac{1}{2} \sum_{(i_1,i_2)\in[n]\times[m]} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{i_1,i_2}} \cdot (\sigma^Y_{i_1,i_2;i_1,i_2} - \sigma^X_{i_1,i_2;i_1,i_2})$$
$$- \frac{1}{2} \sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]\\(i_1,i_2)\neq(j_1,j_2)}} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}} \cdot (-2\sigma^Y_{i_1,i_2;j_1,j_2} + 2\sigma^X_{i_1,i_2;j_1,j_2})$$
$$+ \frac{1}{2} \sum_{(j_1,j_2)\in[n]\times[m]} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{j_1,j_2}\partial x_{j_1,j_2}} \cdot (\sigma^Y_{j_1,j_2;j_1,j_2} - \sigma^X_{j_1,j_2;j_1,j_2}).$$

By Lemma 3 we know that for all $(a_1, a_2) \in [n] \times [m]$,

$$\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{a_1,a_2}\partial x_{a_1,a_2}} = - \sum_{\substack{(b_1,b_2)\in[n]\times[m]\\(b_1,b_2)\neq(a_1,a_2)}} \frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{a_1,a_2}\partial x_{b_1,b_2}},$$

which allows us to conclude. □

Lemma 5. *For all $x \in \mathbb{R}^{n\times m}$,*

$$\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(x) \begin{cases} \geq 0 & \text{if } i_1 \neq j_1 \\ \leq 0 & \text{if } i_1 = j_1 \text{ and } i_2 \neq j_2. \end{cases}$$

Proof. Notice that we can rewrite (6) in the following way:

$$\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}} = \frac{\beta}{q_{i_1}} \Big[\delta p^{i_2}_{i_1} p^{j_2}_{j_1} \Big(\frac{1}{q_{j_1}} - \mathbf{1}_{\{i_1=j_1\}}\Big) + \mathbf{1}_{\{i_1=j_1\}} \big(p^{i_2,j_2}_{i_1} - p^{i_2}_{i_1} p^{j_2}_{j_1}\big)\Big]. \tag{7}$$

Obviously when $i_1 \neq j_1$ the expression is positive. When $i_1 = j_1$, notice that $1/q_{i_1} - 1 \leq 0$ since $q_{i_1} \geq 1$, being a sum of non-negative summands and one of those (corresponding to $\ell = i_1$) is exactly equal to 1 . It thus remains to prove only that $p^{i_2,j_2}_{i_1} - p^{i_2}_{i_1} p^{j_2}_{i_1} \leq 0$ whenever $i_2 \neq j_2$. Writing $y_\ell := e^{\beta x_{i_1,\ell}} > 0$ and multiplying both sides by $\sum_{L\in\mathfrak{L}_k} \exp\big(\beta \sum_{\ell\in L} x_{i_1,\ell}\big)$, this is equivalent to

$$\sum_{L\in\mathfrak{L}_k} \prod_{\ell\in L} y_\ell \sum_{L'\in\mathfrak{L}^{ij}_k} \prod_{\ell'\in L'} y_{\ell'} \leq \sum_{L\in\mathfrak{L}^{i}_k} \prod_{\ell\in L} y_\ell \sum_{L'\in\mathfrak{L}^{j}_k} \prod_{\ell'\in L'} y_{\ell'},$$

where we also renamed i_2 and j_2 as i and j, respectively, for simplicity of notation. Since we can decompose the first double sum as

$$\sum_{L\in\mathfrak{L}_k}\sum_{L'\in\mathfrak{L}_k^{ij}} = \sum_{L\in\mathfrak{L}_k^{ij}}\sum_{L'\in\mathfrak{L}_k^{ij}} + \sum_{L\in\mathfrak{L}_k^{i\hat{j}}}\sum_{L'\in\mathfrak{L}_k^{ij}} + \sum_{L\in\mathfrak{L}_k^{\hat{i}j}}\sum_{L'\in\mathfrak{L}_k^{ij}} + \sum_{L\in\mathfrak{L}_k^{\hat{i}\hat{j}}}\sum_{L'\in\mathfrak{L}_k^{ij}},$$

and the second as

$$\sum_{L\in\mathfrak{L}_k^{i}}\sum_{L'\in\mathfrak{L}_k^{j}} = \sum_{L\in\mathfrak{L}_k^{ij}}\sum_{L'\in\mathfrak{L}_k^{ij}} + \sum_{L\in\mathfrak{L}_k^{i\hat{j}}}\sum_{L'\in\mathfrak{L}_k^{ij}} + \sum_{L\in\mathfrak{L}_k^{ij}}\sum_{L'\in\mathfrak{L}_k^{\hat{i}j}} + \sum_{L\in\mathfrak{L}_k^{i\hat{j}}}\sum_{L'\in\mathfrak{L}_k^{\hat{i}j}},$$

it appears that the only comparison that needs to be checked is

$$\sum_{L\in\mathfrak{L}_k^{\hat{i}\hat{j}}}\sum_{L'\in\mathfrak{L}_k^{ij}}\prod_{\ell\in L} y_\ell \prod_{\ell'\in L'} y_{\ell'} \leq \sum_{L\in\mathfrak{L}_k^{i\hat{j}}}\sum_{L'\in\mathfrak{L}_k^{\hat{i}j}}\prod_{\ell\in L} y_\ell \prod_{\ell'\in L'} y_{\ell'}.$$

By simplifying a factor $y_i y_j$ on both sides, this is inequality is equivalent to

$$\sum_{L\in\mathfrak{L}_k}\sum_{L'\in\mathfrak{L}_{k-2}}\prod_{\ell\in L} y_\ell \prod_{\ell'\in L'} y_{\ell'} \leq \sum_{L\in\mathfrak{L}_{k-1}}\sum_{L'\in\mathfrak{L}_{k-1}}\prod_{\ell\in L} y_\ell \prod_{\ell'\in L'} y_{\ell'},$$

with the caveat that here $\mathfrak{L}_k$ (respectively $\mathfrak{L}_{k-1}, \mathfrak{L}_{k-2}$) indicates the collection of subsets of cardinality k (respectively $k-1, k-2$) out of $m-2$ total indexes (and not m as previously). We will show that every double product of the LHS appears with larger multiplicity in the RHS. Let L be a multiset of $2k$ indices (i.e. a set with $2k$ elements that can be repeated, out of a set of $m-2$ total indexes). Then, the quantity $\prod_{\ell\in L} y_\ell$ appears in the LHS if and only if the two following conditions are met:

- no index ℓ appears in L more than twice,
- there are at most $k-2$ repeated indices.

Suppose that exactly $r \leq k-2$ indices are repeated in L. Those indexes have to appear both in each L and L' of the RHS. Then, in the LHS there are $\binom{2k-2r-2}{k-r-2}$ ways to rearrange the remaining indexes between L and L', while in the RHS there are $\binom{2k-2r-2}{k-r-1}$ ways. Notice that the latter quantity is a central binomial coefficient so it is necessarily larger than the former. □

Lemma 6. *For all $x \in \mathbb{R}^{n\times m}$, it holds that*

$$\sum_{\substack{(i_1,i_2)\in[n]\times[m] \\ (j_1,j_2)\in[n]\times[m] \\ (i_1,i_2)\neq(j_1,j_2)}} \left|\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}\right| \leq \beta\frac{k}{m}\big(m-k+a_n(2m-1)k\delta\big),$$

where $a_n \coloneqq 1-\frac{1}{n} \in [0,1)$.

Proof. Again, we evaluate the contribution of the summands of (6) separately. We make use extensively of Lemma 2. We start with the case $i_1 \neq j_1$:

$$\frac{1}{\beta}\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]\\i_1\neq j_1}}\Big|\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}\Big| = \sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]\\i_1\neq j_1}}\delta\,\frac{p_{i_1}^{i_2}p_{j_1}^{j_2}}{q_{i_1}q_{j_1}} = \delta k^2\Big(1-\sum_{i_1\in[n]}\frac{1}{q_{i_1}^2}\Big) \leq \delta k^2\Big(1-\frac{1}{n}\Big).$$

The last inequality is due to the fact that the terms $1/q_{i_1}$ are positive and sum to 1, in particular by the Cauchy-Schwarz inequality the sum of their squares is minimized when all of them are equal to $1/n$.

For the case $i_1 = j_1$, we make use of the expression (7) in which both summands are non-positive, in such a way that we can write

$$\begin{aligned}\frac{1}{\beta}\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]\\i_1=j_1\\i_2\neq j_2}}\Big|\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}\Big| &= \sum_{\substack{(i_1,i_2)\in[n]\times[m]\\j_2\in[m]\\i_2\neq j_2}}\frac{1}{q_{i_1}}\Big[\delta p_{i_1}^{i_2}p_{i_1}^{j_2}\Big(1-\frac{1}{q_{i_1}}\Big)+\big(p_{i_1}^{i_2}p_{i_1}^{j_2}-p_{i_1}^{i_2,j_2}\big)\Big]\\
&= \sum_{\substack{(i_1,i_2)\in[n]\times[m]\\j_2\in[m]\\i_2\neq j_2}}\Big(\frac{1+\delta}{q_{i_1}}-\frac{\delta}{q_{i_1}^2}\Big)p_{i_1}^{i_2}p_{i_1}^{j_2} - \sum_{\substack{(i_1,i_2)\in[n]\times[m]\\j_2\in[m]\\i_2\neq j_2}}\frac{p_{i_1}^{i_2,j_2}}{q_{i_1}}.\end{aligned}$$

The first sum can be estimated as follows:

$$\begin{aligned}\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\j_2\in[m]\\i_2\neq j_2}}\Big(\frac{1+\delta}{q_{i_1}}-\frac{\delta}{q_{i_1}^2}\Big)p_{i_1}^{i_2}p_{i_1}^{j_2} &= \sum_{(i_1,i_2)\in[n]\times[m]}\Big(\frac{1+\delta}{q_{i_1}}-\frac{\delta}{q_{i_1}^2}\Big)p_{i_1}^{i_2}\sum_{\substack{j_2\in[m]\\i_2\neq j_2}}p_{i_1}^{j_2}\\
&= \sum_{(i_1,i_2)\in[n]\times[m]}\Big(\frac{1+\delta}{q_{i_1}}-\frac{\delta}{q_{i_1}^2}\Big)p_{i_1}^{i_2}(k-p_{i_1}^{i_2})\\
&= \sum_{i_1\in[n]}\Big(\frac{1+\delta}{q_{i_1}}-\frac{\delta}{q_{i_1}^2}\Big)\Big(k^2-\sum_{i_2\in[m]}(p_{i_1}^{i_2})^2\Big)\\
&\leq k^2\Big(1-\frac{1}{m}\Big)\Big(1+\delta\Big(1-\frac{1}{n}\Big)\Big),\end{aligned}$$

where we used the fact that the numbers $\{p_{i_1}^{i_2}\}_{i_2\in[m]}$ sum to k, so that the minimum of $\sum_{i_2\in[m]}(p_{i_1}^{i_2})^2$ is k^2/m, again by the Cauchy-Schwarz inequality. Concerning the last summand, we obtain

$$\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\j_2\in[m]\\i_2\neq j_2}}\frac{p_{i_1}^{i_2,j_2}}{q_{i_1}} = \sum_{(i_1,i_2)\in[n]\times[m]}\frac{1}{q_{i_1}}\sum_{\substack{j_2\in[m]\\i_2\neq j_2}}p_{i_1}^{i_2,j_2} = \sum_{(i_1,i_2)\in[n]\times[m]}(k-1)\frac{p_{i_1}^{i_2}}{q_{i_1}} = k(k-1).$$

Note that the sums over the indexes $i_1 \neq j_1$ (respectively, $i_2 \neq j_2$) make sense only when $n > 1$ (respectively, $m > 1$) but those expressions are 0 anyway, if that is not the case. Merging the three contributions, one deduces the desired conclusion. $\square$

We are now ready to prove Theorem 1.

2.3 Proof of Theorem 1

Once the analytical lemmas presented in the previous section are established, the proof follows from a classical interpolation technique — already exploited e.g. in [3, 18] or [20, Chapter 6]. Without loss of generality, we can assume that X and Y are independent. Let $\mu_{i_1,i_2} = \mathbf{E}(X_{i_1,i_2}) = \mathbf{E}(X_{i_1,i_2})$. For $t \in [0,1]$ we consider the random interpolation matrix $Z_t \in \mathbb{R}^{n\times m}$ whose entries are given by

$$(Z_t)_{i_1,i_2} := \sqrt{1-t}(X_{i_1,i_2} - \mu_{i_1,i_2}) + \sqrt{t}(Y_{i_1,i_2} - \mu_{i_1,i_2}) + \mu_{i_1,i_2}.$$

Note that $Z_0 = X$, $Z_1 = Y$ and $\mathbf{E}((Z_t)_{i_1,i_2}) = \mu_{i_1,i_2}$ for all $t \in [0,1]$ and all $i_1, i_2 \in [n] \times [m]$. We also define $\psi(t) := \mathbf{E}(f_k^{\beta,\delta}(Z_t))$, in such a way that ψ is differentiable with derivative

$$\psi'(t) = \frac{1}{2} \sum_{(j_1,j_2)\in[n]\times[m]} \mathbf{E}\Big(\frac{\partial f_k^{\beta,\delta}}{\partial x_{j_i,j_2}}(Z_t)\Big(\frac{Y_{j_1,j_2} - \mu_{j_1,j_2}}{\sqrt{t}} - \frac{X_{j_1,j_2} - \mu_{j_1,j_2}}{\sqrt{1-t}}\Big)\Big).$$

Moreover, integration by parts yields

$$\mathbf{E}\Big(\frac{\partial f_k^{\beta,\delta}}{\partial x_{j_i,j_2}}(Z_t)(Y_{j_1,j_2} - \mu_{j_1,j_2})\Big) = \sum_{(i_1,i_2)\in[n]\times[m]} \sqrt{t}\,\mathbf{E}\Big(\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(Z_t)\Big)\sigma^Y_{i_1,i_2;j_1,j_2}$$

and

$$\mathbf{E}\Big(\frac{\partial f_k^{\beta,\delta}}{\partial x_{j_i,j_2}}(Z_t)(X_{j_1,j_2} - \mu_{j_1,j_2})\Big) = \sum_{(i_1,i_2)\in[n]\times[m]} \sqrt{1-t}\,\mathbf{E}\Big(\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(Z_t)\Big)\sigma^X_{i_1,i_2;j_1,j_2}.$$

Plugging both previous identities into the initial one, we obtain

$$\psi'(t) = \frac{1}{2} \sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}} \mathbf{E}\Big(\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(Z_t)\Big)(\sigma^Y_{i_1,i_2;j_1,j_2} - \sigma^X_{i_1,i_2;j_1,j_2}).$$

Note that, by construction,

$$\mathbf{E}\big((f_k^{\beta,\delta}(Y)\big) - \mathbf{E}\big(f_k^{\beta,\delta}(X)\big) = \psi(1) - \psi(0) = \int_0^1 \psi'(t)\,\mathrm{d}t.$$

Using Lemma 4 in combination with Lemma 5 shows that under the conditions on the signs of $\gamma^X_{i_1,i_2;j_1,j_2} - \gamma^Y_{i_1,i_2;j_1,j_2}$ as per assumption, $\psi' \geq 0$, hence $\mathbf{E}\big(f_k^{\beta,\delta}(X)\big) \leq \mathbf{E}\big(f_k^{\beta,\delta}(Y)\big)$, from which the second claim of the theorem follows by letting $\beta \to \infty$, thanks to Lemma 1. Again, Lemma 1 combined this time with Lemma 6 and

$$\left|\mathbf{E}(f_k^{\beta,\delta}(X)) - \mathbf{E}(f_k^{\beta,\delta}(Y))\right| \le \sup_{t\in[0,1]} |\psi'(t)| \le \frac{\beta k}{4m}(m-k+a_n(2m-1)k\delta)\gamma,$$

shows that

$$\begin{aligned}\Big|\mathbf{E}\Big(\min_{i_1\in[n]}\sum_{h=m-k+1}^{m} X_{i_1,(h)}\Big)-\mathbf{E}\Big(\min_{i_1\in[n]}\sum_{h=m-k+1}^{m} Y_{i_1,(h)}\Big)\Big| \\ \le \frac{\beta k}{4m}\big(m-k+a_n(2m-1)k\delta\big)\gamma + \frac{1}{\beta\delta}\log n + \frac{1}{\beta}\log\binom{m}{k},\end{aligned}$$

that for $k \in [m-1]$ is minimized by

$$\beta = 2\sqrt{\frac{m\log\binom{m}{k}}{(m-k)k\gamma}} \qquad \text{and} \qquad \delta = \sqrt{\frac{(m-k)\log n}{ka_n(2m-1)\log\binom{m}{k}}},$$

yielding bound (3). For $k = m$ the bound is trivially true since the LHS is 0. □

3 Comparison of min-max statistics of two random matrices, one of which is Gaussian

3.1 The language of Malliavin calculus

The reader is referred e.g. to the monograph [20] for a detailed discussion of the concepts presented in this subsection.

Let $\mathfrak{H}$ be a real separable Hilbert space, and write $\langle\cdot,\cdot\rangle_{\mathfrak{H}}$ to indicate the corresponding inner product. In what follows, we will write $G = \{G(h) : h \in \mathfrak{H}\}$ to denote an *isonormal Gaussian process* over $\mathfrak{H}$, that is: G is a (real) centered Gaussian family indexed by $\mathfrak{H}$ and such that $\mathbf{E}\big(G(h)G(g)\big) = \langle h, g\rangle_{\mathfrak{H}}$, for all $h, g \in \mathfrak{H}$. Denoting by $\sigma(G)$ the sigma-field generated by G, every $F \in L^2(\sigma(G))$ admits a *Wiener-Itô chaos expansion* of the form

$$F = \mathbf{E}(F) + \sum_{q=1}^{\infty} I_q(f_q), \tag{8}$$

where f_q is an element of the symmetric qth tensor product $\mathfrak{H}^{\odot q}$ (which is uniquely determined by F), and $I_q(f_q)$ is the q-th *multiple Wiener-Itô integral* of f_q with respect to G. One writes $F \in \mathbb{D}^{1,2}$ if

$$\sum_{q\ge 1} q q! \|f_q\|^2_{\mathfrak{H}^{\otimes q}} < \infty.$$

For $F \in \mathbb{D}^{1,2}$, we denote by DF the *Malliavin derivative* of F. Recall that DF is by definition a random element with values in $\mathfrak{H}$. The operator D satisfies a crucial *chain rule*: if φ is a mapping on $\mathbb{R}^m$ of class C^1 with bounded derivatives and if $F_1, \ldots, F_m \in \mathbb{D}^{1,2}$, then $\varphi(F_1, \ldots, F_m) \in \mathbb{D}^{1,2}$, and also

$$D\varphi(F_1, \ldots, F_m) = \sum_{i=1}^{m} \partial_i \varphi(F_1, \ldots, F_m) DF_i.$$

For general $p > 2$, we write $F \in \mathbb{D}^{1,p}$ if $F \in L^p(\sigma(G)) \cap \mathbb{D}^{1,2}$ and $\mathbf{E}\big(\|DF\|_{\mathfrak{H}}^p\big) < \infty$. The adjoint of D, customarily referred to as the *divergence operator* or the *Skorohod integral*, is denoted by δ and satisfies the duality formula,

$$\mathbf{E}\big(\delta(u)F\big) = \mathbf{E}\big(\langle u, DF\rangle_{\mathfrak{H}}\big)$$

for all $F \in \mathbb{D}^{1,2}$, whenever $u \colon \Omega \to \mathfrak{H}$ is contained in the domain $\mathrm{Dom}(\delta)$ of δ.

The *generator of the Ornstein-Uhlenbeck semigroup*, written L, is defined by the relation $LF = -\sum_{q\geq 1} qI_q(f_q)$ for every F as in (8) such that $\sum_{q\geq 1} q^2 q! \|f_q\|^2_{\mathfrak{H}^{\otimes q}} < \infty$. The *pseudo-inverse* of L, written L^{-1}, is the operator defined, as $L^{-1}F = -\sum_{q\geq 1} \frac{1}{q} I_q(f_q)$, for all $F \in L^2(\sigma(G))$ as in (8). The fundamental relation linking the objects introduced above is the identity

$$F = \mathbf{E}(F) - \delta(DL^{-1}F),$$

which is valid for any $F \in L^2(\sigma(G))$ (this relation implies in particular that, for every $F \in L^2(\sigma(G))$, $DL^{-1}F \in \mathrm{Dom}(\delta)$).

The notation and setting introduced above will prevail for the rest of the section; also, we will systematically assume that the underlying Hilbert space $\mathfrak{H}$ has infinite dimension.

3.2 Main estimates

We now fix the following objects: for $n \geq 1$ and $m \geq 2$, $X = (X_{i_1,i_2})_{(i_1,i_2)\in[n]\times[m]}$ is a centered Gaussian random matrix with covariance matrix $(\sigma_{i_1,i_2;j_1,j_2})_{(i_1,i_2),(j_1,j_2)\in[n]\times[m]}$ (without loss of generality, we can assume that X is extracted from the isonormal Gaussian process G); $F = (F_{i_1,i_2})_{(i_1,i_2)\in[n]\times[m]}$ is a centered random matrix with entries $F_{i_1,i_2} \in \mathbb{D}^{1,2}$. We also write σ_{i_1,i_2} as shorthand for $\sqrt{\mathbf{Var}(X_{i_1,i_2})}$, $\underline{\sigma}_{i_1} := \min_{i_2\in[m]} \sigma_{i_1,i_2}$, $\underline{\sigma} = \min_{(i_1,i_2)\in[n]\times[m]} \sigma_{i_1,i_2}$ and $\overline{\sigma} := \max_{(i_1,i_2)\in[n]\times[m]} \sigma_{i_1,i_2}$.

For simplicity, we will now work with statistics such as the ones appearing on the left-hand side of (3) only in the case $k = 1$. To this end, recall that we write $\min\max X$ to indicate $\min_{i_1\in[n]} \max_{i_2\in[m]} X_{i_1,i_2}$ and analogously for F. Our main findings are contained in the statement of the forthcoming Theorem 2, providing an upper bound on the Kolmogorov distance between the distributions of the min-max's of X and F.

In this section, we make the mild assumption that the covariance structure of the random matrix X is such that

$$\mathbf{P}\Big(\big|\{(i_1,i_2) \in [n]\times[m] : X_{i_1,i_2} = \min\max X\}\big| > 1\Big) = 0; \tag{A}$$

in other words: we require that, with probability one, there exists a unique pair (i_*, j_*) such that $X_{i_*,j_*} = \min\max X$. This is, for instance, the case when $\mathrm{corr}(X_{i_1,i_2}, X_{j_1,j_2}) < 1$ for all distinct pairs (i_1, i_2) and (j_1, j_2), or when X is the matrix associated with the order statistics of a vector W, whose components verify $\mathrm{corr}(W_i, W_j) < 1$ for all $i \neq j$, built as described in the proof of Corollary 1. Indeed, by construction, the argument of the min-max of such a matrix is always unique with probability 1.

The next statement is the main achievement of the present section. In the special case $n = 1$, it generalizes both [4, Theorem 2] and [13, Theorem 2.1]. Note that reference [4] only deals with the case in which both F and X are Gaussian.

Theorem 2. *Let the above assumptions prevail, suppose that $\underline{\sigma} > 0$ and let*

$$\Delta := \mathbf{E}\left(\max_{\substack{(i_1,i_2)\in[n]\times[m] \\ (j_1,j_2)\in[n]\times[m]}} \left| \langle DF_{i_1,i_2}, -DL^{-1}F_{j_1,j_2} \rangle - \sigma_{i_1,i_2;j_1,j_2} \right| \right). \tag{9}$$

(a) Let $a_{m,i} := \mathbf{E}\big(\max_{i_2\in[m]} \frac{X_{i,i_2}}{\sigma_{i,i_2}}\big)$, $\alpha_{nm} := \frac{1}{n}\sum_{i=1}^{n} a_{m,i}$ and $p_{nm} := n/\log nm$. Suppose that there exist constants $\zeta, \zeta' > 0$ such that $\zeta \leq \underline{\sigma} \leq \overline{\sigma} \leq \zeta'$. Then, there exists a constant $C > 0$, depending only on ζ, ζ', such that

$$\begin{aligned} &\sup_{x\in\mathbb{R}} \left| \mathbf{P}(\min\max F \leq x) - \mathbf{P}(\min\max X \leq x) \right| \\ &\qquad\qquad \leq Cn\sqrt{\log nm}(\log m)^{1/4} \max\big(1, \sqrt{\alpha_{nm}}, \log(1/\Delta)^{1/4}\big)\sqrt{\Delta}. \end{aligned}$$

(b) There exists a universal constant $\tilde{C} > 0$ such that

$$\sup_{x\in\mathbb{R}} \left| \mathbf{P}(\min\max F \leq x) - \mathbf{P}(\min\max X \leq x) \right| \leq \frac{\tilde{C}}{\underline{\sigma}} n\sqrt{\Delta \log nm \log m}. \tag{10}$$

Remark. If F is Gaussian, then the quantity Δ appearing in (9) is simply the maximal discrepancy – in absolute value – between the entries of the covariance matrices. In this special case, our bounds can be compared with [7, Theorem 2.1]. In particular, specializing such a result to maxima ($r = d$ in the notation of [7]) yields an estimate on the left-hand side of (10) where the mapping $(n, m) \mapsto n(\log m)^{1/2}(\log nm)^{1/2}$ is replaced by an application of the type $(n, m) \mapsto n^a m^b$, with $a, b > 1$, and $\Delta^{1/2}$ is replaced by a index of discrepancy between the two covariance matrices which is of the order Δ, for Δ converging to zero. We also observe that — reasoning as in [13, Corollary 2.1] — the bounds in the statement of Theorem 2 continue to hold if the matrices F and X are replaced by those with entries $|F_{i_1,i_2}|$ and $|X_{i_1,i_2}|$, respectively (up to a change in the exact value of the absolute constants $C, \tilde{C}$).

We proceed with the proof of Theorem 2. First, we need a bound on the second derivatives of the composition of a smooth function with the approximation function $f_k^{\beta,\delta}$ defined in Lemma 1 (note that, in the statement below, we consider a generic $k \in [m]$ despite only $k = 1$ being relevant for the present section).

Lemma 7. *Let $g\colon \mathbb{R}\to\mathbb{R}$ be twice continuously differentiable with bounded first and second derivatives. Then for all $(i_1,i_2),(j_1,j_2)\in[n]\times[m]$,*

$$\frac{\partial^2(g\circ f_k^{\beta,\delta})}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}} = g''(f_k^{\beta,\delta})\frac{p_{i_1}^{i_2}}{q_{i_1}}\frac{p_{j_1}^{j_2}}{q_{j_1}} + g'(f_k^{\beta,\delta})\beta\Big[\delta\frac{p_{i_1}^{i_2}p_{j_1}^{j_2}}{q_{i_1}q_{j_1}} + \frac{\mathbf{1}_{\{i_1=j_1\}}}{q_{i_1}}\big(-(1+\delta)p_{i_1}^{i_2}p_{j_1}^{j_2} + p_{i_1}^{i_2,j_2}\big)\Big].$$

In particular

$$\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}}\Big|\frac{\partial^2(g\circ f_k^{\beta,\delta})}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}\Big| \le k^2\|g''\|_\infty + 2\beta k(1+\delta k)\|g'\|_\infty.$$

Proof. By the chain rule we have

$$\frac{\partial^2(g\circ f_k^{\beta,\delta})}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}} = g''(f_k^{\beta,\delta})\frac{\partial f_k^{\beta,\delta}}{\partial x_{i_1,i_2}}\frac{\partial f_k^{\beta,\delta}}{\partial x_{i_1,i_2}} + g'(f_k^{\beta,\delta})\frac{\partial^2 f_k^{\beta,\delta}}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}},$$

which yields the statement because of the computations above. The last inequality follows from the computations of Lemma 6, with the difference that here also the diagonal indexes are taken into account. □

The next statement generalizes Lemma 5 in [4].

Lemma 8. *Let W be a Gaussian random matrix in $\mathbb{R}^{n\times m}$ with $\mathbf{Var}(W_{i_1,i_2})=1$ for all $(i_1,i_2)\in[n]\times[m]$ and (A) holds. Then the distribution of $\min\max W$ admits a density with respect to the Lebesgue measure given by*

$$g_{n,m}(z) = \varphi(z)\sum_{i=1}^n H_i(z)G_i(z),$$

where

$$H_i(z) := \mathbf{P}\Big(\max_{\ell\in[m]} W_{i,\ell} = \min\max W \;\Big|\; \max_{\ell\in[m]} W_{i,\ell} = z\Big)$$

and

$$G_i(z) := \sum_{\ell=1}^m e^{\mathbf{E}(W_{i,\ell})z - \mathbf{E}(W_{i,\ell})^2}\mathbf{P}\Big(W_{i,k} = \max_{\ell\in[m]} W_{i,\ell} \;\Big|\; W_{i,k}=z\Big).$$

Moreover $z\mapsto e^{\mathbf{E}(W_{i,\ell})z-\mathbf{E}(W_{i,\ell})^2}\mathbf{P}\big(W_{i,k} = \max_{\ell\in[m]}W_{i,\ell} \mid W_{i,k}=z\big)$ is non-decreasing as soon as $\mathbf{E}(W_{i,k})\ge 0$.

Proof. Exploiting assumption (A), one hast that, for every real t,

$$\mathbf{P}\big(\min\max W\le t\big) = \sum_{i=1}^n \mathbf{P}\Big(\max_{\ell\in[m]}W_{i,\ell} = \min\max W \cap \max_{\ell\in[m]}W_{i,\ell}\le t\Big).$$

Writing

$$\mathbf{P}\Big(\max_{\ell\in[m]} W_{i,\ell} = \min\max W \,\cap\, \max_{\ell\in[m]} W_{i,\ell} \le t\Big)$$
$$= \int_{-\infty}^{t} \mathbf{P}\Big(\max_{\ell\in[m]} W_{i,\ell} = \min\max W \;\Big|\; \max_{\ell\in[m]} W_{i,\ell} = z\Big)\mu_i(\mathrm{d}z),$$

where μ_i stands for the law of $\max_{\ell\in[m]} W_{i,\ell}$, we deduce the desired conclusion from Lemmas 5 and 6 in [4]. □

The following is a generalization of [4, Theorem 3]. For every $\varepsilon > 0$ the anti-concentration function of a r.v. Y is defined as

$$\mathcal{L}(Y,\varepsilon) \coloneqq \sup_{x\in\mathbb{R}} \mathbf{P}\big(|Y - x| \le \varepsilon\big).$$

If Y if absolutely continuous with essentially bounded density f then it follows from the definition that

$$\mathcal{L}(Y,\varepsilon) \le 2\varepsilon\, \|f\|_\infty,$$

where $\|\cdot\|_\infty$ is the essential supremum.

Lemma 9 (Anti-concentration inequality, first variant)**.** *There exists $C > 0$ that depends only on $\underline{\sigma}$ and $\overline{\sigma}$ such that for all $\varepsilon > 0$*

$$\mathcal{L}(\min\max X,\varepsilon) \le C\varepsilon\Big(\sum_{i=1}^{n} a_{m,i} + n\max\big(1, \sqrt{\log(\underline{\sigma}/\varepsilon)}\big)\Big).$$

Proof. We divide the proof in two steps.

(i) Reduction to unit variance. Let $x \ge 0$ arbitrary and let

$$W_{i_1,i_2} \coloneqq \frac{X_{i_1,i_2} - x}{\sigma_{i_1,i_2}} + \frac{x}{\underline{\sigma}}.$$

Then $\mu_{i_1,i_2} \coloneqq \mathbf{E}(W_{i_1,i_2}) = x\big(\frac{1}{\underline{\sigma}} - \frac{1}{\sigma_{i_1,i_2}}\big) \ge 0$ and $\mathbf{Var}(W_{i_1,i_2}) = 1$. Let $Z \coloneqq \min\max W$. Since the function min max is non-decreasing in each argument, we have

$$\begin{aligned}\mathbf{P}\big(|\min\max X - x| \le \varepsilon\big) &\le \mathbf{P}\Big(\Big|\min_{i_1}\max_{i_2} \frac{X_{i_1,i_2} - x}{\sigma_{i_1,i_2}}\Big| \le \frac{\varepsilon}{\underline{\sigma}}\Big)\\ &\le \sup_{y\in\mathbb{R}} \mathbf{P}\Big(\Big|\min_{i_1}\max_{i_2} \frac{X_{i_1,i_2} - x}{\sigma_{i_1,i_2}} + \frac{x}{\underline{\sigma}} - y\Big| \le \frac{\varepsilon}{\underline{\sigma}}\Big) \qquad (11)\\ &= \sup_{y\in\mathbb{R}} \mathbf{P}\Big(|Z - y| \le \frac{\varepsilon}{\underline{\sigma}}\Big).\end{aligned}$$

(ii) We proceed with bounding the density of Z. Since $W_{i_1,i_2} \sim \mathcal{N}(\mu_{i_1,i_2}, 1)$, by Lemma 8, assuming

(A) we have that the density of Z has the form

$$g_{n,m}(z) = \varphi(z) \sum_{i=1}^{n} H_i(z) G_i(z) \leq \varphi(z) \sum_{i=1}^{n} G_i(z) \tag{12}$$

We know from [4, Lemma 7] that

$$\varphi(z) G_i(z) \leq 2 \max(z, 1) \exp\Big(-\frac{\max(z - \overline{z} - a_{m,i}, 0)^2}{2}\Big) \leq 2(\overline{z} + a_{m,i} + 1),$$

where $\overline{z} := x\big(\frac{1}{\underline{\sigma}} - \frac{1}{\overline{\sigma}}\big)$, hence

$$g_{n,m}(z) \leq 2 \sum_{i=1}^{n} (\overline{z} + a_{m,i} + 1).$$

In particular, for all $y \in \mathbb{R}$ and $t > 0$ we have

$$\mathbf{P}\Big(|Z - y| \leq \frac{\varepsilon}{\underline{\sigma}}\Big) \leq 4 \frac{\varepsilon}{\underline{\sigma}} \sum_{i=1}^{n} (\overline{z} + a_{m,i} + 1)$$

and using step (i) we get

$$\mathbf{P}\big(|\min\max X - x| \leq \varepsilon\big) \leq \sup_{y \in \mathbb{R}} \mathbf{P}\Big(|Z - y| \leq \frac{\varepsilon}{\underline{\sigma}}\Big) \leq 4 \frac{\varepsilon}{\underline{\sigma}} \sum_{i=1}^{n} (\overline{z} + a_{m,i} + 1)$$

Repeating the argument with $x < 0$ one gets instead, one gets

$$\mathbf{P}\big(|\min\max X - x| \leq \varepsilon\big) \leq 4 \frac{\varepsilon}{\underline{\sigma}} \Big(n|x|\Big(\frac{1}{\underline{\sigma}} - \frac{1}{\overline{\sigma}}\Big) + \sum_{i=1}^{n} 1 + a_{m,i}\Big).$$

If $\underline{\sigma} = \overline{\sigma} = \sigma$ then $\mathbf{P}\big(|\min\max X - x| \leq \varepsilon\big) \leq \frac{4\varepsilon}{\sigma} \sum_{i=1}^{n} 1 + a_{m,i}$. On the contrary if $\underline{\sigma} < \overline{\sigma}$, note that the claim is true trivially for $\underline{\sigma}/\varepsilon < 1$. If $\varepsilon \leq \underline{\sigma}$ instead, for $|x| \geq \varepsilon + \overline{\sigma}\big(\frac{1}{n}\sum_{i=1}^{n} a_{m,i} + \sqrt{2\log(\underline{\sigma}/\varepsilon)}\big)$, since the min-max is a Lipschitz function, we can use the Gaussian deviation inequality (see [14, Theorem 7.1])

$$\begin{aligned}
\mathbf{P}\big(|\min\max X - x| \leq \varepsilon\big) &\leq \mathbf{P}\big(\min\max X \geq |x| - \varepsilon\big) \\
&\leq \mathbf{P}\Big(\min\max X \geq \frac{\overline{\sigma}}{n} \sum_{i=1}^{n} a_{m,i} + \overline{\sigma}\sqrt{2\log(\underline{\sigma}/\varepsilon)}\Big) \\
&\leq \mathbf{P}\Big(\min\max X \geq \mathbf{E}(\min\max X) + \overline{\sigma}\sqrt{2\log(\underline{\sigma}/\varepsilon)}\Big) \leq \frac{\varepsilon}{\underline{\sigma}},
\end{aligned}$$

where we used the fact that $\frac{\overline{\sigma}}{n}\sum_{i=1}^{n} a_{m,i} \geq \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}\big(\max_{i_2 \in [m]} X_{i,i_2}\big) \geq \mathbf{E}(\min\max X)$, since $\min\max X \leq \max_{i_2 \in [m]} X_{i,i_2}$ for all $i \in [n]$. When $|x| \leq \varepsilon + \overline{\sigma}\big(\frac{1}{n}\sum_{i=1}^{n} a_{m,i} + \sqrt{2\log(\underline{\sigma}/\varepsilon)}\big) \leq |x| \leq$

$\underline{\sigma} + \overline{\sigma}\big(\frac{1}{n}\sum_{i=1}^n a_{m,i} + \sqrt{2\log(\underline{\sigma}/\varepsilon)}\big)$ instead, we get

$$\begin{aligned}\mathbf{P}\big(|\min\max X - x| \leq \varepsilon\big) &\leq 4\frac{\varepsilon}{\underline{\sigma}}\Big(n|x|\Big(\frac{1}{\underline{\sigma}} - \frac{1}{\overline{\sigma}}\Big) + \sum_{i=1}^n 1 + a_{m,i}\Big) \\ &\leq 4\frac{\varepsilon}{\underline{\sigma}}\Big(n\Big(1 - \frac{\underline{\sigma}}{\overline{\sigma}}\Big) + n\Big(\frac{\overline{\sigma}}{\underline{\sigma}} - 1\Big)\sqrt{2\log(\underline{\sigma}/\varepsilon)} + \Big(\frac{\overline{\sigma}}{\underline{\sigma}} - 1\Big)\sum_{i=1}^n a_{m,i}\Big) \\ &\leq C\varepsilon\Big(\sum_{i=1}^n a_{m,i} + n\max\big(1, \sqrt{\log(\underline{\sigma}/\varepsilon)}\big)\Big),\end{aligned}$$

which concludes the proof. □

Remark. Rewriting the bound of Lemma 9 as

$$\mathcal{L}(\min\max X, \varepsilon) \leq Cn\varepsilon(\alpha_{nm} + \max\big(1, \sqrt{\log(\underline{\sigma}/\varepsilon)}\big),$$

we see that the multiplicative factor n cannot be improved. In fact, suppose that the matrix X is composed by n i.i.d. rows, which are copies of a m-dimensional standard Gaussian row vector X', and let $\alpha := \mathbf{E}(\max X')$. Then

$$\mathbf{P}(\min\max X \geq z) = \mathbf{P}(\max X' \geq z)^n,$$

and

$$\frac{\mathrm{d}}{\mathrm{d}z}\mathbf{P}(\min\max X \geq z) = n\mathbf{P}(\max X' \geq z)^{n-1}\frac{\mathrm{d}}{\mathrm{d}z}\mathbf{P}(\max X' \geq z).$$

Exploiting the sub-Gaussian deviation inequality for $\max X'$ we get that the last expression is less or equal than

$$n\exp\Big(-\frac{(n-1)\max(z-\alpha,0)^2}{2}\Big) \times 2\max(z,1)\exp\Big(-\frac{\max(z-\alpha,0)^2}{2}\Big),$$

which is uniformly bounded from above by $2n(\alpha+1)$, the desired order.

The following is a generalization of [5, Lemma 4.3].

Lemma 10 (Anti-concentration inequality, second variant)**.** *Let X be a (possibly non-centered) Gaussian random. For all $\varepsilon > 0$*

$$\mathcal{L}(\min\max X, \varepsilon) \leq 2\sqrt{2}\varepsilon\sum_{i_1=1}^n \frac{1}{\underline{\sigma}_{i_1}}\big(\sqrt{2} + \sqrt{\log m}\big) \leq 2\sqrt{2}\frac{\varepsilon}{\underline{\sigma}}n\big(\sqrt{2} + \sqrt{\log m}\big).$$

Proof. Let Σ_{i_1} be the covariance matrix of the row vector $X_{i_1,\cdot}$, in particular $X_{i_1,\cdot} \overset{\mathrm{d}}{=} \Sigma_{i_1}^{1/2}Z_{i_1} + \mu_{i_1}$ for some $Z_{i_1} \sim \mathcal{N}(0, \mathrm{I}_m)$ and $\mu_{i_1} \in \mathbb{R}^m$. Note that the i_2-th row of $\Sigma_{i_1}^{1/2}$ can be written as $\sigma_{i_1,i_2}v_{i_1,i_2}$

for some unit-norm row vector $v_{i_1,i_2} \in \mathbb{R}^{1\times m}$, which yields

$$B_x := \Big\{ \min_{i_1\in[n]} \max_{i_2\in[m]} (\Sigma_{i_1}^{1/2} Z_{i_1} + \mu_{i_1})_{i_2} \le x \Big\} = \bigcup_{i_1\in[n]} \Big\{ \forall i_2 \in [m]\ v_{i_1,i_2} Z_{i_1} \le \frac{x - \mu_{i_1,i_2}}{\sigma_{i_1,i_2}} \Big\} =: \bigcup_{i_1\in[n]} C_{i_1,x}.$$

In particular, since $\min\max X$ is absolutely continuous, its density f is given by

$$f(x) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \mathbf{P}(B_{x+\varepsilon} \setminus B_x),$$

for almost all $x \in \mathbb{R}$. For all $i_1 \in [n]$ we have

$$C_{i_1,x+\varepsilon} = \Big\{ \forall i_2 \in [m]\ v_{i_1,i_2} Z_{i_1} \le \frac{x + \varepsilon - \mu_{i_1,i_2}}{\sigma_{i_1,i_2}} \Big\} \subseteq \Big\{ \forall i_2 \in [m]\ v_{i_1,i_2} Z_{i_1} \le \frac{x - \mu_{i_1,i_2}}{\sigma_{i_1,i_2}} + \frac{\varepsilon}{\underline{\sigma}_{i_1}} \Big\} = C_{i_1,x}^{\varepsilon/\underline{\sigma}_{i_1}},$$

hence

$$B_{x+\varepsilon} \setminus B_x = \bigcup_{i_1\in[n]} C_{i_1,x+\varepsilon} \setminus \bigcup_{i_1\in[n]} C_{i_1,x} \subseteq \bigcup_{i_1\in[n]} C_{i_1,x}^{\varepsilon/\underline{\sigma}_{i_1}} \setminus \bigcup_{i_1\in[n]} C_{i_1,x} \subseteq \bigcup_{i_1\in[n]} C_{i_1,x}^{\varepsilon/\underline{\sigma}_{i_1}} \setminus C_{i_1,x}.$$

By the union bound, we deduce that

$$f(x) \le \lim_{\varepsilon\to 0} \sum_{i_1=1}^{n} \frac{1}{\varepsilon} \mathbf{P}\big(C_{i_1,x}^{\varepsilon/\underline{\sigma}_{i_1}} \setminus C_{i_1,x}\big).$$

Using Nazarov's inequality (see [17]) on each term of the last sum gives that

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \mathbf{P}\big(C_{i_1,x}^{\varepsilon/\underline{\sigma}_{i_1}} \setminus C_{i_1,x}\big) \le \frac{2\sqrt{2}}{\underline{\sigma}_{i_1}} \big(\sqrt{2} + \sqrt{\log m}\big),$$

which allows us to conclude. □

Lemma 11. *For all $b \in \mathbb{R}$ and all $\varepsilon > 0$ there exists a twice continuously differentiable function $g = g_{b,\varepsilon}$ such that $\mathbf{1}_{(-\infty,b]} \le g \le \mathbf{1}_{(-\infty,b+\varepsilon]}$ and $\|g'\|_\infty \le 2/\varepsilon$, $\|g''\|_\infty \le 6/\varepsilon^2$.*

Proof. Following the example of [ref], one can see if one defines the function as

$$g_{b,\varepsilon}(z) := \begin{cases} 1 & \text{if } z \le b, \\ 30 \displaystyle\int_{\frac{z-b}{\varepsilon}}^{1} t^2(1-t)^2 \mathrm{d}t & \text{if } b < z < b + \varepsilon, \\ 0 & \text{if } z \ge b + \varepsilon, \end{cases}$$

it is straightforward to compute the bounds of its first two derivatives. □

The next statement represents a generalization to $n > 1$ of [4, Theorem 1] (case of F Gaussian) and [13, Theorem 2.1] (general case). Its proof uses ideas inspired by [6], a work which achieves an improvement of [4, Theorem 2] and [13, Theorem 2.1].

Proposition 1. *For all functions g as in Lemma 11, it holds*

$$\big|\mathbf{E}\big(g\circ f_1^{\beta,\delta}(F)\big)-\mathbf{E}\big(g\circ f_1^{\beta,\delta}(X)\big)\big|\le\Big(\frac{3}{\varepsilon^2}+\beta(1+\delta)\frac{2}{\varepsilon}\Big)\Big(\varepsilon+\frac{\log m}{\beta}+\frac{\log n}{\beta\delta}\Big)\frac{2n\big(2+\sqrt{2\log m}\big)}{\underline{\sigma}}\Delta.$$

Proof. We may assume that F and X are independent, without loss of generality. Consider their interpolation given by

$$Z(t):=\sqrt{t}F+\sqrt{1-t}X,$$

for all $t\in[0,1]$. Let $\varphi:=g_{x,\varepsilon}\circ f_1^{\beta,\delta}$. Then, $\frac{\partial^2\varphi}{\partial_{x_{i_1,i_2}}\partial_{x_{j_1,j_2}}}(y)=0$ whenever $f_1^{\beta,\delta}(y)\notin[x,x+\varepsilon]$. In view of Lemma 1 this also implies that

$$\frac{\partial^2\varphi}{\partial_{x_{i_1,i_2}}\partial_{x_{j_1,j_2}}}(Z(t))=0\quad\text{if}\quad\min\max Z(t)\notin\Big[x-\frac{\log m}{\beta},x+\varepsilon+\frac{\log n}{\beta\delta}\Big]=:I. \tag{13}$$

Now consider the function $\Psi(t):=\mathbf{E}\big(\varphi(Z(t))\big)$. Then Ψ is differentiable in $[0,1]$ with derivative equal to

$$\Psi'(t)=\frac{1}{2}\sum_{(j_1,j_2)\in[n]\times[m]}\mathbf{E}\Big(\frac{\partial\varphi}{\partial x_{j_1,j_2}}(Z(t))\Big(\frac{F_{j_1,j_2}}{\sqrt{t}}-\frac{X_{j_1,j_2}}{\sqrt{1-t}}\Big)\Big).$$

By independence and integration by parts we deduce that

$$\sum_{(j_1,j_2)\in[n]\times[m]}\mathbf{E}\Big(\frac{\partial\varphi}{\partial x_{j_1,j_2}}(Z(t))\frac{X_{j_1,j_2}}{\sqrt{1-t}}\Big)=\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}}\mathbf{E}\Big(\frac{\partial^2\varphi}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(Z(t))\sigma_{i_1,i_2;j_1,j_2}\Big).$$

Analogously, arguing as in the proof of [20, Theorem 6.1.1] yields that

$$\sum_{(j_1,j_2)\in[n]\times[m]}\mathbf{E}\Big(\frac{\partial\varphi}{\partial x_{j_1,j_2}}(Z(t))\frac{F_{j_1,j_2}}{\sqrt{t}}\Big)=\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}}\mathbf{E}\Big(\frac{\partial^2\varphi}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(Z(t))\langle DF_{i_1,i_2},-DL^{-1}F_{j_1,j_2}\rangle\Big).$$

Hence, we obtain

$$\Psi'(t)=\frac{1}{2}\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}}\mathbf{E}\Big(\frac{\partial^2\varphi}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(Z(t))\big(\langle DF_{i_1,i_2},-DL^{-1}F_{j_1,j_2}\rangle-\sigma_{i_1,i_2;j_1,j_2}\big)\Big).$$

Because of (13), the above identity can be also rewritten as

$$\Psi'(t)=\frac{1}{2}\sum_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}}\mathbf{E}\Big(\mathbf{1}_I(\min\max Z(t))\frac{\partial^2\varphi}{\partial x_{i_1,i_2}\partial x_{j_1,j_2}}(Z(t))\big(\langle DF_{i_1,i_2},-DL^{-1}F_{j_1,j_2}\rangle-\sigma_{i_1,i_2;j_1,j_2}\big)\Big).$$

Now we can use Lemma 7 with $k=1$ and Lemma 11 to perform the bound

$$|\Psi'(t)| \le \Big(\frac{3}{\varepsilon^2} + \beta(1+\delta)\frac{2}{\varepsilon}\Big)\mathbf{E}\Big(\mathbf{1}_I(\min\max Z(t)) \max_{i_1,i_2,j_1,j_2}\big(\langle DF_{i_1,i_2}, -DL^{-1}F_{j_1,j_2}\rangle - \sigma_{i_1,i_2;j_1,j_2}\big)\Big).$$

Since X is independent of the collection $\{F; (DF_{i_1,i_2})_{i_1,i_2\in[n]\times[m]}; (DL^{-1}F_{j_1,j_2})_{j_1,j_2\in[n]\times[m]}\}$, using the law of total expectation we can write

$$\begin{aligned}&\mathbf{E}\Big(\mathbf{1}_I(\min\max Z(t)) \max_{i_1,i_2,j_1,j_2}\big(\langle DF_{i_1,i_2}, -DL^{-1}F_{j_1,j_2}\rangle - \sigma_{i_1,i_2;j_1,j_2}\big)\Big)\\ &\qquad = \mathbf{E}\Big(\mathbf{P}\big(\min\max Z(t)\in I \;\big|\; F\big)\mathbf{E}\Big(\max_{i_1,i_2,j_1,j_2}\big(\langle DF_{i_1,i_2}, -DL^{-1}F_{j_1,j_2}\rangle - \sigma_{i_1,i_2;j_1,j_2}\big) \;\big|\; F\Big)\Big).\end{aligned}$$

Now, for all $\mu\in\mathbb{R}^{n\times m}$ fixed we have

$$\begin{aligned}\mathbf{P}\big(\min\max Z(t)\in I \;\big|\; F=\mu\big) &\le \mathcal{L}\Big(\min\max(\sqrt{t}\mu + \sqrt{1-t}X), \frac{1}{2}\Big(\varepsilon + \frac{\log m}{\beta} + \frac{\log n}{\beta\delta}\Big)\Big)\\ &\le \Big(\varepsilon + \frac{\log m}{\beta} + \frac{\log n}{\beta\delta}\Big)\frac{n\big(\sqrt{2}+\sqrt{\log m}\big)}{2\underline{\sigma}\sqrt{1-t}},\end{aligned}$$

where in the last inequality we used Lemma 10. Since the previous bound is uniform in μ we can eventually conclude that

$$\begin{aligned}\big|\mathbf{E}\big(\varphi(F)\big) - \mathbf{E}\big(\varphi(X)\big)\big| &\le \int_0^1 |\Psi'(t)|\,\mathrm{d}t\\ &\le \Big(\frac{3}{\varepsilon^2} + \beta(1+\delta)\frac{2}{\varepsilon}\Big)\Big(\varepsilon + \frac{\log m}{\beta} + \frac{\log n}{\beta\delta}\Big)\frac{n\big(2+\sqrt{2\log m}\big)}{\underline{\sigma}}\Delta\int_0^1\frac{\mathrm{d}t}{\sqrt{1-t}},\\ &= \Big(\frac{3}{\varepsilon^2} + \beta(1+\delta)\frac{2}{\varepsilon}\Big)\Big(\varepsilon + \frac{\log m}{\beta} + \frac{\log n}{\beta\delta}\Big)\frac{2n\big(2+\sqrt{2\log m}\big)}{\underline{\sigma}}\Delta,\end{aligned}$$

which concludes the proof. □

Lemma 12. *For all $\varepsilon>0$ and all $a,b\in\mathbb{R}$, we have*

$$\mathbf{P}\big(\min\max F\le b\big) \le \mathbf{P}(\min\max X \le b+3\varepsilon) + Cn\log nm\sqrt{\log m}\,\frac{\Delta}{\varepsilon\underline{\sigma}}, \tag{14}$$

$$\mathbf{P}\big(\min\max F\ge a\big) \le \mathbf{P}(\min\max X \ge a-3\varepsilon) + Cn\log nm\sqrt{\log m}\,\frac{\Delta}{\varepsilon\underline{\sigma}}, \tag{15}$$

where $C>0$ is a universal constant.

Proof. Fix $\varepsilon>0$ and let $\beta = \frac{\log m}{\varepsilon}$, $\delta = \frac{\log n}{\log m}$, so that $\frac{\log m}{\beta} = \frac{\log n}{\beta\delta} = \varepsilon$. Then, by Lemma 1 we get

$$\mathbf{P}\big(\min\max F\le b\big) \le \mathbf{P}\big(f_1^{\beta,\delta}(F)\le b+\varepsilon\big) = \mathbf{E}\big(\mathbf{1}_{(-\infty,b+\varepsilon]}(f_1^{\beta,\delta}(F))\big).$$

Now we can apply Lemma 11. For such $g = g_{b+\varepsilon,\varepsilon}$ by monotonicity we get

$$\mathbf{E}\big(\mathbf{1}_{(-\infty,b+\varepsilon]}(f_1^{\beta,\delta}(F))\big) \le \mathbf{E}\big(g(f_1^{\beta,\delta}(F))\big).$$

Since $\beta(1+\delta) = \frac{\log nm}{\varepsilon}$, by Proposition 1 we get

$$\begin{aligned}\left|\mathbf{E}\big(g(f_1^{\beta,\delta}(F))\big) - \mathbf{E}\big(g(f_1^{\beta,\delta}(X))\big)\right| &\leq \Big(\frac{3}{2\varepsilon^2} + \frac{2\log nm}{\varepsilon^2}\Big)(3\varepsilon)2n\frac{2+\sqrt{2\log m}}{\underline{\sigma}}\Delta \\ &= 3n(3+4\log nm)(2+\sqrt{2\log m})\frac{\Delta}{\varepsilon\underline{\sigma}} \\ &\leq Cn\log nm\sqrt{\log m}\frac{\Delta}{\varepsilon\underline{\sigma}},\end{aligned}$$

for some absolute constant $C > 0$. In analogy with above, it also holds

$$\mathbf{E}\big(g(f_1^{\beta,\delta}(X))\big) \leq \mathbf{E}\big(\mathbf{1}_{(-\infty,b+2\varepsilon]}(f_1^{\beta,\delta}(X))\big) \leq \mathbf{P}(\min\max X \leq b+3\varepsilon),$$

which allows us to obtain (14). Analogously, we get (15) by using the reflected function $g_{-a+\varepsilon,\varepsilon}(-\cdot)$. □

Finally, we need one more technical Lemma, similar to the one in [13, Lemma A.3].

Lemma 13. *Consider two random variables U and V. Suppose that there exist two positive constants ε_1 and ε_2 such that for all $a, b \in \mathbb{R}$*

$$\mathbf{P}(U \leq b) \leq \mathbf{P}(V \leq b + \varepsilon_1) + \varepsilon_2$$

and

$$\mathbf{P}(U \geq a) \leq \mathbf{P}(V \geq a - \varepsilon_1) + \varepsilon_2.$$

Then

$$\sup_{x\in\mathbb{R}}\big|\mathbf{P}(U \leq x) - \mathbf{P}(V \leq x)\big| \leq \sup_{x\in\mathbb{R}}\mathbf{P}(|V-x| \leq \varepsilon_1) + \varepsilon_2.$$

Proof. Using the first assumption we can write for all $x \in \mathbb{R}$,

$$\begin{aligned}\mathbf{P}(U \leq x) - P(V \leq x) &= \big(\mathbf{P}(U \leq x) - P(V \leq x+\varepsilon_1)\big) + \big(P(V \leq x+\varepsilon_1) - P(V \leq x)\big) \\ &\leq \varepsilon_2 + \sup_{x\in\mathbb{R}}\mathbf{P}(|V-x| \leq \varepsilon_1).\end{aligned}$$

The second assumption provides the converse inequality. □

We are now ready to prove our main results.

Proof of Theorem 2 (a). Note that if $\Delta \geq 1$ the result is trivially true. So we can assume $\Delta \in (0,1)$.

Lemma 12 allows to use Lemma 13 with $\varepsilon_1 = 3\varepsilon$ and $\varepsilon_2 = Cn\log nm\sqrt{\log m}\,\frac{\Delta}{\varepsilon\underline{\sigma}}$, to get

$$\begin{aligned}&\sup_{x\in\mathbb{R}}\big|\mathbf{P}(\min\max F \leq x) - \mathbf{P}(\min\max X \leq x)\big| \\ &\qquad\leq \sup_{x\in\mathbb{R}}\mathbf{P}(|\min\max X - x| \leq 3\varepsilon) + Cn\log nm\sqrt{\log m}\,\frac{\Delta}{\varepsilon\underline{\sigma}}\end{aligned} \tag{16}$$

We can now use Lemma 9 to estimate the first summand, hence

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\big|\mathbf{P}(\min\max F\le x)-\mathbf{P}(\min\max X\le x)\big|\\
&\qquad\le C'\Big[3\varepsilon\Big(\sum_{i=1}^{n}a_{m,i}+n\max\big(1,\sqrt{\log(\underline{\sigma}/3\varepsilon)}\big)\Big)+n\log nm\sqrt{\log m}\,\frac{\Delta}{\underline{\sigma}\varepsilon}\Big]\\
&\qquad\le C''n\Big[\varepsilon\max\big(1,\alpha_{nm},\sqrt{\log(1/\varepsilon)}\big)+\log nm\sqrt{\log m}\,\frac{\Delta}{\varepsilon}\Big].
\end{aligned}$$

We can estimate the right hand side by choosing

$$\varepsilon=\frac{\sqrt{\log nm}(\log m)^{1/4}\sqrt{\Delta}}{\max\big(1,\sqrt{\alpha_{nm}},\log(1/\Delta)^{1/4}\big)},$$

which yields

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\big|\mathbf{P}(\min\max F\le x)-\mathbf{P}(\min\max X\le x)\big|\\
&\qquad\le\frac{C''n\sqrt{\log nm}(\log m)^{1/4}\sqrt{\Delta}}{\max\big(1,\sqrt{\alpha_{nm}},\log(1/\Delta)^{1/4}\big)}\Big[\max\big(1,\alpha_{nm},\sqrt{\log(1/\varepsilon)}\big)+\max\big(1,\alpha_{nm},\sqrt{\log(1/\Delta)}\big)\Big]\\
&\qquad\le\frac{2C''n\sqrt{\log nm}(\log m)^{1/4}\sqrt{\Delta}}{\max\big(1,\sqrt{\alpha_{nm}},\log(1/\Delta)^{1/4}\big)}\max\big(1,\alpha_{nm},\sqrt{\log(1/\varepsilon)},\sqrt{\log(1/\Delta)}\big).
\end{aligned}$$

Since

$$\log\Big(\frac{1}{\varepsilon}\Big)=\frac{1}{2}\log\frac{\max\big(1,\alpha_{nm},\sqrt{\log(1/\Delta)}\big)}{\Delta}=\frac{1}{2}\Big[\max\Big(0,\log\alpha_{nm},\frac{1}{2}\log\log\Big(\frac{1}{\Delta}\Big)\Big)+\log\Big(\frac{1}{\Delta}\Big)\Big],$$

we have that

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\big|\mathbf{P}(\min\max F\le x)-\mathbf{P}(\min\max X\le x)\big|\\
&\qquad\le\frac{C'''n\sqrt{\log nm}(\log m)^{1/4}\sqrt{\Delta}}{\max\big(1,\sqrt{\alpha_{nm}},\log(1/\Delta)^{1/4}\big)}\max\big(1,\alpha_{nm},\sqrt{\log(1/\Delta)}\big)\\
&\qquad=C'''n\sqrt{\log nm}(\log m)^{1/4}\max\big(1,\sqrt{\alpha_{nm}},\log(1/\Delta)^{1/4}\big)\sqrt{\Delta},
\end{aligned}$$

which concludes the first claim. □

Proof of Theorem 2 (b). As in eq. (16) we have We can apply Lemma 10 to obtain

$$\sup_{x\in\mathbb{R}}\big|\mathbf{P}(\min\max F\le x)-\mathbf{P}(\min\max X\le x)\big|\le\frac{C}{\underline{\sigma}}\Big(\varepsilon n\sqrt{\log m}+\frac{n\log nm\sqrt{\log m}\Delta}{\varepsilon}\Big)$$

The last expression is minimized by choosing $\varepsilon=\sqrt{\Delta\log nm}$, which yields

$$\sup_{x\in\mathbb{R}}\big|\mathbf{P}(\min\max F\le x)-\mathbf{P}(\min\max X\le x)\big|\le\frac{2C}{\underline{\sigma}}n\sqrt{\Delta\log nm\log m},$$

concluding the proof. □

4 Application to matrices of multiple stochastic integrals

We will now apply our previous findings to matrices of multiple Wiener-Itô integrals, as introduced in Section 3.1 (whose setting will prevail throughout).

4.1 A general estimate

Let $q, N \in \mathbb{N}$ and consider three sequences of natural numbers $d = d(N)$, $n = n(N)$ and $m = m(N)$. For every $(i_1, i_2) \in [n] \times [m]$, we consider a random variable of the type

$$F_{i_1,i_2} = F^N_{i_1,i_2} := I_q(f^N_{i_1,i_2}),$$

where I_q indicates a multiple stochastic integral of order $q \geq 2$ and $f_{i_1,i_2} = f^N_{i_1,i_2} \in \mathfrak{H}^{\odot q}$ (when there is no risk of confusion, and in order to simplify the presentation, we will sometimes avoid to write the superscript N).

Proposition 2. *Suppose that for all $N \in \mathbb{N}$, $X^N = (X^N_{i_1,i_2})_{(i_1,i_2)\in[n]\times[m]}$ is a centered Gaussian random matrix with covariance matrix $(\sigma^N_{i_1,i_2;j_1,j_2})_{(i_1,i_2)\in[n]\times[m]}$ and $F^N = (F^N_{i_1,i_2})_{(i_1,i_2)\in[n]\times[m]}$ is the random matrix described as above. Suppose moreover that there exists a constant $c > 0$ such that $\underline{\sigma}^N \geq c$ for all $N \in \mathbb{N}$ (where we used the same notation introduced at the beginning of Section 3.2). If*

$$A := \max_{\substack{(i_1,i_2)\in[n]\times[m] \\ (j_1,j_2)\in[n]\times[m]}} \left|\sigma^N_{i_1,i_2;j_1,j_2} - \mathbf{E}[F_{i_1,i_2}F_{j_1,j_2}]\right| n^2 (\log m)(\log nm)$$

and

$$B := \max_{(i_1,i_2)\in[n]\times[m]} \left(\mathbf{E}(F^4_{i_1,i_2}) - 3\mathbf{E}(F^2_{i_1,i_2})^2\right) n^4 (\log m)^2 (\log nm)^{2q},$$

then there exists a constant $C > 0$ independent of N such that

$$\sup_{x\in\mathbb{R}} \left|\mathbf{P}(\min\max F^N \leq x) - \mathbf{P}(\min\max X^N \leq x)\right| \leq C(A^{1/2} + B^{1/4}).$$

Remark. The content of Proposition 2 can be regarded as further confirmation of the so-called *(multidimensional) fourth moment phenomenon* (see e.g. [20, Chapters 5 and 6]). According to this notion, if $(F_n)_{n\in\mathbb{N}}$ is a sequence of random vectors whose components belong to Gaussian Wiener chaoses of fixed orders, then F_n verifies a multidimensional central limit theorem if and only if the covariance matrices of F_n converges pointwise, and the fourth cumulants of its components converge to zero. The main contribution of Proposition 2 is that of providing (in the case of min-max statistics) a bound with explicit dimensional dependencies. See [19] for a constantly updated repository of papers connected to fourth moment theorems and related results.

Remark. Note that for $n = 1$ and $q = 2$ we recover [13, Theorem 3.1].

Proof. We know from [13, Lemma 2.2] that, for Δ as defined in (9), one has that

$$\Delta \leq \max_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}} \big|\sigma^N_{i_1,i_2;j_1,j_2} - \mathbf{E}[F_{i_1,i_2}F_{j_1,j_2}]\big| \\ + C_q \log^{q-1}(2n^2m^2 - 1 + e^{q-2}) \max_{(i_1,i_2)\in[n]\times[m]} \sqrt{\mathbf{E}(F^4_{i_1,i_2}) - 3\mathbf{E}(F^2_{i_1,i_2})^2},$$

for some constant C_q depending only on q. Note that since q is fixed we can bound $\log^{q-1}(2n^2m^2 - 1 + e^{q-2}) \leq \tilde{c}\log^{q-1}(nm)$ for some constant $\tilde{c} > 0$. The conclusion is reached by applying Theorem 2 (b) as follows

$$\sup_{x\in\mathbb{R}} \big|\mathbf{P}(\min\max F^N \leq x) - \mathbf{P}(\min\max X^N \leq x)\big| \\ \leq C_1 \Big(\max_{\substack{(i_1,i_2)\in[n]\times[m]\\(j_1,j_2)\in[n]\times[m]}} \big|\sigma^N_{i_1,i_2;j_1,j_2} - \mathbf{E}[F_{i_1,i_2}F_{j_1,j_2}]\big| n^2(\log m)(\log nm)\Big)^{1/2} \\ + C_2 \Big(\max_{(i_1,i_2)\in[n]\times[m]} \big(\mathbf{E}(F^4_{i_1,i_2}) - 3\mathbf{E}(F^2_{i_1,i_2})^2\big) n^4 (\log m)^2 (\log nm)^{2q}\Big)^{1/4},$$

where $C_1, C_2 > 0$ are constants that do not depend on N, thanks to the fact that $\underline{\sigma}^N$ is bounded from below by an absolute constant. □

4.2 An illustration

We will now briefly illustrate our findings with an example inspired by the statistical procedures for testing the absence of lead-lag effects in time series, as put forward in [13, Section 4.1]. See the remark at the end of this section for a statistical interpretation of our findings.

We start by considering a 4-dimensional Gaussian process $Z(t) = (B_1(t), B_2(t), \tilde{B}_1(t), \tilde{B}_2(t))$ on the real line, with the following characteristics: (a) each coordinate of Z is a standard Brownian motion issued from zero, (b) B_i and $\tilde{B}_i$ are independent, for $i = 1, 2$, (c) the dependence among other pairs of coordinates of Z is arbitrary. We also write $B := (B_1, B_2)$ and $\tilde{B} := (\tilde{B}_1, \tilde{B}_2)$, and denote by $W^\theta(\cdot) = \tilde{B}(\cdot - \theta)$ the process $\tilde{B}$ translated by $\theta \in \mathbb{R}$.

For some $T, b, w > 0$, we now suppose to observe the process B, respectively W^θ, at a finite set of points in time $\mathcal{T}_B = \{0, \frac{T}{bN}, \frac{2T}{bN}, ..., \frac{\lfloor bN\rfloor T}{bN}\}$, respectively $\mathcal{T}_{W^\theta} = \{0, \frac{T}{wN}, \frac{2T}{wN}, ..., \frac{\lfloor wN\rfloor T}{wN}\}$, in such a way that $|\mathcal{T}_B| \sim bN$ and $|\mathcal{T}_W| \sim wN$. For every $\theta \in \mathbb{R}$, we also introduce the following two (centered) statistics $U^N_1(\theta)$, $U^N_2(\theta)$, that can be seen as special cases of the general class defined in [13, Introduction and Section 4.1],

$$U^N_1(\theta) = \sum_{i=1}^{\lfloor bN\rfloor} \sum_{j=1}^{\lfloor wN\rfloor} \Big(B_1\Big(\frac{iT}{bN}\Big) - B_1\Big(\frac{(i-1)T}{bN}\Big)\Big) \\ \times \Big(W^\theta_1\Big(\frac{jT}{wN}\Big) - W^\theta_1\Big(\frac{(j-1)T}{wN}\Big)\Big)\mathbf{1}_{\{(\frac{(i-1)T}{bN},\frac{iT}{bN}]\cap(\frac{(j-1)T}{wN},\frac{jT}{wN}]\neq\emptyset\}}$$

and

$$U_2^N(\theta) = \sum_{i=1}^{\lfloor bN \rfloor} \sum_{j=1}^{\lfloor wN \rfloor} \Big(B_2\Big(\frac{iT}{bN}\Big) - B_2\Big(\frac{(i-1)T}{bN}\Big)\Big) \times \Big(W_2^\theta\Big(\frac{jT}{wN}\Big) - W_2^\theta\Big(\frac{(j-1)T}{wN}\Big)\Big) \mathbf{1}_{\{(\frac{(i-1)T}{bN}, \frac{iT}{bN}] \cap (\frac{(j-1)T}{wN}, \frac{jT}{wN}] \neq \emptyset\}}.$$

We are interested in the fluctuations of the following statistic

$$\sqrt{N} \min_{i \in \{1,2\}} \max_{\theta \in \Theta_N} |U_i^N(\theta)|,$$

where Θ_N is an index set such that $|\Theta_N| = m(N) \in \mathbb{N}$. In order to study the asymptotic properties of the aforementioned object, it is appropriate to apply Proposition 2 to the matrix

$$F^N = \big(\sqrt{N} \big| U_i^N(\theta) \big| \big)_{(i,\theta) \in [2] \times \Theta_N},$$

setting $q = 2$ and $n \equiv 2$ (see the final claim in the Remark following Theorem 2).

Proposition 3. *Suppose that for all $N \in \mathbb{N}$, $X^N = (X_{i,\theta}^N)_{(i,\theta) \in [2] \times \Theta_N}$ is a $2 \times m$ centered Gaussian random matrix whose columns have the same covariance matrix as the respective columns of the random matrix F^N, and denote by $|X|^N$ the matrix whose entries are given by the absolute values of the corresponding entries of X^N. Then there is an absolute constant $c > 0$*

$$\sup_{x \in \mathbb{R}} \big| \mathbf{P}(\min \max |X|^N \leq x) - \mathbf{P}(\min \max F^N \leq x) \big| \leq c \, \frac{(\log m)^{3/2}}{N^{1/4}}.$$

Proof. Note that the construction of $\mathcal{T}_B$ and $\mathcal{T}_{W^\theta}$ ensures that assumptions [A1] and [A2] of [13, Section 4.1] are met. Moreover, as in [13, Lemma B.7], we have that

$$\max_{i,\theta} \mathbf{E}((\sqrt{N} U_i^N(\theta))^4) - 3\mathbf{E}((\sqrt{N} U_i^N(\theta))^2)^2 \leq \frac{c}{N}$$

for some constant $c > 0$ depending only on ρ_1 and ρ_2. Since n is fixed and $q = 2$, we recover the claimed inequality from Proposition 2. □

Proposition 3 implies that $\min \max F^N$ is asymptotically close to the min-max of a suitable Gaussian random matrix as long as $\log |\Theta_N| = o(N^{1/6})$.

Remark. There is no conceptual difficulty in extending the previous convergence results to the case in which the correlation between B_i and $\tilde{B}_i$ equals some non zero parameter ρ_i, $i = 1, 2$. In this case, given a fixed nonzero $a \in (-1, 1)$, the corresponding modification of the statistics $U_1^N(\theta)$, $U_2^N(\theta)$ can in principle be used to solve the following statistical hypothesis testing problem:

$$H_0 : \rho_1 = 0 \text{ or } \rho_2 = 0, \quad \text{(null hypothesis)}$$

$$H_1 : \rho_1 = \rho_2 = a.$$

We regard this line of investigation as a separate topic, and leave it open for further investigation.

Acknowledgments

The authors would like to thank the anonymous referees for their comments and for suggesting an alternative proof of Proposition 1 which led to an improvement on the bound of Theorem 2. The authors were supported by the FNR grant FoRGES (R-AGR-3376-10).

References

[1] Adler, R. J. and Taylor, J. E. *Random fields and geometry.* Springer, Berlin, 2007.

[2] Alexander, R. "Lipschitzian mappings and total mean curvature of polyhedral surfaces. I". *Trans. Amer. Math. Soc.* 288.2 (1985), pp. 661–678.

[3] Chatterjee, S. *An error bound in the Sudakov-Fernique inequality.* arXiv:math/0510424.

[4] Chernozhukov, V., Chetverikov, D., and Kato, K. "Comparison and anti-concentration bounds for maxima of Gaussian random vectors". *Probab. Theory Related Fields* 162.1-2 (2015), pp. 47–70.

[5] Chernozhukov, V., Chetverikov, D., and Kato, K. "Empirical and multiplier bootstraps for suprema of empirical processes of increasing complexity, and related Gaussian couplings". *Stochastic Processes and their Applications* 126.12 (2016), pp. 3632–3651.

[6] Chernozhukov, V., Chetverikov, D., Kato, K., and Koike, Y. "Improved central limit theorem and bootstrap approximation in high dimension". *Ann. Statist. (forthcoming)* (2022).

[7] Debicki, K., Hashorva, E., Ji, L., and Ling, C. "Comparison Inequalities for Order Statistics of Gaussian Arrays". *ALEA* 14 (2017), pp. 93–116.

[8] Fernique, X. "Regularité des trajectoires des fonctions aléatoires gaussiennes". *École d'Été de Probabilités de Saint-Flour, IV-1974.* 1975, 1–96. Lecture Notes in Math., Vol. 480.

[9] Gordon, Y. "Some inequalities for Gaussian processes and applications". *Israel J. Math.* 50.4 (1985), pp. 265–289.

[10] Gordon, Y. "Elliptically contoured distributions". *Probab. Theory Related Fields* 76.4 (1987), pp. 429–438.

[11] Gordon, Y. "Majorization of Gaussian processes and geometric applications". *Probab. Theory Related Fields* 91.2 (1992), pp. 251–267.

[12] Kahane, J.-P. "Une inégalité du type de Slepian et Gordon sur les processus gaussiens". *Israel J. Math.* 55.1 (1986), pp. 109–110.

[13] Koike, Y. “Gaussian approximation of maxima of Wiener functionals and its application to high-frequency data”. *Ann. Statist.* 47.3 (2019), pp. 1663–1687.

[14] Ledoux, M. *The concentration of measure phenomenon*. Vol. 89. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2001, pp. x+181.

[15] Liu, Y. and Xie, J. “Accurate and efficient p-value calculation via Gaussian approximation: A novel Monte-Carlo method”. *J. Americ. Statistic. Assoc.* 114 (2019a).

[16] Liu, Y. and Xie, J. “Supplement to Liu and Xie (2019a)” (2019b).

[17] Nazarov, F. “On the maximal perimeter of a convex set in $\mathbb{R}^n$ with respect to a Gaussian measure”. *Geometric aspects of functional analysis*. Vol. 1807. Lecture Notes in Math. Springer, Berlin, 2003, pp. 169–187.

[18] Nourdin, I., Peccati, G., and Viens, F. G. “Comparison inequalities on Wiener space”. *Stochastic Process. Appl.* 124.4 (2014), pp. 1566–1581.

[19] Nourdin, I. *Malliavin-Stein*. `https://sites.google.com/site/malliavinstein/home`.

[20] Nourdin, I. and Peccati, G. *Normal approximations with Malliavin calculus*. Vol. 192. Cambridge Tracts in Mathematics. From Stein’s method to universality. Cambridge University Press, Cambridge, 2012, pp. xiv+239.

[21] Sudakov, V. N. “Gaussian random processes, and measures of solid angles in Hilbert space”. *Dokl. Akad. Nauk SSSR* 197 (1971), pp. 43–45.

[22] Sudakov, V. N. “Geometric problems in the theory of infinite-dimensional probability distributions”. *Proc. Steklov Inst. Math.* 2 (1979), pp. i–v, 1–178.

[23] Vitale, R. A. “Some comparisons for Gaussian processes”. *Proc. Amer. Math. Soc.* 128.10 (2000), pp. 3043–3046.