Electronic Journal of Probability

Electron. J. Probab. **31** (2026), article no. 13, 1–46.
ISSN: 1083-6489 https://doi.org/10.1214/26-EJP1472

Tall random matrices with chaotic entries: Approximate isometry and covariance estimation*

Giovanni Luca Torrisi†

Abstract

Let $\mathbf{W}_n$, $n \in \mathbb{N}$, be the Wishart matrix associated to an $n \times d_n$ random matrix $\mathbf{F}_n$ with independent rows $(\mathbf{F}_n)_i$, $i = 1, \dots, n$, whose components are in the q_i-th, $q_i \in \mathbb{N}$, Wiener chaos of an isonormal Gaussian process X_i. In the case when the random matrix $\mathbf{F}_n$ is "tall", i.e., $d_n = o(n)$, as $n \to \infty$, we provide sufficient conditions on the kernels and the orders of the Wiener chaoses which ensure that for any $\varepsilon > 0$ (small enough) there exists n_ε such that $\mathbb{P}(\|\mathbf{W}_n - \mathbf{I}_{d_n}\|_{op} \le \varepsilon) \ge \mathcal{L}_{n,\varepsilon}$ for any $n > n_\varepsilon$, and $\mathcal{L}_{n,\varepsilon} \to 1$, as $n \to \infty$. Here $\mathbf{I}_{d_n}$ denotes the $d_n \times d_n$ identity matrix and the symbol $\|\cdot\|_{op}$ denotes the operator norm of a square matrix. We explicitly compute the lower bound $\mathcal{L}_{n,\varepsilon}$ in terms of n, d_n and ε. As a by-product, we get that the random matrix $\mathbf{F}_n/\sqrt{n}$ is an approximate isometry with high probability, as $n \to \infty$. We apply the main result to a covariance estimation problem, which is of potential interest in data analysis. The theory is illustrated by means of matrices $\mathbf{F}_n$ whose entries are suitable stochastic integrals of Ornstein-Uhlenbeck processes.

Keywords: random matrices; Wiener chaos; approximate isometry; high-dimensional covariance estimation.
MSC2020 subject classifications: 60B20; 60H07.
Submitted to EJP on November 2, 2024, final version accepted on January 5, 2026.

1 Introduction

Wishart matrices, introduced in [44], are random matrices of particular importance in statistics as they constitute special classes of sample covariance matrices which find applications in various fields. Let $\mathbf{F}_{n,d} = (F_{ij})_{1 \le i \le n,\, 1 \le j \le d}$, $n, d \in \mathbb{N} := \{1, 2, \dots\}$, be a random matrix with real entries. Its associated Wishart matrix is the symmetric $d \times d$ random matrix $\mathbf{W}_{n,d} := \frac{1}{n}\mathbf{F}_{n,d}^\top \mathbf{F}_{n,d}$, where $\mathbf{F}_{n,d}^\top$ denotes the transpose of the matrix $\mathbf{F}_{n,d}$. If the random variables F_{ij} are i.i.d., centered with unit variance, d is fixed and $n \to \infty$, then (i) by the Strong Law of Large Numbers $\mathbf{W}_{n,d}$ converges almost surely to the

*The author was supported by the European Union's Horizon 2020 research and innovation program WARIFA, under grant agreement No 101017385. The author thanks two anonymous referees for their careful reading of the paper and the INdAM group GNAMPA for the financial support.
†Consiglio Nazionale delle Ricerche. Rome, Italy. E-mail: giovanniluca.torrisi@cnr.it

$d \times d$ identity matrix $\mathbf{I}_d$ (ii) by the Multivariate Central Limit Theorem the renormalized Wishart matrix $\sqrt{n}(\mathbf{W}_{n,d} - \mathbf{I}_d)$ converges in distribution to a random vector distributed according to a centered d-dimensional Gaussian law, provided that the fourth moment of F_{11} is finite.

For many years, the framework where d is fixed and n goes to infinity was sufficient for applications. Nowadays, due to the increasing need of handling large data sets, the high-dimensional regime, where both n and d are large simultaneously, is instead the usual context, see e.g. [14]. In the high-dimensional setting, the Law of Large Numbers and the Central Limit Theorem cannot be applied anymore, and to study the asymptotic behavior of the Wishart matrix, one has to employ different techniques. It is well-known that if $n, d \to \infty$ in such a way that $\frac{d}{n} \to c \in (0, \infty)$, then the empirical measure of the eigenvalues of $\mathbf{W}_{n,d}$ converges weakly to the Marchenko-Pastur law [19]. This central result has been refined in various directions; e.g. the independence assumption among the entries of $\mathbf{F}_{n,d}$ has been relaxed in a number of ways, see the important contributions [1, 29, 45], among many others.

However, in some applications n and d do not have the same asymptotic order, and therefore the Marchenko-Pastur Theorem is not appropriate to describe the asymptotics of the corresponding large Wishart matrix. The limiting distribution for families of eigenvalue statistics of general sample covariance matrices in the regimes $d/n \to \infty$ and $d/n \to 0$ has been investigated in the interesting articles [35] and [2], respectively.

Recently, many authors took another approach to analyze large Wishart matrices. Let $\mathbf{Z}_d = (Z_{ij})_{1 \le i \le d,\, 1 \le j \le d}$ be the GOE (Gaussian Orthogonal Ensemble) random matrix, i.e., the random variables Z_{ii} are distributed according to the Gaussian law with mean 0 and variance 2, the random variables Z_{ij}, $i < j$, are distributed according to the standard Normal law, $Z_{ij} = Z_{ji}$ for $j > i$, and the random variables $\{Z_{ij} :\ i \le j\}$ are independent. It has been independently proved in [6, 13] that, if the random variables F_{ij} are i.i.d. standard Gaussian, then the total variation distance between the law of the renormalized Wishart matrix $\sqrt{n}(\mathbf{W}_{n,d} - \mathbf{I}_d)$ and the law of $\mathbf{Z}_d$ tends to zero when $\frac{d^3}{n} \to 0$, as $n, d \to \infty$. In [6] it is also proved that if $\frac{n}{d^3} \to 0$, as $n, d \to \infty$, then the total variation distance between the laws of $\sqrt{n}(\mathbf{W}_{n,d} - \mathbf{I}_d)$ and $\mathbf{Z}_d$ tends to 1, in this sense the condition $d^3 = o(n)$ is sharp. In [36] it is proved that if $\frac{n}{d^3} \to c \in (0, \infty)$, as $n, d \to \infty$, then the total variation distance between the laws of $\sqrt{n}(\mathbf{W}_{n,d} - \mathbf{I}_d)$ and $\mathbf{Z}_d$ tends to $\mathrm{Erf}\left(\frac{1}{4\sqrt{3c}}\right)$, as $n, d \to \infty$. Here $\mathrm{Erf}(\cdot)$ denotes the Gauss error function; note that, as either $c \to \infty$ or $c \to 0$ one has that either the total variation distance between the laws of $\sqrt{n}(\mathbf{W}_{n,d} - \mathbf{I}_d)$ and $\mathbf{Z}_d$ tends to 0 or to 1, meaning that the "phase transition" from Wishart to GOE is smooth. The achievements in [6] have been generalized in [7] to the case where the entries F_{ij} are i.i.d. from a log-concave distribution.

The statistical inference based on observations belonging to Wiener chaoses of arbitrary order has been intensively studied, see e.g. [8, 9, 32, 41] among other contributions. This has stimulated mathematical research on Wishart matrices associated to random matrices with chaotic entries. Let $\mathfrak{H}$ be a real separable Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$, and let $\{e_{ij}\}_{i,j \in \mathbb{N}} \subset \mathfrak{H}$ be a family such that $\langle e_{ij}, e_{i'j'} \rangle_{\mathfrak{H}} = \mathbf{1}_{\{i=i'\}} s(j - j')$, where $s : \mathbb{Z} \to \mathbb{R}$, $\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$, is a correlation function satisfying $s(0) = 1$. Suppose that the entries of $\mathbf{F}_{n,d}$ are of the form $F_{ij} = I_1^X(e_{ij})$, where X is an isonormal Gaussian process over $\mathfrak{H}$ and $I_q^X(\cdot)$, $q \in \mathbb{N}$, denotes an element of the q-th Wiener chaos of X (we refer the reader to the Appendix A for any unexplained notions and notation). Letting $\mathbf{W}_{n,d}$ be the Wishart matrix associated to the chaotic random matrix $\mathbf{F}_{n,d}$, it is proved in [25] that the Wasserstein distance between the law of $\sqrt{n}(\mathbf{W}_{n,d} - \mathbf{I}_d)$ and the law of $\mathbf{G}_{n,d}$ is bounded above by $\sqrt{\frac{192 d^3}{n} \left(\sum_{k \in \mathbb{Z}:\ |k| \le n} |s(k)|^{4/3} \right)^3}$. Here $\mathbf{G}_{n,d} = (G_{ij})_{1 \le i \le d,\, 1 \le j \le d}$ is a symmetric random matrix, whose associated random vector $(G_{11},\ldots,G_{1d},\ldots,G_{d1},\ldots,G_{dd})$ is distributed according to the d^2-dimensional Gaussian law with mean $\mathbf{0}$ and the same covariance matrix as the random vector associated to the renormalized Wishart matrix $\sqrt{n}(\mathbf{W}_{n,d}-\mathbf{I}_d)$. Therefore if $s\in\ell^{4/3}(\mathbb{Z})$, then the Wasserstein distance is $O\left(\sqrt{\frac{d^3}{n}}\right)$, and so, similarly to [6, 13], one has that the renormalized Wishart matrix is close to $\mathbf{G}_{n,d}$ if $\frac{d^3}{n}\to 0$, as $n,d\to\infty$. Note that, in the setting of [25], the random variables F_{ij} are standard Gaussian, but only the rows $(\mathbf{F}_{n,d})_i$, $i=1,\ldots,n$, are i.i.d. This means that the context of [6, 13] has been extended to a situation where the entries of $\mathbf{F}_{n,d}$ are Gaussian and partially correlated. The row independence has been relaxed in other results of [25], and in [24]. Another important achievement of [25] concerns the study of the asymptotic behavior of Wishart matrices associated to the random matrix $\mathbf{F}_{n,d}$ with entries $F_{ij}:=I_1^X(e_{ij})$, where the family $\{e_{ij}\}_{i,j\in\mathbb{N}}\subset\mathfrak{H}$ is such that $\langle e_{ij},e_{i'j'}\rangle_{\mathfrak{H}}=\mathbf{1}_{\{i=i'\}}s_H(j-j')$, being $s_H(k)$ the correlation function of the fractional Brownian motion with Hurst parameter $H\in(0,1)$. In particular, for each value of $H\in(0,\frac{3}{4}]$, it is estimated the Wasserstein distance between the law of $\sqrt{n}(\mathbf{W}_{n,d}-\mathbf{I}_d)$ and the law of the corresponding random matrix $\mathbf{G}_{n,d}$. Interestingly, when $H\in(3/4,1)$ a new phenomenon appears: the Wishart matrix associated with $\mathbf{F}_{n,d}$, properly renormalized, converges to the so-called Rosenblatt-Wishart random matrix $\mathbf{R}_d=(R_{ij})_{1\le i\le d,\,1\le j\le d}$, a new random matrix model, introduced in [25], whose entries are suitable elements of the second Wiener chaos of X and whose diagonal elements R_{ii} are all independent from each other and have the Rosenblatt distribution.

The fluctuations of Wishart matrices associated to random matrices with non-Gaussian entries are investigated in [4]. Let $\mathbf{F}_{n,d}$ be the $n\times d$ random matrix with entries $F_{ij}:=I_{q_i}^X(f_{ij})$, $f_{ij}\in\mathfrak{H}^{\odot q_i}$, where $q_i\in\mathbb{N}$ are such that $\max_{1\le i\le n}q_i\le Q$, for some $Q\in\mathbb{N}$ not depending on n, and assume that the random variables F_{ij} are independent (but not necessarily identically distributed) and that the kernels are chosen in such a way that $q_i\|f_{ij}\|^2_{\mathfrak{H}^{\otimes q_i}}=1$ and $\mathbb{E}F_{ij}^2=3$. Let $\mathbf{W}_{n,d}$ be the Wishart matrix associated to $\mathbf{F}_{n,d}$. It is proved in [4] that the Wasserstein distance between the law of the renormalized Wishart matrix $\sqrt{n}(\mathbf{W}_{n,d}-\mathbf{I}_d)$ and the law of the GOE matrix $\mathbf{Z}_d$ is $O\left(\sqrt{\frac{d^3}{n}}\right)$. Another direction of research considered in [4] is to study the asymptotic behavior of the suitable renormalized Wishart matrix associated to a random matrix $\mathbf{F}_{n,d}=(F_{ij})_{1\le i\le n,\,1\le j\le d}$ whose elements are partially correlated and belong to the second Wiener chaos. More precisely, $F_{ij}:=R_j^{(i)}-R_{j-1}^{(i)}$, where $R^{(i)}=\{R_j^{(i)}\}_{0\le j\le d}$, $1\le i\le n$, are n independent Rosenblatt processes with a same Hurst parameter $H\in(1/2,1)$. In this case the suitably renormalized Wishart matrix converges in law, as $n\to\infty$, to a diagonal matrix whose diagonal elements are random variables distributed according to the Rosenblatt law; the authors are also able to quantify the Wasserstein distance associated to this limit.

In the recent contribution [5] are investigated the high-dimensional asymptotics of a properly renormalized Wishart matrix associated to a Gaussian matrix with correlated and non-stationary entries. Two distinct regimes emerge. In the first one, the law of the properly renormalized Wishart matrix becomes close in Wasserstein distance to the law of the GOE random matrix. In the second one, the law of the properly renormalized Wishart matrix becomes close in Wasserstein distance to the law of the Rosenblatt-Wishart random matrix.

The proofs in [4, 5, 24, 25] employ the Stein-Malliavin calculus, a new technique, introduced to quantify Nualart and Peccati's Fourth Moment Theorem [27], that has been extensively developed by the authors of [20, 21], as well as other researchers, see e.g. [10, 16, 18, 22, 23, 30, 33, 34, 37, 38, 39, 40]. In this paper we use this technique, combined with some ideas borrowed from high-dimensional probability [42, 43], to address a question that, to the best of our knowledge, has not been studied so far in the context of random matrices with chaotic entries. Let $\mathbf{F}_{n,d}$ be an $n \times d$ random matrix whose rows $(\mathbf{F}_{n,d})_i$, $1 \leq i \leq n$, are independent, centered, sub-Gaussian and isotropic random vectors in $\mathbb{R}^d$, and denote by $\mathbf{W}_{n,d}$ the associated Wishart matrix. It is well-known that, for every $t \geq 0$,

$$\mathbb{P}\left(\|\mathbf{W}_{n,d} - \mathbf{I}_d\|_{op} \leq (\max_{1\leq i\leq n} \|(\mathbf{F}_{n,d})_i\|_{\psi_2})^2 \max\left\{C\left(\sqrt{\frac{d}{n}} + \frac{t}{\sqrt{n}}\right), C^2\left(\sqrt{\frac{d}{n}} + \frac{t}{\sqrt{n}}\right)^2\right\}\right) \geq 1 - 2\mathrm{e}^{-t^2}$$

for a positive constant $C > 0$ not depending on n and d (see e.g. Theorem 4.6.1 p. 91 in [42]). Here $\|\cdot\|_{op}$ denotes the operator norm of a square matrix (see Section 4 for a formal definition) and $\|\cdot\|_{\psi_2}$ denotes the sub-Gaussian norm of a sub-Gaussian random vector (see Definitions 3.4.1 and 2.5.6 in [42]). Let $\varepsilon > 0$ be arbitrarily fixed, take $d = d_n = o(n)$, $t = t_n = o(\sqrt{n})$, as $n \to \infty$, and suppose that there exists a positive constant $K > 0$ (not depending on n) such that $\max_{1\leq i\leq n} \|(\mathbf{F}_n)_i\|_{\psi_2} \leq K$, where $\mathbf{F}_n := \mathbf{F}_{n,d_n}$ (in particular this condition is satisfied if e.g. the entries of $\mathbf{F}_n$ are centered Gaussian with variances bounded above by a same absolute constant or the entries of $\mathbf{F}_n$ are bounded random variables whose modulus is bounded above by a same absolute constant). Then there exists n_ε such that

$$\mathbb{P}\left(\|\mathbf{W}_n - \mathbf{I}_{d_n}\|_{op} \leq \varepsilon\right) \geq 1 - 2\mathrm{e}^{-t_n^2}, \quad \forall\, n > n_\varepsilon \tag{1.1}$$

where $\mathbf{W}_n := \mathbf{W}_{n,d_n}$.

In this paper we prove a relation similar to (1.1) in the high-dimensional regime when $d_n = o(n)$, as $n \to \infty$, and the rows $(\mathbf{F}_n)_i$, $1 \leq i \leq n$, are still independent, but their components are in the q_i-th, $q_i \in \mathbb{N}$, Wiener chaos of an isonormal Gaussian process X_i. Note that for matrices with chaotic entries of order bigger than or equal to 1 the conditions which guarantee (1.1), mentioned above, are in general not satisfied. As a by-product, we provide sufficient conditions which ensure that the "tall" random matrices $\mathbf{F}_n/\sqrt{n}$, with $\mathbf{F}_n$ having suitable chaotic entries, are approximate isometries. We apply the main result to a covariance estimation problem, which is of potential interest in data analysis. We leave to future research not only the study of the regimes where either d_n is asymptotically proportional to n or $d_n/n \to \infty$, as $n \to \infty$, but also the investigation of analogous results when the elements of the random matrix $\mathbf{F}_n$ belong to chaoses on Poisson and Rademacher spaces (see [16, 30, 18, 22, 34]).

The article is organized as follows. The main results concerning approximate isometries and covariance estimation are stated in Section 2, see Theorem 2.2 (approximate isometries) and Theorem 2.5 (covariance estimation). Their proofs exploit a basic inequality, which is also stated in Section 2, see Theorem 2.1. As already mentioned, one of the tools exploited in the proofs is the Malliavin-Stein method. In Section 3 we use this approach to provide a bound on the Kolmogorov distance between the Chi-Square law with n degrees of freedom and the law of the sum of the squares of n independent and centered random variables taking values on real Hilbert spaces and satisfying an integration by parts formula, see Corollary 3.4. Although results concerning the Chi-Square (or more generally, the Gamma) approximation already exist (see e.g. [20]), we found them not appropriate for our aims. For this reason, we provided new approximation bounds on the Chi-Square law which are both of independent interest and perfectly tailored to our scopes. In Section 4, we combined the estimates given in Section 3 with typical arguments of high-dimensional probability, such as approximation by δ-nets, to prove an abstract lower bound on the probability $\mathbb{P}(\|\mathbf{W}_{n,d} - \mathbf{I}_d\|_{op} \leq \varepsilon)$, for any $\varepsilon > 0$, see Theorem 4.4. Here $\mathbf{W}_{n,d}$ is the Wishart matrix associated to a random matrix $\mathbf{F}_{n,d}$, whose rows

$(\mathbf{F}_{n,d})_i$, $i=1,\dots,n$, are independent, centered and such that certain linear combinations of their elements satisfy an integration by parts formula. Section 5 contains an intermediate step towards the main results. Indeed, applying Theorem 4.4. we provide an abstract lower bound on the probability $\mathbb{P}(\|\mathbf{W}_{n,d}-\mathbf{I}_d\|_{op}\leq\varepsilon)$, $\varepsilon>0$, when the entries of $\mathbf{F}_{n,d}$ are chaotic. In Section 6, we prove the above mentioned basic inequality of Theorem 2.1. In Section 7, we prove Theorem 2.2., i.e., we prove that, for any $\varepsilon>0$ (small enough), the event $\{\|\mathbf{W}_n-\mathbf{I}_{d_n}\|_{op}\leq\varepsilon\}$ occurs with high probability, and indeed we are able to give an explicit lower bound on the quantity $\mathbb{P}(\|\mathbf{W}_n-\mathbf{I}_{d_n}\|_{op}\leq\varepsilon)$, for all n large enough. Here $\mathbf{W}_n$ is the Wishart matrix associated to a "tall" random matrix $\mathbf{F}_n:=\mathbf{F}_{n,d_n}$ with $d_n=o(n)$, as $n\to\infty$, and with independent rows $(\mathbf{F}_n)_i$ whose components belong to the q_i-th Wiener chaos of an isonormal Gaussian process X_i. In Section 8, we prove Theorem 2.5. Roughly speaking, we address the following covariance estimation problem. Let $\mathbf{I}_X(d)$ be a d-dimensional random vector whose components belong to the q-th Wiener chaos of an isonormal Gaussian process X. Let $\mathbf{\Sigma}_d$ be the covariance matrix of $\mathbf{I}_X(d)$ and let $\widehat{\mathbf{\Sigma}}_{n,d}$ be the sample covariance matrix based on n data (i.e., n independent samples from $\mathbf{I}_X(d)$). We prove that, for any $\varepsilon>0$ (small enough), the event $\{R_n\leq\varepsilon\}$ occurs with high probability, as n grows large, and indeed we are able to give an explicit lower bound on the quantity $\mathbb{P}(R_n\leq\varepsilon)$, for all n large enough. Here R_n is the relative error concerning the approximation of $\mathbf{\Sigma}_{d_n}$ by $\widehat{\mathbf{\Sigma}}_{n,d_n}$ (in the operator norm) and $d_n=o(n)$, as $n\to\infty$. Finally, in Section 9, we apply Theorem 2.2 to random matrices with independent rows and entries given by suitable stochastic integrals of the Ornstein-Uhlenbeck process (see Theorem 9.1), and we apply Theorem 2.5 to random vectors with entries given by suitable stochastic integrals of the Ornstein-Uhlenbeck process (see Theorem 9.4). We include the Appendices A, B, C, D and E, where we give some preliminaries on the Malliavin calculus on the Wiener space, and we prove some ancillary results.

2 Statement of the main results

In this section we state the main theoretical results of the paper. Throughout this section, $\mathfrak{H}_1,\dots,\mathfrak{H}_n$ are real separable Hilbert spaces and $X_1,\dots,X_n$ are independent isonormal Gaussian processes, X_i over $\mathfrak{H}_i$, $i=1,\dots,n$, $n\in\mathbb{N}$. As already mentioned in the introduction, $I_q^{X_i}(\cdot)$ denotes an element of the q-th Wiener chaos, $q\in\mathbb{N}$, of the isonormal Gaussian process X_i. Here again, we refer the reader to the Appendix A for any unexplained notions and notation.

2.1 A basic relation

Herafter, we denote by $s_{\min}(\mathbf{A})$ (respectively, $s_{\max}(\mathbf{A})$) the smallest (respectively, the biggest) singular value of a rectangular matrix $\mathbf{A}$ (see Section 4 for a formal definition), by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^p$, for some $p\in\mathbb{N}$, and by $f\otimes_r g$, $r\in\mathbb{N}$, the contraction of f and g of order r (see the Appendix A).

Theorem 2.1. Let $n\in\mathbb{N}\setminus\{1,2\}$, $d,q_i\in\mathbb{N}$, $i=1,\dots,n$, and let $\mathbf{F}_{n,d}$ denote the $n\times d$ random matrix $\mathbf{F}_{n,d}:=(I_{q_i}^{X_i}(f_{ij}))_{1\leq i\leq n,\,1\leq j\leq d}$, where $f_{ij}\in\mathfrak{H}_i^{\odot q_i}$, $i=1,\dots,n$, $j=1,\dots,d$. Let $\varphi_{1,n,d}$, $\varphi_{2,n,d}$ and $\varphi_{3,n,d}$ be three non-negative quantities such that:

$$\sup_{1\leq i\leq n}\sup_{1\leq j\leq d}|1-(q_i!)\|f_{ij}\|^2_{\mathfrak{H}_i^{\otimes q_i}}|\leq\varphi_{1,n,d},\tag{2.1}$$

$$\text{If } d\geq 2, \text{ then } \sup_{1\leq i\leq n}\sup_{1\leq j\neq k\leq d}|\langle f_{ij},f_{ik}\rangle_{\mathfrak{H}_i^{\otimes q_i}}|\leq\varphi_{2,n,d},\tag{2.2}$$

$$\text{If } \{i:\,q_i\geq 2\}\neq\emptyset, \text{ then } \sup_{1\leq i\leq n:\,q_i\geq 2}\sup_{1\leq j\leq d}\sup_{1\leq r\leq q_i-1}\|f_{ij}\otimes_r f_{ij}\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}}\leq\varphi_{3,n,d}.\tag{2.3}$$

We set $\varphi_{2,n,d}:=0$ if $d=1$ and $\varphi_{3,n,d}:=0$ if $q_i=1$ for any i.

Assume that there exists $Q \in \mathbb{N}$ (*not depending on* n) *such that*

$$\max_{1\le i\le n} q_i \le Q. \tag{2.4}$$

Then, for every $\varepsilon \in (0,1)$, *we have*

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \le s_{\min}\left(\frac{\mathbf{F}_{n,d}}{\sqrt{n}}\right) \le s_{\max}\left(\frac{\mathbf{F}_{n,d}}{\sqrt{n}}\right) \le 1+\varepsilon\right)\\
&\quad \ge \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \le \left\|\frac{1}{\sqrt{n}}\mathbf{F}_{n,d}\mathbf{u}\right\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \textit{for all } \mathbf{u}\in\mathbb{R}^d\right)\\
&\quad \ge \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top \mathbf{F}_{n,d} - \mathbf{I}_d\right\|_{op} \le \varepsilon\right)\\
&\quad \ge 1 - 5\inf_{\delta\in(0,1/2)}\left(\frac{2}{\delta}+1\right)^d\left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_{n,d}} + \Phi_{n,d}\right),
\end{aligned} \tag{2.5}$$

where

$$\Phi_{n,d} := \kappa_Q n d \max\{\varphi_{1,n,d}, (d-1)\varphi_{2,n,d}, \varphi_{3,n,d} + (d-1)\varphi_{3,n,d}^2\}$$

and

$$\kappa_Q := 1 + Q! + Q(Q-1)[(Q-1)!]^2\sqrt{(2Q-2)!}. \tag{2.6}$$

2.2 Approximate isometry

Roughly speaking, the next theorem shows that if the kernels f_{ij} depend on n, i.e., $f_{ij} = f_{ij}^{(n)}$, and are properly chosen, and $d = d_n = o(n)$, as $n \to \infty$, then the lower bound (2.5) can be made explicit and tends to 1, as $n \to \infty$.

Hereafter, we shall consider the polynomial of third degree

$$f(\delta) := \delta(2+\delta)(1-2\delta), \quad \delta \in (0, 1/2).$$

Note that $f(\cdot)$ is concave, continuous, equal to zero in 0 and $1/2$, strictly increasing on $(0,\delta^*)$, $\delta^* := \frac{1}{6}(\sqrt{21}-3)$, strictly decreasing on $(\delta^*, 1/2)$ and

$$f(\delta^*) = \frac{7\sqrt{7/3}}{6} - \frac{3}{2} > 0. \tag{2.7}$$

Theorem 2.2. *Let* $n \in \mathbb{N}\setminus\{1,2\}$, $q_i, d_n \in \mathbb{N}$, $i = 1,\dots,n$, *and let* $\mathbf{F}_n$ *denote the* $n\times d_n$ *random matrix* $\mathbf{F}_n := (I_{q_i}^{X_i}(f_{ij}^{(n)}))_{1\le i\le n,\,1\le j\le d_n}$, *where* $f_{ij}^{(n)} \in \mathfrak{H}_i^{\circ q_i}$, $i = 1,\dots,n$, $j = 1,\dots,d_n$.

Assume (2.4) *and* $d_n = o(n)$, *as* $n\to\infty$, *let*

$$0 < \varepsilon < \frac{f(\delta^*)}{4\sqrt{\kappa_Q}\left(\mathrm{e}^{-3/2}\sqrt{12\kappa_Q} + \sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right)} < 1$$

be arbitrarily fixed and let $\{\varphi_{h,n}\}_n$, $h = 1,2,3$, *be three sequences of non-negative numbers that satisfy the following conditions* (i) *and* (ii):
(i) *There exists an index* $n'_\varepsilon \ge 2$ *such that*

$$d_n^3 \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\} \le \varepsilon^6 n^{\frac{3}{2}}\mathrm{e}^{-\frac{n\varepsilon^2}{4}}, \quad \forall\, n > n'_\varepsilon. \tag{2.8}$$

(ii) *There exists an index* $n'' \ge 2$ *such that, for any* $n > n''$, *the relations* (2.1), (2.2) *and* (2.3) *hold with* $f_{ij} = f_{ij}^{(n)}$, $d = d_n$ *and* $\varphi_{h,n,d_n} = \varphi_{h,n}$, $h = 1,2,3$.

Here κ_Q *is defined by* (2.6), $f(\delta^*)$ *is given by* (2.7) *and we set* $\varphi_{2,n} := 0$ *if* $d_n = 1$, *and* $\varphi_{3,n} := 0$ *if* $q_i = 1$ *for any* i.

Then, there exists $n_\varepsilon \geq 2$ *such that*

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \leq s_{\min}\left(\frac{\mathbf{F}_n}{\sqrt{n}}\right) \leq s_{\max}\left(\frac{\mathbf{F}_n}{\sqrt{n}}\right) \leq 1+\varepsilon\right) \\
&\quad \geq \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \leq \left\|\frac{1}{\sqrt{n}}\mathbf{F}_n\mathbf{u}\right\| \leq (1+\varepsilon)\|\mathbf{u}\|, \quad \textit{for all } \mathbf{u}\in\mathbb{R}^{d_n}\right) \\
&\quad \geq \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_n)^\top \mathbf{F}_n - \mathbf{I}_{d_n}\right\|_{op} \leq \varepsilon\right) \\
&\quad \geq 1 - 5\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)^{d_n}\left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2} + c_\varepsilon \frac{n}{d_n}\mathrm{e}^{-n\varepsilon^2/8}\right), \quad \forall\, n > n_\varepsilon. \qquad (2.9)
\end{aligned}$$

Here

$$\delta_1(n,\varepsilon) := 2\sqrt{\frac{7}{12}}\cos\left(\frac{5\pi}{3} - \frac{1}{3}\arctan\left(2\frac{\sqrt{\frac{343}{1728} - \frac{1}{4}\left(\frac{3}{4} + \frac{2(\frac{d_n}{n}+c_\varepsilon)}{\varepsilon^2}\right)^2}}{\frac{3}{4} + \frac{2(\frac{d_n}{n}+c_\varepsilon)}{\varepsilon^2}}\right)\right) \in (0,1/2) \qquad (2.10)$$

and

$$c_\varepsilon := \varepsilon^3\sqrt{\kappa_Q}\left(\mathrm{e}^{-3/2}\sqrt{12\kappa_Q} + \sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right). \qquad (2.11)$$

It turns out that

$$\begin{aligned}
&\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)^{d_n}\left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2} + c_\varepsilon\frac{n}{d_n}\mathrm{e}^{-n\varepsilon^2/8}\right) \\
&\qquad = \mathrm{e}^{-\frac{n\varepsilon^2}{8}(1-2\overline{\delta}_\varepsilon)^2(1+o(1))} + c_\varepsilon\frac{n}{d_n}\mathrm{e}^{-\frac{n\varepsilon^2}{8}(1+o(1))}, \quad \textit{as } n\to\infty \qquad (2.12)
\end{aligned}$$

where

$$\overline{\delta}_\varepsilon := 2\sqrt{\frac{7}{12}}\cos\left(\frac{5\pi}{3} - \frac{1}{3}\arctan\left(2\frac{\sqrt{\frac{343}{1728} - \frac{1}{4}\left(\frac{3}{4} + \frac{2c_\varepsilon}{\varepsilon^2}\right)^2}}{\frac{3}{4} + \frac{2c_\varepsilon}{\varepsilon^2}}\right)\right) \in (0,1/2).$$

Therefore the lower bound (2.9) *tends to* 1 *as* $n\to\infty$.

Remark 2.3. It is worthwhile to emphasize that the “smallness” condition (2.8) may imply (e.g. in some specific examples, see Remark 9.2) that $d_n^3 = o(n)$, as $n\to\infty$. As discussed in the Introduction, this is exactly the regime under which various statistical distances between renormalized Wishart matrices, based on suitable classes of chaotic random variables, and random matrices belonging to suitable Gaussian Ensembles, tend to zero, as $n\to\infty$, see [4, 6, 13, 24, 25].

Remark 2.4. Theorem 2.2 provides sufficient conditions which guarantee that, for n large enough, the random matrix $\mathbf{F}_n/\sqrt{n}$ is an approximate isometry with high probability. Note that the assumption (2.8) is satisfied if

$$d_n^3\max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\} = o\left(n^{\frac{3}{2}}\mathrm{e}^{-\frac{n\varepsilon^2}{4}}\right), \quad \text{as } n\to+\infty.$$

A close inspection of the proof of Theorem 2.2 gives the theoretical expression of the index n_ε in (2.9), see Remark 7.2. In specific cases, the expression of n_ε can be explicitly computed: see Remark 9.3 for a concrete application to random matrices whose entries are (suitable) stochastic integrals of Ornstein-Uhlenbeck processes.

2.3 Covariance estimation

Suppose that we are analyzing some data which are represented as n points $\mathbf{D}_{1,d}, \dots, \mathbf{D}_{n,d}$, $n \in \mathbb{N}$, of dimension $d \in \mathbb{N}$ sampled independently from an unknown vector of the form

$$\mathbf{I}_X(d) := (I_q^X(f_1), \dots, I_q^X(f_d)),$$

for some $q \in \mathbb{N}$, some isonormal Gaussian process X over a real separable Hilbert space $\mathfrak{H}$ and some kernels $f_j \in \mathfrak{H}^{\circ q}$, $j = 1, \dots, d$. One of the most popular data exploration tool is the Principal Component Analysis, for which, when we do not have access to the full distribution of $\mathbf{I}_X(d)$, covariance estimation is a fundamental step.

Let $\mathbf{\Sigma}_d := \mathbb{E}\mathbf{I}_X(d)^\top \mathbf{I}_X(d)$ be the covariance matrix of $\mathbf{I}_X(d)$. By the isometry formula (A.1) we have

$$\mathbf{\Sigma}_d = (q!\langle f_j, f_k\rangle_{\mathfrak{H}^{\circ q}})_{1 \le j,k \le d}.$$

Thus estimating the covariance of $\mathbf{I}_X(d)$ means estimating the quantities $q!\langle f_j, f_k\rangle_{\mathfrak{H}^{\circ q}}$, $j, k = 1, \dots, d$. The standard estimator of $\mathbf{\Sigma}_d$ is the sample covariance matrix

$$\widehat{\mathbf{\Sigma}}_{n,d} := \frac{1}{n}\sum_{i=1}^{n} \mathbf{D}_{i,d}^\top \mathbf{D}_{i,d}.$$

Note that $\widehat{\mathbf{\Sigma}}_{n,d}$ is unbiased, i.e., $\mathbb{E}\widehat{\mathbf{\Sigma}}_{n,d} = \mathbf{\Sigma}_d$, and, for $d \in \mathbb{N}$ fixed, consistent, indeed, by the Strong Law of Large Numbers

$$\widehat{\mathbf{\Sigma}}_{n,d} \to \mathbf{\Sigma}_d \quad \text{almost surely, as } n \to \infty.$$

If $d = d_n$ depends on n, then consistency is no more guaranteed by the Strong Law of Large Numbers, and a natural question to address is the following: for which values of n, does the sample covariance $\widehat{\mathbf{\Sigma}}_{n,d_n}$ accurately estimate $\mathbf{\Sigma}_{d_n}$?

Let $f_i^{(n)} \in \mathfrak{H}^{\circ q}$, $i = 1, \dots, d_n$, $n \in \mathbb{N} \setminus \{1, 2\}$, $q, d_n \in \mathbb{N}$, be kernels. We denote by $\mathbf{\Sigma}^{(n)}$ the covariance matrix of $\mathbf{I}_X^{(n)}(d_n) := (I_q^X(f_1^{(n)}), \dots, I_q^X(f_{d_n}^{(n)}))$, i.e.,

$$\mathbf{\Sigma}^{(n)} := (q!\langle f_j^{(n)}, f_k^{(n)}\rangle_{\mathfrak{H}^{\circ q}})_{1 \le j,k \le d_n} \tag{2.13}$$

and by $\widehat{\mathbf{\Sigma}}^{(n)}$ the sample covariance of $\mathbf{\Sigma}^{(n)}$, i.e.,

$$\widehat{\mathbf{\Sigma}}^{(n)} := \frac{1}{n}\sum_{i=1}^{n} (\mathbf{D}_i^{(n)})^\top \mathbf{D}_i^{(n)},$$

being $\mathbf{D}_1^{(n)}, \dots, \mathbf{D}_n^{(n)}$ n independent samples from $\mathbf{I}_X^{(n)}(d_n)$. We denote by

$$R_n := \frac{\|\widehat{\mathbf{\Sigma}}^{(n)} - \mathbf{\Sigma}^{(n)}\|_{op}}{\|\mathbf{\Sigma}^{(n)}\|_{op}} \tag{2.14}$$

the relative error concerning the approximation of $\mathbf{\Sigma}^{(n)}$ by the sample covariance $\widehat{\mathbf{\Sigma}}^{(n)}$.

The next theorem shows that if the kernels $f_j^{(n)}$ are properly chosen and $d_n = o(n)$, as $n \to \infty$, then, for any $\varepsilon > 0$, the event $\{R_n \le \varepsilon\}$ occurs with high probability, as $n \to \infty$.

Theorem 2.5. *Let $n \in \mathbb{N} \setminus \{1, 2\}$, $q, d_n \in \mathbb{N}$ and let R_n be defined by* (2.14).

Assume $d_n = o(n)$, as $n \to \infty$, let

$$0 < \varepsilon < \frac{f(\delta^*)}{4\sqrt{\kappa_q}\left(\mathrm{e}^{-3/2}\sqrt{12\kappa_q} + \sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right)} < 1$$

be arbitrarily fixed, and let $\{\widetilde{\varphi}_{h,n}\}_n$*,* $h = 1, 2, 3$*, be three sequences of non-negative numbers that satisfy the following conditions* (i) *and* (ii)*:*
(i) *There exists an index* $n'_\varepsilon \geq 2$ *such that*

$$d_n^{10} \max\{\widetilde{\varphi}_{1,n}, \widetilde{\varphi}_{2,n}, \widetilde{\varphi}_{3,n}\} \leq \frac{1}{C_1(q)} \varepsilon^6 n^{3/2} \mathrm{e}^{-\frac{n\varepsilon^2}{4}}, \quad \forall\, n > n'_\varepsilon. \tag{2.15}$$

(ii) *There exists an index* $n'' \geq 2$ *such that* $\forall\, n > n''$

$$\sup_{1 \leq j \leq d_n} |1 - q! \|f_j^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}| \leq \widetilde{\varphi}_{1,n}, \tag{2.16}$$

$$\textit{If } d_n \geq 2\textit{, then } \sup_{1 \leq j \neq k \leq d_n} |\langle f_j^{(n)}, f_k^{(n)} \rangle_{\mathfrak{H}^{\otimes q}}| \leq \widetilde{\varphi}_{2,n}, \tag{2.17}$$

$$\textit{If } q \geq 2\textit{, then } \sup_{1 \leq j \leq d_n} \sup_{1 \leq r \leq q-1} \|f_j^{(n)} \otimes_r f_j^{(n)}\|_{\mathfrak{H}^{\otimes 2(q-r)}} \leq \widetilde{\varphi}_{3,n}. \tag{2.18}$$

Furthermore, assume that there exists and index $n''' \geq 2$ *such that the covariance matrices* $\mathbf{\Sigma}^{(n)}$ *are invertible for every* $n > n'''$*.*

Here κ_q *is defined by* (2.6) *with* q *in place of* Q*,* $f(\delta^*)$ *is given by* (2.7)*,*

$$C_1(q) := \frac{67}{4} + \frac{55}{4} q! + \frac{63}{2} (q!)^2 + 10 (q!)^3, \tag{2.19}$$

and we set $\widetilde{\varphi}_{2,n} := 0$ *if* $d_n = 1$*, and* $\widetilde{\varphi}_{3,n} := 0$ *if* $q = 1$*.*

Then there exists $n_\varepsilon \geq 2$ *such that*

$$\mathbb{P}(R_n \leq \varepsilon) \geq 1 - 5 \left(\frac{2}{\delta'_1(n, \varepsilon)} + 1 \right)^{d_n} \left(\mathrm{e}^{-\frac{n}{8} \varepsilon^2 (1 - 2\delta'_1(n,\varepsilon))^2} + c'_\varepsilon \frac{n}{d_n} \mathrm{e}^{-n\varepsilon^2/8} \right), \quad \forall\, n > n_\varepsilon. \tag{2.20}$$

Here the quantities $\delta'_1(n, \varepsilon)$ *and* c'_ε *are defined, respectively, as* $\delta_1(n, \varepsilon)$ *in* (2.10) *and* c_ε *in* (2.11) *with* q *in place of* Q*. It turns out that relation* (2.12) *holds with* $\delta'_1(n, \varepsilon)$ *and* c'_ε *in place of* $\delta_1(n, \varepsilon)$ *and* c_ε*, respectively, and therefore the rightmost term in* (2.20) *tends to* 1*, as* $n \to \infty$*.*

Remark 2.6. Theorem 2.5 provides sufficient conditions which guarantee that, for n large enough, the relative error R_n is small with high probability. Note that the assumption (2.15) is satisfied if

$$d_n^{10} \max\{\widetilde{\varphi}_{1,n}, \widetilde{\varphi}_{2,n}, \widetilde{\varphi}_{3,n}\} = o\left(n^{3/2} \mathrm{e}^{-\frac{n\varepsilon^2}{4}} \right), \quad \text{as } n \to \infty.$$

Here again, a close inspection of the proof of Theorem 2.5 gives the theoretical expression of the index n_ε in (2.20), see Remark 8.2. In specific cases, the expression of n_ε can be explicitly computed: see Remark 9.5 for a concrete application to random vectors whose components are (suitable) stochastic integrals of Ornstein-Uhlenbeck processes.

2.4 The intuition behind the results

The following Proposition 2.7 is an immediate consequence of some findings in [20], [28] and [31] (see also [21] and [23]). It is proved in the Appendix B for completeness. Hereon, for $d \in \mathbb{N}$, we denote by δ_{jk}, $1 \leq j, k \leq d$, the Kronecker symbol.

Proposition 2.7. *Let* $d, q \in \mathbb{N}$ *and* X *an isonormal Gaussian process on a real separable Hilbert space* $\mathfrak{H}$*. For* $j = 1, \ldots, d$ *and* $m \in \mathbb{N}$*, let* $g_j^{(m)} \in \mathfrak{H}^{\odot q}$*. Assume that, for any* $j, k = 1, \ldots, d$*,*

$$q! \langle g_j^{(m)}, g_k^{(m)} \rangle_{\mathfrak{H}^{\otimes q}} \to \delta_{jk}, \quad \textit{as } m \to \infty. \tag{2.21}$$

Moreover, assume that if $q \geq 2$, then, for any $j = 1, \dots, d$ and any $r = 1, \dots, q-1$,

$$\|g_j^{(m)} \otimes_r g_j^{(m)}\|_{\mathfrak{H}^{\otimes 2(q-r)}} \to 0, \quad \text{as } m \to \infty. \tag{2.22}$$

Then $(I_q^X(g_1^{(m)}), \dots, I_q^X(g_d^{(m)}))$ converges in distribution to a random vector distributed according to the d-dimensional standard Gaussian law, as $m \to \infty$.

We can now explain the intuition behind our results. Let $m, n, d \in \mathbb{N}$, $q_i \in \mathbb{N}$, $i = 1, \dots, n$, and $\mathbf{F}_{m,n,d} = (I_{q_i}^{X_i}(f_{ij}^{(m)}))_{1 \leq i \leq n,\, 1 \leq j \leq d}$, where $f_{ij}^{(m)} \in \mathfrak{H}_i^{\circ q_i}$. Suppose that, for any $i = 1, \dots, n$:

$$\lim_{m\to\infty} q_i! \langle f_{ij}^{(m)}, f_{ik}^{(m)} \rangle_{\mathfrak{H}_i^{\otimes q_i}} = \delta_{jk}, \quad \forall\, j,k = 1, \dots, d$$

and

$$\text{If } q_i \geq 2, \text{ then } \lim_{m\to\infty} \|f_{ij}^{(m)} \otimes_r f_{ij}^{(m)}\|_{\mathfrak{H}_i^{\otimes 2(q_i - r)}} = 0, \quad \forall\, j = 1, \dots, d \text{ and } \forall\, r = 1, \dots, q_i - 1.$$

By Proposition 2.7 and the fact that the rows of $\mathbf{F}_{m,n,d}$ are independent (since the isonormal Gaussian processes $X_1, \dots, X_n$ are independent, as assumed at the beginning of Section 2), it follows that, as $m \to \infty$, $\mathbf{F}_{m,n,d}$ converges in law to the $n \times d$ random matrix $\mathbf{N}_{n,d}$ whose entries are independent and distributed according to the 1-dimensional standard Gaussian law. Clearly, relation (1.1) holds with $\mathbf{F}_n := \mathbf{N}_{n,d_n}$. Therefore, intuitively, under suitable (not obvious) assumptions, one may hope that a relation similar to (1.1) holds with $\mathbf{F}_n := (I_{q_i}^{X_i}(f_{ij}^{(n)}))_{1 \leq i \leq n,\, 1 \leq j \leq d_n}$.

3 Chi-Square approximation

The following characterization of the gamma distribution (see [17]) is the starting point of Stein's method for gamma approximation. The random variable X has the gamma distribution with parameters (r, λ), denoted by $\Gamma_{r,\lambda}$, $r, \lambda \in \mathbb{R}^+ := (0, \infty)$, if and only if

$$\mathbb{E}[Xf''(X) + (r - \lambda X)f'(X)] = 0,$$

for all twice differentiable functions $f : \mathbb{R}^+ \to \mathbb{R}$, which are such that the expectations $\mathbb{E}|Zf''(Z)|$, $\mathbb{E}|f'(Z)|$ and $\mathbb{E}|Zf'(Z)|$ are finite for $Z \sim \Gamma_{r,\lambda}$. Hereafter, we write $Y \sim \mathcal{L}$ if Y is a random element distributed according to the probability law $\mathcal{L}$.

Let $h : \mathbb{R}^+ \to \mathbb{R}$ be a measurable function such that $\mathbb{E}|h(Z)| < \infty$, and again $Z \sim \Gamma_{r,\lambda}$. It is straightforward to verify that

$$f'(x) := \frac{1}{xp(x)} \int_0^x (h(t) - \Gamma_{r,\lambda} h) p(t) \mathrm{d}t, \quad p(x) := \frac{\lambda^r}{\Gamma(r)} x^{r-1} \mathrm{e}^{-\lambda x}, \quad x \in \mathbb{R}^+ \tag{3.1}$$

solves the Stein equation

$$xf''(x) + (r - \lambda x)f'(x) = h(x) - \Gamma_{r,\lambda} h, \quad x \in \mathbb{R}^+. \tag{3.2}$$

Hereafter, given a probability law $\mathcal{L}$ and a measurable function h, we often denote $\mathbb{E}h(Y)$, where $Y \sim \mathcal{L}$, by $\mathcal{L}h$; we also denote by Γ the Euler gamma function.

The following lemma collects results from [11] and [17]. We set

$$\begin{aligned}\mathcal{C}_\lambda := \{&h : \mathbb{R}^+ \to \mathbb{R} : \ h \text{ is absolutely continuous and } \exists c > 0, a < \lambda \text{ such that} \\ &|h(x)| \leq c\mathrm{e}^{ax}, \forall x \in \mathbb{R}^+\}.\end{aligned}$$

We also let

$$\mathcal{M}_b := \{h : \mathbb{R}^+ \to \mathbb{R} \text{ measurable and bounded}\}$$

and $\|\cdot\|_\infty := \sup_{x>0} |\cdot|$.

Lemma 3.1. *Let f' be the solution* (3.1) *of the Stein equation* (3.2). *The following relations hold*:
(i) *If $h \in \mathcal{C}_\lambda$, then $\|f'\|_\infty \leq \frac{\|h'\|_\infty}{\lambda}$.*
(ii) *If $h \in \mathcal{M}_b$, then $\sup_{x>0} |xf''(x)| \leq 4\|h\|_\infty$.*

In this section, all the random variables are real-valued and defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout this paper we shall consider the set of functions $\mathcal{C}_b^1$ defined by

$$\mathcal{C}_b^1 := \{g : \mathbb{R} \to \mathbb{R} \text{ such that } g \text{ is continuously differentiable with } g' \text{ bounded}\}. \tag{3.3}$$

Furthermore, $\mathfrak{H}$ denotes a generic real Hilbert space, with inner product $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$ and norm $\langle y, y \rangle_{\mathfrak{H}}^{1/2} = \|y\|_{\mathfrak{H}}$, and the symbol $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathfrak{H})$ denotes the class of those $\mathfrak{H}$-valued random elements Y that are $\mathcal{F}$-measurable and such that $\mathbb{E}\|Y\|_{\mathfrak{H}}^2 < \infty$. If $\mathfrak{H} = \mathbb{R}$, we simply write $L^2(\Omega, \mathcal{F}, \mathbb{P})$ in place of $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$.

Let X, Y be two $\mathbb{R}^+$-valued random variables. Throughout this paper, we shall consider the distance $d_{\{1\}}$ and the Kolmogorov distance d_K between the laws of X and Y, which are defined as

$$d_{\{1\}}(X, Y) := \sup_{h \in \mathcal{H}_{\{1\}}} |\mathbb{E}h(X) - \mathbb{E}h(Y)|$$

and

$$d_K(X, Y) := \sup_{x \in \mathbb{R}^+} |\mathbb{P}(X \leq x) - \mathbb{P}(Y \leq x)|,$$

respectively. Here,

$$\mathcal{H}_{\{1\}} := \{h : \mathbb{R}^+ \to \mathbb{R} \text{ such that } h \text{ is absolutely continuous and } \|h\|_\infty, \|h'\|_\infty \leq 1\}.$$

Since Lipschitz continuous functions are absolutely continuous it follows that

$$d_{\{1\}} \geq d_{bW}, \tag{3.4}$$

where d_{bW} denotes the bounded Wasserstein distance (also known as Fortet-Mourier or Dudley metric), i.e.,

$$d_{bW}(X, Y) := \sup_{h \in \mathcal{H}_{bW}} |\mathbb{E}h(X) - \mathbb{E}h(Y)|,$$

where

$$\mathcal{H}_{bW} := \{h : \mathbb{R}^+ \to \mathbb{R} \text{ such that } h \text{ is Lipschitz and } \|h\|_\infty, \|h'\|_\infty \leq 1\}.$$

Note that both $d_{\{1\}}$ and d_K metrize the convergence in distribution.

Hereon, we denote by χ_n^2 the Chi-Square distribution with $n \in \mathbb{N}$ degrees of freedom, For a random variable X defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and with values on a measurable space $(E, \mathcal{E})$, we put $\mathbb{P}_X(B) := \mathbb{P} \circ X^{-1}(B)$, $B \in \mathcal{E}$.

3.1 Chi-Square approximation in the $d_{\{1\}}$-distance

The next theorem provides a general bound for the Chi-Square approximation in the $d_{\{1\}}$-distance.

Theorem 3.2. *For $n \in \mathbb{N}$, let $F_1, \ldots, F_n$ be independent and centered real-valued random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathfrak{H}_1, \ldots, \mathfrak{H}_n$ be real Hilbert spaces. For any $i = 1, \ldots, n$, we assume that there exist random elements $C^{(i)}(F_i), \Delta^{(i)}(F_i) \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{H}_i)$ such that*

$$\mathbb{E}F_i g(F_i) = \mathbb{E}g'(F_i) \langle C^{(i)}(F_i), \Delta^{(i)}(F_i) \rangle_{\mathcal{H}_i}, \quad g \in \mathcal{C}_b^1. \tag{3.5}$$

Then

$$d_{\{1\}}(S_n, Z_n) \leq 5 \sum_{i=1}^{n} \mathbb{E}|1 - \langle C^{(i)}(F_i), \Delta^{(i)}(F_i)\rangle_{\mathfrak{H}_i}|, \tag{3.6}$$

where $S_n := \sum_{i=1}^n F_i^2$ *and* $Z_n \sim \chi_n^2$*.*

Remark 3.3. We refer the reader to the Appendix A for any unexplained notion and notation in this remark. To orient the reader on the verification of the crucial assumption (3.5), we note that, if the random variables F_i in the statement of Theorem 3.2 lie in $\mathbb{D}^{1,2}$, then, by formula (A.5), we have that the integration by parts (3.5) holds with $C^{(i)}(F_i) = -DL^{-1}F_i$ and $\Delta^{(i)}(F_i) = DF_i$, where D is the Malliavin derivative operator and L^{-1} is the inverse of the Ornstein-Uhlenbeck operator. In particular, letting X denote an isonormal Gaussian process over a real separable Hilbert space $\mathfrak{H}$ and $I_q^X(f)$, $q \in \mathbb{N}$, $f \in \mathfrak{H}^{\odot q}$, an element of the q-th Wiener chaos, if $F_i = I_q^X(f)$, then, by formulas (A.6) and (A.7), the integration by parts (3.5) holds with $C^{(i)}(F_i) = I_{q-1}^X(f)$ and $\Delta^{(i)}(F_i) = qI_{q-1}^X(f)$.

Proof. (Theorem 3.2) Since $\Gamma_{n/2,1/2} = \chi_n^2$, we shall make use of the Stein equation (3.2). Let $h \in \mathcal{H}_{\{1\}}$ be arbitrarily fixed and let f_h' be the solution of the Stein equation (3.2) with $r = n/2$ and $\lambda = 1/2$. Setting $\mathbf{F}_{\hat{i}}^2 := (F_1^2, \ldots, F_{i-1}^2, F_{i+1}^2, \ldots, F_n^2)$, we have

$$\begin{aligned}
\mathbb{E}[h(S_n)] - \chi_n^2 h &= \mathbb{E}\left[S_n f_h''(S_n) + \frac{1}{2}(n - S_n) f_h'(S_n)\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\mathbb{E}\left[F_i^2 f_h''(S_n) + \frac{1}{2}(1 - F_i^2) f_h'(S_n) \;\middle|\; F_1^2, \ldots, F_{i-1}^2, F_{i+1}^2, \ldots, F_n^2\right] \\
&= \sum_{i=1}^{n} \int_{(\mathbb{R}^+)^{n-1}} \mathbb{E}\left[F_i^2 f_h''(S_n) + \frac{1}{2}(1 - F_i^2) f_h'(S_n) \;\middle|\; \mathbf{F}_{\hat{i}}^2 = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)\right] \\
&\qquad\qquad \mathbb{P}_{\mathbf{F}_{\hat{i}}^2}(\mathrm{d}x_1, \ldots, \mathrm{d}x_{i-1}, \mathrm{d}x_{i+1}, \ldots, \mathrm{d}x_n).
\end{aligned} \tag{3.7}$$

By independence, for any $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n > 0$, we have

$$\begin{aligned}
&\mathbb{E}\left[F_i^2 f_h''(S_n) + \frac{1}{2}(1 - F_i^2) f_h'(S_n) \;\middle|\; \mathbf{F}_{\hat{i}}^2 = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)\right] \\
&\quad = \mathbb{E}\left[F_i^2 f_h''\left(F_i^2 + \sum_{j \neq i}^{1,n} x_j\right) + \frac{1}{2}(1 - F_i^2) f_h'\left(F_i^2 + \sum_{j \neq i}^{1,n} x_j\right)\right].
\end{aligned}$$

Setting $x_{(i)} := \sum_{j \neq i}^{1,n} x_j > 0$ and $f_h^{x_{(i)}}(\cdot) := f_h(\cdot + x_{(i)})$, we have

$$\begin{aligned}
&\mathbb{E}\left[F_i^2 f_h''(S_n) + \frac{1}{2}(1 - F_i^2) f_h'(S_n) \;\middle|\; \mathbf{F}_{\hat{i}}^2 = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)\right] \\
&\quad = \mathbb{E}\left[F_i^2 (f_h^{x_{(i)}})''(F_i^2) + \frac{1}{2}(1 - F_i^2)(f_h^{x_{(i)}})'(F_i^2)\right].
\end{aligned} \tag{3.8}$$

Set $u_h^{x_{(i)}}(x) := \frac{1}{4} f_h^{x_{(i)}}(x^2)$, $x \in \mathbb{R}$, and note that

$$(u_h^{x_{(i)}})'(x) = \frac{1}{2} x (f_h^{x_{(i)}})'(x^2), \quad (u_h^{x_{(i)}})''(x) = x^2 (f_h^{x_{(i)}})''(x^2) + \frac{1}{2}(f_h^{x_{(i)}})'(x^2). \tag{3.9}$$

Note that the mapping $x \in \mathbb{R} \mapsto (u_h^{x_{(i)}})''(x)$ is continuous. Indeed, the mapping $x \in \mathbb{R} \mapsto (f_h^{x_{(i)}})'(x^2) = f_h'(x^2 + x_{(i)})$ is continuous (since f_h' is continuous on $\mathbb{R}^+$) and the mapping

$x \in \mathbb{R} \mapsto x^2(f_h^{x_{(i)}})''(x^2) = x^2 f_h''(x^2 + x_{(i)})$ is continuous (since f_h'' is continuous on $\mathbb{R}^+$; see the Stein equation (3.2)). We shall show later on that

$$\sup_{x\in\mathbb{R}} |(u_h^{x_{(i)}})''(x)| \leq 5 \tag{3.10}$$

and therefore $(u_h^{x_{(i)}})' \in \mathcal{C}_b^1$. By the relations (3.9) and the assumption (3.5), for any $i = 1, \ldots, n$, we have

$$\begin{aligned}\mathbb{E}\left[F_i^2(f_h^{x_{(i)}})''(F_i^2) + \frac{1}{2}(1-F_i^2)(f_h^{x_{(i)}})'(F_i^2)\right] &= \mathbb{E}[(u_h^{x_{(i)}})''(F_i) - F_i(u_h^{x_{(i)}})'(F_i)] \\ &= \mathbb{E}(u_h^{x_{(i)}})''(F_i)(1 - \langle C^{(i)}(F_i), \Delta^{(i)}(F_i)\rangle_{\mathfrak{H}_i}).\end{aligned} \tag{3.11}$$

Therefore by (3.8), (3.10) and (3.11), for any $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n > 0$, we have

$$\begin{aligned}&\left|\mathbb{E}\left[F_i^2 f_h''(S_n) + \frac{1}{2}(1-F_i^2)f_h'(S_n) \mid \mathbf{F}_{\hat{i}}^2 = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)\right]\right| \\ &\qquad \leq 5\mathbb{E}|1 - \langle C^{(i)}(F_i), \Delta^{(i)}F_i\rangle_{\mathfrak{H}_i}|.\end{aligned} \tag{3.12}$$

The claim follows taking first the modulus in (3.7) and using the inequality (3.12) (note that the r.h.s. of this inequality does not depend on $x_1, \ldots, x_{i-1}, x_i, \ldots, x_n$ and $h \in \mathcal{H}_{\{1\}}$) and then taking the supremum over all $h \in \mathcal{H}_{\{1\}}$.

It remains to prove (3.10). By the second relation in (3.9) and the fact that $x_{(i)} > 0$, we have

$$\begin{aligned}\sup_{x\in\mathbb{R}} |(u_h^{x_{(i)}})''(x)| &\leq \sup_{x\in\mathbb{R}} |x^2(f_h^{x_{(i)}})''(x^2)| + \frac{1}{2}\sup_{x\in\mathbb{R}} |(f_h^{x_{(i)}})'(x^2)| \\ &\leq \sup_{x\in\mathbb{R}} |(x^2 + x_{(i)})f_h''(x^2 + x_{(i)})| + \frac{1}{2}\sup_{x\in\mathbb{R}} |f_h'(x^2 + x_{(i)})| \\ &= \sup_{x>0} |xf_h''(x)| + \frac{1}{2}\|f_h'\|_\infty.\end{aligned}$$

Since $h \in \mathcal{H}_{\{1\}}$, we have $h \in \mathcal{C}_{1/2}$ and $\|h\|_\infty, \|h'\|_\infty \leq 1$. So by Lemma 3.1$(i)$ (note that in our case $r = n/2$ and $\lambda = 1/2$) it follows $\|f_h'\|_\infty \leq 2$, and by Lemma 3.1(ii) it follows $\sup_{x>0} |xf_h''(x)| \leq 4$. The proof is completed. □

3.2 Chi-Square approximation in the Kolmogorov distance

In this subsection we give a bound on the Kolmogorov distance between the law of S_n (defined as in the statement of Theorem 3.2) and the Chi-Square distribution with n degrees of freedom.

The following corollary holds, where we adopt the convention $0^0 := 1$.

Corollary 3.4. *Let the assumptions and notation of Theorem 3.2 prevail. Then, for any $n \in \mathbb{N} \setminus \{1\}$, we have*

$$\begin{aligned}d_K(S_n, Z_n) \leq \sqrt{10A_n}&\sqrt{\sum_{i=1}^n \mathbb{E}|1 - \langle C^{(i)}(F_i), \Delta^{(i)}(F_i)\rangle_{\mathfrak{H}_i}|} \\ &+ \frac{5}{2}\sum_{i=1}^n \mathbb{E}|1 - \langle C^{(i)}(F_i), \Delta^{(i)}(F_i)\rangle_{\mathfrak{H}_i}|,\end{aligned} \tag{3.13}$$

where

$$A_n := \frac{1}{2^{n/2}\Gamma(n/2)}(n-2)^{n/2-1}\mathrm{e}^{-(n-2)/2}. \tag{3.14}$$

For $n \in \mathbb{N} \setminus \{1,2\}$, a further upper bound is obtained noticing that

$$\sqrt{10A_n} \leq \frac{\sqrt{5}}{\pi^{1/4}} \frac{1}{(n-2)^{1/4}}. \tag{3.15}$$

The proof of Corollary 3.4 exploits the following lemma proved in [12] (take $m = 1$ in Proposition 2.3(i) therein and note that the distance d_1 in [12] coincides with the distance d_{bW}).

Lemma 3.5. *Let X be any real-valued random variable and let Y be a continuous real-valued random variable with probability density function p. If there exists a positive constant $A > 0$ such that $p(y) \leq A$ for all $y \in \mathbb{R}$, then*

$$d_K(X,Y) \leq \sqrt{2A}\sqrt{d_{bW}(X,Y)} + \frac{d_{bW}(X,Y)}{2}.$$

Proof. (Corollary 3.4) The inequality (3.13) immediately follows by Theorem 3.2, Lemma 3.5 and the inequality (3.4). Indeed, letting p_n denote the density of Z_n, for any $x \in \mathbb{R}$ and $n \geq 2$, we have

$$p_n(x) = \mathbf{1}_{\{x \geq 0\}} \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} \mathrm{e}^{-x/2} \leq \frac{1}{2^{n/2}\Gamma(n/2)} (n-2)^{n/2-1} \mathrm{e}^{-(n-2)/2} = A_n. \tag{3.16}$$

The inequality (3.15) follows noticing that

$$A_n \leq \frac{1}{2\sqrt{\pi(n-2)}}, \quad n \geq 3 \tag{3.17}$$

which, in turn, follows from the relation $\Gamma(x+1) \geq \sqrt{2\pi} x^{x+1/2} \mathrm{e}^{-x}$, $x > 0$. □

4 A general lower bound

4.1 Preliminaries

We start introducing some notation and definitions. Let $\mathbf{A}$ be a real $n \times d$ matrix, $n, d \in \mathbb{N}$. We recall that the operator norm of $\mathbf{A}$ is defined as

$$\|\mathbf{A}\|_{op} := \sup_{\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}} \frac{\|\mathbf{A}\mathbf{y}\|}{\|\mathbf{y}\|} = \max_{\mathbf{y} \in S^{d-1}} \|\mathbf{A}\mathbf{y}\|,$$

where S^{d-1} is the unit sphere of $\mathbb{R}^d$. We recall that $s_{\min}(\mathbf{A})$ and $s_{\max}(\mathbf{A})$, the smallest and the biggest singular value of $\mathbf{A}$, respectively, are defined by

$$s_{\min}(\mathbf{A}) := \sqrt{\lambda_{\min}(\mathbf{A}^\top \mathbf{A})} \quad \text{and} \quad s_{\max}(\mathbf{A}) := \sqrt{\lambda_{\max}(\mathbf{A}^\top \mathbf{A})},$$

where $\lambda_{\min}(\mathbf{A}^\top \mathbf{A})$ and $\lambda_{\max}(\mathbf{A}^\top \mathbf{A})$ denote the smallest and the biggest eigenvalue of the $d \times d$ matrix $\mathbf{A}^\top \mathbf{A}$, respectively.

Let $(\mathfrak{T}, \mathfrak{m})$ be a metric space, $\mathfrak{T}' \subseteq \mathfrak{T}$ and $\delta > 0$ a positive constant. A subset $N_\delta \subseteq \mathfrak{T}'$ is called δ-net of $\mathfrak{T}'$ if every point in $\mathfrak{T}'$ is within a distance δ of some point of N_δ, i.e., $\mathfrak{T}'$ can be covered by balls with centers in N_δ and radii δ. The smallest possible cardinality of a δ-net of $\mathfrak{T}'$ is called covering number and it is denoted $\mathcal{N}(\mathfrak{T}', \mathfrak{m}, \delta)$.

The next lemmas will be exploited later on. They can be found in [42], see Lemma 4.2.13 p. 78, Lemma 4.1.5 p. 74 and Exercise 4.4.3(b) p. 84, therein, respectively.

Lemma 4.1. *Let $d_E(\cdot,\cdot) := \|\cdot - \cdot\|$ denote the Euclidean distance on $\mathbb{R}^d$, $d \in \mathbb{N}$. For any $\delta > 0$, it holds $\mathcal{N}(S^{d-1}, d_E, \delta) \leq \left(\frac{2}{\delta} + 1\right)^d$.*

Lemma 4.2. Let $\mathbf{A}$ be a real $n \times d$ matrix, $n, d \in \mathbb{N}$, and $\varepsilon \in (0,1)$. Suppose that

$$\|\mathbf{A}^\top \mathbf{A} - \mathbf{I}_d\|_{op} \le \varepsilon.$$

Then

$$(1-\varepsilon)\|\mathbf{y}\| \le \|\mathbf{A}\mathbf{y}\| \le (1+\varepsilon)\|\mathbf{y}\|, \quad \textit{for all } \mathbf{y} \in \mathbb{R}^d$$

and consequently

$$1-\varepsilon \le s_{\min}(\mathbf{A}) \le s_{\max}(\mathbf{A}) \le 1+\varepsilon.$$

Lemma 4.3. Let $\mathbf{A}$ be a real symmetric $d \times d$ matrix, $d \in \mathbb{N}$. Then, for any δ-net N_δ, $\delta \in (0,1/2)$, of the unit sphere S^{d-1}, we have

$$\sup_{\mathbf{x}\in N_\delta} |\langle \mathbf{A}\mathbf{x}, \mathbf{x}\rangle| \le \|\mathbf{A}\|_{op} \le \frac{1}{1-2\delta} \sup_{\mathbf{x}\in N_\delta} |\langle \mathbf{A}\mathbf{x}, \mathbf{x}\rangle|,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^d$.

4.2 Lower bound

In this section, all the random variables are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we denote by N_δ, $\delta > 0$, a δ-net of S^{d-1}, $d \in \mathbb{N}$, with cardinality less than or equal to $\left(\frac{2}{\delta}+1\right)^d$, whose existence is guaranteed by Lemma 4.1.

Theorem 4.4. For $n, d \in \mathbb{N}$, let F_{ij}, $1 \le i \le n$, $1 \le j \le d$, be real-valued random variables, let $\mathbf{F}_{n,d}$ be the random matrix $\mathbf{F}_{n,d} := (F_{ij})_{1\le i\le n,\, 1\le j\le d}$ and let $\delta \in (0,1/2)$. Assume that the random vectors $\{(\mathbf{F}_{n,d})_i\}_{1\le i\le n}$, $(\mathbf{F}_{n,d})_i := (F_{i1}, \ldots, F_{id})$, are independent and centered, and that, for every $\mathbf{x} = (x_1, \ldots, x_d) \in N_\delta$ and $i = 1, \ldots, n$, the random variable

$$F_i(\mathbf{x}) := \sum_{j=1}^{d} F_{ij} x_j \tag{4.1}$$

satisfies (3.5), i.e., there exist random elements $C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x})) \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{H}_i)$, where $\mathcal{H}_i$ is a real Hilbert space, such that

$$\mathbb{E}F_i(\mathbf{x})g(F_i(\mathbf{x})) = \mathbb{E}g'(F_i(\mathbf{x}))\langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathcal{H}_i}, \quad g \in \mathcal{C}_b^1. \tag{4.2}$$

Then, letting A_n be the term defined in (3.14), for any $\varepsilon \in (0,1)$ and $n \in \mathbb{N} \setminus \{1\}$, we have

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \le s_{\min}(\mathbf{F}_{n,d}/\sqrt{n}) \le s_{\max}(\mathbf{F}_{n,d}/\sqrt{n}) \le 1+\varepsilon\right) \\
&\quad \ge \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \le \left\|\frac{1}{\sqrt{n}}\mathbf{F}_{n,d}\mathbf{u}\right\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \textit{for all } \mathbf{u} \in \mathbb{R}^d\right) \\
&\quad \ge \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top \mathbf{F}_{n,d} - \mathbf{I}_d\right\|_{op} \le \varepsilon\right) \ge \mathfrak{L}_{n,d,\varepsilon,\delta},
\end{aligned}$$

where

$$\mathfrak{L}_{n,d,\varepsilon,\delta} := 1 - 4\left(\frac{2}{\delta}+1\right)^d \mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} - 2\left(\frac{2}{\delta}+1\right)^d \sqrt{M_{n,d}(\delta)}\sqrt{10A_n} - 5\left(\frac{2}{\delta}+1\right)^d M_{n,d}(\delta) \tag{4.3}$$

and

$$M_{n,d}(\delta) := \sum_{i=1}^{n} \max_{\mathbf{x}\in N_\delta} \mathbb{E}|1 - \langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|.$$

For $n \in \mathbb{N} \setminus \{1, 2\}$, a further lower bound is obtained by the inequality (3.15).

The proof of this theorem exploits the following result, which is shown later on in this section.

Proposition 4.5. For $n,d \in \mathbb{N}$, let F_{ij}, $1 \le i \le n$, $1 \le j \le d$, be real-valued random variables and let $\mathbf{F}_{n,d}$ be the random matrix $\mathbf{F}_{n,d} := (F_{ij})_{1\le i\le n,\,1\le j\le d}$. Assume that the random vectors $\{(\mathbf{F}_{n,d})_i\}_{1\le i\le n}$, $(\mathbf{F}_{n,d})_i := (F_{i1},\dots,F_{id})$, are independent and centered, and that, for every $\mathbf{x} = (x_1,\dots,x_d) \in S^{d-1}$ and $i = 1,\dots,n$, the random variable $F_i(\mathbf{x})$ defined by (4.1) *satisfies* (3.5)*, i.e., there exist random elements $C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x})) \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathcal{H}_i)$, where $\mathcal{H}_i$ is a real Hilbert space, such that* (4.2) *holds.*

Then, letting A_n be the term in (3.14)*, for any $\varepsilon \in (0,1)$, $\mathbf{x} \in S^{d-1}$ and $n \in \mathbb{N}\setminus\{1\}$, we have*

$$\begin{aligned}\mathbb{P}\left(1-\varepsilon \le \frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2 \le 1+\varepsilon\right) \ge 1 - 4\mathrm{e}^{-\frac{n}{8}\varepsilon^2}&\\ - 2\sqrt{10A_n}\sqrt{\sum_{i=1}^n \mathbb{E}|1-\langle C^{(i)}(F_i(\mathbf{x})),\Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|}&\\ + 5\sum_{i=1}^n \mathbb{E}|1-\langle C^{(i)}(F_i(\mathbf{x})),\Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|.&\end{aligned}$$

For $n \in \mathbb{N}\setminus\{1,2\}$, a further lower bound is obtained by the inequality (3.15)*.*

Proof. (Theorem 4.4) For every $\varepsilon \in (0,1)$ and $\delta \in (0,1/2)$, by Proposition 4.5 and the union bound, we have

$$\begin{aligned}\mathbb{P}\left(\frac{1}{1-2\delta}\max_{\mathbf{x}\in N_\delta}\left|\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2-1\right| > \varepsilon\right) &\le \sum_{\mathbf{x}\in N_\delta}\mathbb{P}\left(\left|\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2-1\right| > \varepsilon(1-2\delta)\right)\\ &\le 4\left(\frac{2}{\delta}+1\right)^d \mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} + 2\left(\frac{2}{\delta}+1\right)^d\sqrt{M_{n,d}(\delta)}\sqrt{10A_n} + 5\left(\frac{2}{\delta}+1\right)^d M_{n,d}(\delta),\end{aligned} \tag{4.4}$$

where we used that the cardinality of N_δ is less than or equal to $\left(\frac{2}{\delta}+1\right)^d$ and the quantity $M_{n,d}(\delta)$ is defined in the statement of Theorem 4.4. By Lemma 4.3 (and the fact that $N_\delta \subset S^{d-1}$) it follows

$$\begin{aligned}\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top\mathbf{F}_{n,d}-\mathbf{I}_d\right\|_{op} &\le \frac{1}{1-2\delta}\max_{\mathbf{x}\in N_\delta}\left|\left\langle\left(\frac{1}{n}(\mathbf{F}_{n,d})^\top\mathbf{F}_{n,d}-\mathbf{I}_d\right)\mathbf{x},\mathbf{x}\right\rangle\right|\\ &= \frac{1}{1-2\delta}\max_{\mathbf{x}\in N_\delta}\left|\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2-1\right|, \quad \mathbb{P}\text{-a.s..}\end{aligned}$$

Combining this relation with (4.4), for every $\varepsilon \in (0,1)$ and $\delta \in (0,1/2)$, we have

$$\mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top\mathbf{F}_{n,d}-\mathbf{I}_d\right\|_{op} \le \varepsilon\right) \ge \mathfrak{L}_{n,d,\varepsilon,\delta}.$$

The claim follows by Lemma 4.2. For the sake of completeness, in the Appendix C we verify the measurability of the event

$$\mathcal{E} := \{(1-\varepsilon)\|\mathbf{u}\| \le \|(\mathbf{F}_{n,d}/\sqrt{n})\mathbf{u}\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u}\in\mathbb{R}^d\}. \tag{4.5}$$

□

The proof of Proposition 4.5 exploits the next Lemma 4.6, i.e., the Chernoff bound for the Chi-Square law. For the sake of completeness, we provide the proof of this lemma in the Appendix D.

Lemma 4.6. *Let $Z_n \sim \chi_n^2$, $n \in \mathbb{N}$. Then, for any $\varepsilon \in (0,1)$ and $n \in \mathbb{N}$, we have*

$$\max\{\mathbb{P}(Z_n > (1+\varepsilon)n), \mathbb{P}(Z_n < (1-\varepsilon)n)\} \leq 2\mathrm{e}^{-\frac{n}{8}\varepsilon^2}.$$

Proof. (Proposition 4.5) For $\mathbf{x} = (x_1, \dots, x_d) \in S^{d-1}$, we have

$$\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{d} F_{ij}x_j\right)^2 = \frac{1}{n}\sum_{i=1}^{n} F_i(\mathbf{x})^2 =: S_n(\mathbf{x})/n.$$

For any $\varepsilon \in (0,1)$, $\mathbf{x} \in S^{d-1}$ and $n \geq 2$, by the inequality (3.13) of Corollary 3.4 (note that, by construction, the random variables $F_1(\mathbf{x}), \dots, F_n(\mathbf{x})$ are independent and centered) and Lemma 4.6 we have

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2 > 1+\varepsilon\right) &= \mathbb{P}\left(S_n(\mathbf{x}) > n(1+\varepsilon)\right) \\
&\leq \mathbb{P}(Z_n > (1+\varepsilon)n) + \sqrt{10A_n}\sqrt{\sum_{i=1}^{n}\mathbb{E}|1 - \langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|} \\
&\qquad + \frac{5}{2}\sum_{i=1}^{n}\mathbb{E}|1 - \langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}| \\
&\leq 2\mathrm{e}^{-\frac{n}{8}\varepsilon^2} + \sqrt{10A_n}\sqrt{\sum_{i=1}^{n}\mathbb{E}|1 - \langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|} \\
&\qquad + \frac{5}{2}\sum_{i=1}^{n}\mathbb{E}|1 - \langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|. \qquad (4.6)
\end{aligned}$$

Similarly, for any $\varepsilon \in (0,1)$, $\mathbf{x} \in S^{d-1}$ and $n \geq 2$, by (3.13) and Lemma 4.6 we have

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2 < 1-\varepsilon\right) &= \mathbb{P}\left(S_n(\mathbf{x}) < n(1-\varepsilon)\right) \\
&\leq 2\mathrm{e}^{-\frac{n}{8}\varepsilon^2} + \sqrt{10A_n}\sqrt{\sum_{i=1}^{n}\mathbb{E}|1 - \langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|} \\
&\qquad + \frac{5}{2}\sum_{i=1}^{n}\mathbb{E}|1 - \langle C^{(i)}(F_i(\mathbf{x})), \Delta^{(i)}(F_i(\mathbf{x}))\rangle_{\mathfrak{H}_i}|. \qquad (4.7)
\end{aligned}$$

The claim follows by (4.6) and (4.7) noticing that, for any $\varepsilon \in (0,1)$, $\mathbf{x} \in S^{d-1}$ and $n \geq 2$,

$$\mathbb{P}\left(1-\varepsilon \leq \frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2 \leq 1+\varepsilon\right) = 1 - \mathbb{P}\left(\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2 > 1+\varepsilon\right) - \mathbb{P}\left(\frac{1}{n}\|\mathbf{F}_{n,d}\mathbf{x}\|^2 < 1-\varepsilon\right).$$

□

Remark 4.7. The independence of the random variables $F_1, \dots, F_n$ is a crucial assumption of Theorem 3.2 and its Corollary 3.4, which, at the moment, we do not know how to relax. Consequently, in the statement of Theorem 4.4 we have to assume that the random vectors $\{(\mathbf{F}_{n,d})_i\}_{1\leq i\leq n}$ are independent (indeed the proof of this theorem relies on Proposition 4.5 whose proof, in turn, exploits Corollary 3.4). In conclusion, to allow even a mild dependence among the random vectors $\{(\mathbf{F}_{n,d})_i\}_{1\leq i\leq n}$ in Theorem 4.4 (and, consequently, to allow even a mild dependence among the rows of a chaotic random matrix), one should allow a mild dependence among the random variables $F_1, \dots, F_n$ in Theorem 3.2. We leave such an extension to a future research.

5 Random matrices with chaotic entries

Hereafter, we shall consider the following framework. For $i = 1, \dots, n$, $n \in \mathbb{N}$, let $\mathfrak{H}_i$ be real separable Hilbert spaces and let $X_1 = \{X_1(u)\}_{u \in \mathfrak{H}_1}, \dots, X_n = \{X_n(u)\}_{u \in \mathfrak{H}_n}$ be independent isonormal Gaussian processes, X_i over $\mathfrak{H}_i$. We suppose that X_i is defined on a complete probability space $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$, with $\mathcal{F}_i := \sigma(X_i)$, $i = 1, \dots, n$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space defined by $\Omega := \prod_{i=1}^n \Omega_i$, $\mathcal{F} := \bigotimes_{i=1}^n \mathcal{F}_i$, $\mathbb{P} := \bigotimes_{i=1}^n \mathbb{P}_i$, and define $X_1, \dots, X_n$ over Ω setting $X_i(\omega) := X_i(\omega_i)$, $\omega := (\omega_1, \dots, \omega_n) \in \Omega$. All the random quantities considered in this section are defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

For $i = 1, \dots, n$, let $\mathcal{S}_i$ be the set of all smooth cylindrical random variables of the form $F = g(X_i(u_1), \dots, X_i(u_m))$, $m \in \mathbb{N}$, $g : \mathbb{R}^m \to \mathbb{R}$ infinitely differentiable with compact support and $u_k \in \mathfrak{H}_i$, $k = 1, \dots, m$. We define

$$D^{(i)} F := \sum_{k=1}^m \frac{\partial g}{\partial x_k}(X_i(u_1), \dots, X_i(u_m)) u_k, \quad i = 1, \dots, n,$$

and let $\mathbb{D}_i^{1,2}$ be the closure of $\mathcal{S}_i$ with respect to the norm $\|\cdot\|_{1,2,i}$ defined by

$$\|F\|_{1,2,i}^2 := \mathbb{E}F^2 + \mathbb{E}\|D^{(i)} F\|_{\mathfrak{H}_i}^2, \quad i = 1, \dots, n.$$

Using an obvious notation, we denote by L_i^{-1} the inverse of the Ornstein-Uhlenbeck operator L_i (see the Appendix A for details). As in Section 4, we still denote by N_δ, $\delta > 0$, a δ-net of S^{d-1}, $d \in \mathbb{N}$, with cardinality less than or equal to $\left(\frac{2}{\delta} + 1\right)^d$, whose existence is guaranteed by Lemma 4.1.

The following theorem holds.

Theorem 5.1. *For $n, d \in \mathbb{N}$, let $F_{ij} \in \mathbb{D}_i^{1,2}$, $1 \le i \le n$, $1 \le j \le d$, be centered random variables, let $\mathbf{F}_{n,d}$ be the random matrix $\mathbf{F}_{n,d} := (F_{ij})_{1 \le i \le n, 1 \le j \le d}$ and let $\delta \in (0, 1/2)$. For every $\mathbf{x} = (x_1, \dots, x_d) \in N_\delta$ and $i = 1, \dots, n$, let the random variable $F_i(\mathbf{x})$ be defined as in* (4.1).

Then, letting A_n be the term in (3.14)*, for any $\varepsilon \in (0,1)$ and $n \in \mathbb{N} \setminus \{1\}$, we have*

$$\begin{aligned} &\mathbb{P}\left(1 - \varepsilon \le s_{\min}(\mathbf{F}_{n,d}/\sqrt{n}) \le s_{\max}(\mathbf{F}_{n,d}/\sqrt{n}) \le 1 + \varepsilon\right) \\ &\ge \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \le \left\|\frac{1}{\sqrt{n}} \mathbf{F}_{n,d} \mathbf{u}\right\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d\right) \\ &\ge \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top \mathbf{F}_{n,d} - \mathbf{I}_d\right\|_{op} \le \varepsilon\right) \ge \mathfrak{L}'_{n,d,\varepsilon,\delta}, \end{aligned}$$

where $\mathfrak{L}'_{n,d,\varepsilon,\delta}$ is defined by (4.3) *with*

$$M_{n,d}(\delta) := \sum_{i=1}^n \max_{\mathbf{x} \in N_\delta} \mathbb{E}\left|1 - \langle -D^{(i)} L_i^{-1} F_i(\mathbf{x}), D^{(i)} F_i(\mathbf{x}) \rangle_{\mathfrak{H}_i}\right|.$$

For $n \in \mathbb{N} \setminus \{1, 2\}$, a further lower bound is obtained by the inequality (3.15).

Proof. We start noticing that, for any $i = 1, \dots, n$ and $\mathbf{x} \in N_\delta$, we have $F_i(\mathbf{x}) \in \mathbb{D}_i^{1,2}$. This follows e.g. by Proposition 2.3.7 in [21]. Indeed $F_{ij} \in \mathbb{D}_i^{1,2}$ for any $1 \le i \le n$ and $1 \le j \le d$, the function $\varphi_{\mathbf{x}}(y_1, \dots, y_d) := \sum_{j=1}^d y_j x_j$ is continuously differentiable on $\mathbb{R}^d$ with bounded partial derivatives for any $\mathbf{x} \in N_\delta$, and $F_i(\mathbf{x}) = \varphi_{\mathbf{x}}(F_{i1}, \dots, F_{id})$ for any $i = 1, \dots, n$ and $\mathbf{x} \in N_\delta$. Since the random variables F_{ij} are centered so it is $F_i(\mathbf{x})$. Therefore by the integration by parts formula (A.5) we have that (3.5) holds with $F_i := F_i(\mathbf{x})$,

$$C^{(i)}(F_i) := -D^{(i)} L_i^{-1} F_i(\mathbf{x}) \quad \text{and} \quad \Delta^{(i)}(F_i) := D^{(i)} F_i(\mathbf{x}).$$

Note that the vectors $\{(\mathbf{F}_{n,d})_i\}_{1 \le i \le n}$, $(\mathbf{F}_{n,d})_i := (F_{i1}, \dots, F_{id})$, are independent since the isonormal Gaussian processes $X_1, \dots, X_n$ are independent. The claim follows by Theorem 4.4. ☐

5.1 Wiener chaoses

For random matrices with chaotic entries we have the following theorem (as usual, we refer the reader to the Appendix A for any unexplained notation).

Theorem 5.2. *Let $n \in \mathbb{N} \setminus \{1\}$, $d, q_i \in \mathbb{N}$, $i = 1, \dots, n$, and define the random matrix $\mathbf{F}_{n,d} := (I_{q_i}^{X_i}(f_{ij}))_{1 \le i \le n,\, 1 \le j \le d}$, where $f_{ij} \in \mathfrak{H}_i^{\circ q_i}$.*

Then, letting A_n be the term in (3.14)*, for any $\varepsilon \in (0,1)$, we have*

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \le s_{\min}(\mathbf{F}_{n,d}/\sqrt{n}) \le s_{\max}(\mathbf{F}_{n,d}/\sqrt{n}) \le 1+\varepsilon\right) \\
&\quad \ge \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \le \left\|\frac{1}{\sqrt{n}}\mathbf{F}_{n,d}\mathbf{u}\right\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d\right) \\
&\quad \ge \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top \mathbf{F}_{n,d} - \mathbf{I}_d\right\|_{op} \le \varepsilon\right) \ge \mathfrak{L}''_{n,d,\varepsilon},
\end{aligned} \tag{5.1}$$

where

$$\mathfrak{L}''_{n,d,\varepsilon} := 1 - \inf_{\delta \in (0,1/2)} \left(\frac{2}{\delta}+1\right)^d \left(4\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} + 2\sqrt{M_{n,d}}\sqrt{10A_n} + 5M_{n,d}\right)$$

and

$$\begin{aligned}
M_{n,d} := \sum_{i=1}^n \Bigg[&\sum_{j=1}^d |1 - q_i!\|f_{ij}\|^2_{\mathfrak{H}^{\otimes q_i}}| + (q_i!)\sum_{j\neq k}^{1,d} |\langle f_{ij}, f_{ik}\rangle_{\mathfrak{H}^{\otimes q_i}}| \\
&+ q_i \sum_{r=1}^{q_i-1} \sqrt{(2q_i-2r)!}(r-1)!\binom{q_i-1}{r-1}^2 \\
&\times \left(\sum_{j=1}^d \|f_{ij} \otimes_r f_{ij}\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} + \sum_{j\neq k}^{1,d} \|f_{ij} \otimes_{q_i-r} f_{ij}\|_{\mathfrak{H}_i^{\otimes 2r}} \|f_{ik} \otimes_{q_i-r} f_{ik}\|_{\mathfrak{H}_i^{\otimes 2r}} \right) \Bigg]
\end{aligned} \tag{5.2}$$

with the conventions $\sum_{j\neq k}^{1,1} \dots := 0$ and $\sum_{r=1}^{0} \dots := 0$.

Proof. Let $N_\delta \subset S^{d-1}$ be a δ-net of S^{d-1}, $\delta \in (0,1/2)$, with cardinality less than or equal to $(\frac{2}{\delta}+1)^d$, which exists due to Lemma 4.1. Note that, for any $i = 1, \dots, n$ and $\mathbf{x} = (x_1, \dots, x_d) \in N_\delta$,

$$F_i(\mathbf{x}) := \sum_{j=1}^d I_{q_i}^{X_i}(f_{ij})x_j = I_{q_i}^{X_i}\left(\sum_{j=1}^d f_{ij}x_j\right).$$

Therefore, if $q_i = 1$, then by formulas (A.6) and (A.7) we have

$$\mathbb{E}|1 - \langle -D^{(i)}L_i^{-1}F_i(\mathbf{x}), D^{(i)}F_i(\mathbf{x})\rangle_{\mathfrak{H}_i}|^2 = \left(1 - \|\sum_{j=1}^d f_{ij}x_j\|^2_{\mathfrak{H}_i}\right)^2 \tag{5.3}$$

and, if $q_i \ge 2$, then by Proposition 3.2 in [20] it follows

$$\begin{aligned}
&\mathbb{E}|1 - \langle -D^{(i)}L_i^{-1}F_i(\mathbf{x}), D^{(i)}F_i(\mathbf{x})\rangle_{\mathfrak{H}_i}|^2 \\
&\le \left(1 - (q_i!)\|\sum_{j=1}^d f_{ij}x_j\|^2_{\mathfrak{H}_i^{\otimes q_i}}\right)^2 + (q_i)^2 \sum_{r=1}^{q_i-1} (2q_i-2r)!((r-1)!)^2 \binom{q_i-1}{r-1}^4 \\
&\qquad \times \|\left(\sum_{j=1}^d f_{ij}x_j\right) \otimes_r \left(\sum_{j=1}^d f_{ij}x_j\right)\|^2_{\mathfrak{H}_i^{\otimes 2(q_i-r)}}.
\end{aligned} \tag{5.4}$$

By Theorem 5.1 with $F_{ij} := I_{q_i}^{X_i}(f_{ij})$, relations (5.3), (5.4) and the Cauchy-Schwarz inequality, for any $\varepsilon \in (0,1)$ and $n \geq 2$, we have

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \leq s_{\min}(\mathbf{F}_{n,d}/\sqrt{n}) \leq s_{\max}(\mathbf{F}_{n,d}/\sqrt{n}) \leq 1+\varepsilon\right) \\
&\quad \geq \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \leq \left\|\frac{1}{\sqrt{n}}\mathbf{F}_{n,d}\mathbf{u}\right\| \leq (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d\right) \\
&\quad \geq \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top \mathbf{F}_{n,d} - \mathbf{I}_d\right\|_{op} \leq \varepsilon\right) \geq \mathfrak{L}^*_{n,d,\varepsilon,\delta}, \qquad (5.5)
\end{aligned}$$

where

$$\mathfrak{L}^*_{n,d,\varepsilon,\delta} := 1 - \left(\frac{2}{\delta}+1\right)^d \left(4\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} + 2\sqrt{M^*_{n,d}(\delta)}\sqrt{10A_n} + 5M^*_{n,d}(\delta)\right),$$

and

$$\begin{aligned}
M^*_{n,d}(\delta) := \sum_{i=1}^n \max_{\mathbf{x}\in N_\delta} \Bigg[\left(1-(q_i!)\|\sum_{j=1}^d f_{ij}x_j\|^2_{\mathfrak{H}_i^{\otimes q_i}}\right)^2 + (q_i)^2 \sum_{r=1}^{q_i-1} (2q_i-2r)!((r-1)!)^2 \binom{q_i-1}{r-1}^4 \\
\times \|\left(\sum_{j=1}^d f_{ij}x_j\right) \otimes_r \left(\sum_{j=1}^d f_{ij}x_j\right)\|^2_{\mathfrak{H}_i^{\otimes 2(q_i-r)}}\Bigg]^{1/2}.
\end{aligned}$$

Using the elementary inequality $\sqrt{a+b} \leq \sqrt{a}+\sqrt{b}$, $a,b \geq 0$, we have

$$\begin{aligned}
&M^*_{n,d}(\delta) \leq \widetilde{M}_{n,d}(\delta) \\
&\quad := \sum_{i=1}^n \max_{\mathbf{x}\in N_\delta} \Bigg[|1-(q_i!)\|\sum_{j=1}^d f_{ij}x_j\|^2_{\mathfrak{H}_i^{\otimes q_i}}| + q_i \sum_{r=1}^{q_i-1} \sqrt{(2q_i-2r)!}(r-1)!\binom{q_i-1}{r-1}^2 \\
&\qquad \times \|\left(\sum_{j=1}^d f_{ij}x_j\right) \otimes_r \left(\sum_{j=1}^d f_{ij}x_j\right)\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}}\Bigg]. \qquad (5.6)
\end{aligned}$$

For any $i = 1, \dots, n$, we have

$$\|\sum_{j=1}^d f_{ij}x_j\|^2_{\mathfrak{H}^{\otimes q_i}} = \langle \sum_{j=1}^d f_{ij}x_j, \sum_{j=1}^d f_{ij}x_j\rangle_{\mathfrak{H}^{\otimes q_i}} = \sum_{j=1}^d \|f_{ij}\|^2_{\mathfrak{H}^{\otimes q_i}} x_j^2 + \sum_{j\neq k}^{1,d} \langle f_{ij}, f_{ik}\rangle_{\mathfrak{H}^{\otimes q_i}} x_j x_k, \qquad (5.7)$$

and [by the definition of the contraction $\otimes_r$ (see formulas (A.3) and (A.4) in the Appendix A) and the triangular inequality]

$$\begin{aligned}
&\|\left(\sum_{j=1}^d f_{ij}x_j\right) \otimes_r \left(\sum_{j=1}^d f_{ij}x_j\right)\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} \\
&\qquad \leq \sum_{j=1}^d \|f_{ij} \otimes_r f_{ij}\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} x_j^2 + \sum_{j\neq k}^{1,d} \|f_{ij} \otimes_r f_{ik}\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} |x_j x_k|.
\end{aligned}$$

Note that, for $j \neq k$, it is easily checked that

$$\|f_{ij} \otimes_r f_{ik}\|^2_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} = \langle f_{ij} \otimes_{q_i-r} f_{ij}, f_{ik} \otimes_{q_i-r} f_{ik}\rangle_{\mathfrak{H}_i^{\otimes 2r}}$$

(see e.g. [20] p. 93), and so by the Cauchy-Schwarz inequality

$$\begin{aligned}
&\| \left(\sum_{j=1}^{d} f_{ij}x_j\right) \otimes_r \left(\sum_{j=1}^{d} f_{ij}x_j\right) \|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} \\
&\qquad \le \sum_{j=1}^{d} \|f_{ij}\otimes_r f_{ij}\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} x_j^2 + \sum_{j\neq k}^{1,d} \|f_{ij}\otimes_{q_i-r} f_{ij}\|_{\mathfrak{H}_i^{\otimes 2r}} \|f_{ik}\otimes_{q_i-r} f_{ik}\|_{\mathfrak{H}_i^{\otimes 2r}} |x_j x_k|.
\end{aligned} \tag{5.8}$$

By (5.6), (5.7) and (5.8) it follows

$$M^*_{n,d}(\delta) \le \widetilde{M}_{n,d}(\delta) \le \overline{M}_{n,d}(\delta), \tag{5.9}$$

where

$$\begin{aligned}
\overline{M}_{n,d}(\delta) := \sum_{i=1}^{n} \max_{\mathbf{x}\in N_\delta} \Bigg[&|1-(q_i!)\sum_{j=1}^{d} \|f_{ij}\|^2_{\mathfrak{H}^{\otimes q_i}} x_j^2| + (q_i!)\sum_{j\neq k}^{1,d} |\langle f_{ij}, f_{ik}\rangle_{\mathfrak{H}^{\otimes q_i}}||x_j x_k| \\
&+ q_i \sum_{r=1}^{q_i-1} \sqrt{(2q_i-2r)!}(r-1)! \binom{q_i-1}{r-1}^2 \\
\times &\left(\sum_{j=1}^{d} \|f_{ij}\otimes_r f_{ij}\|_{\mathfrak{H}_i^{\otimes 2(q_i-r)}} x_j^2 + \sum_{j\neq k}^{1,d} \|f_{ij}\otimes_{q_i-r} f_{ij}\|_{\mathfrak{H}_i^{\otimes 2r}} \|f_{ik}\otimes_{q_i-r} f_{ik}\|_{\mathfrak{H}_i^{\otimes 2r}} |x_j x_k|\right)\Bigg].
\end{aligned}$$

Since $\mathbf{x} = (x_1, \dots, x_d) \in S^{d-1}$, we have $\sum_{j=1}^{d} x_j^2 = 1$ and $|x_j x_k| \le 1$ for any $j, k = 1, \dots, d$. In particular,

$$|1-(q_i!)\sum_{j=1}^{d} \|f_{ij}\|^2_{\mathfrak{H}^{\otimes q_i}} x_j^2| = |\sum_{j=1}^{d} x_j^2 (1-(q_i!)\|f_{ij}\|^2_{\mathfrak{H}^{\otimes q_i}})| \le \sum_{j=1}^{d} |1-(q_i!)\|f_{ij}\|^2_{\mathfrak{H}^{\otimes q_i}}|,$$

and so $\overline{M}_{n,d}(\delta) \le M_{n,d}$ which, combined with (5.9) and (5.5), gives, for any $\varepsilon \in (0,1)$ and $n \ge 2$,

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \le s_{\min}(\mathbf{F}_{n,d}/\sqrt{n}) \le s_{\max}(\mathbf{F}_{n,d}/\sqrt{n}) \le 1+\varepsilon\right) \\
&\quad \ge \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \le \left\|\frac{1}{\sqrt{n}}\mathbf{F}_{n,d}\mathbf{u}\right\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u}\in\mathbb{R}^d\right) \\
&\quad \ge \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top \mathbf{F}_{n,d} - \mathbf{I}_d\right\|_{op} \le \varepsilon\right) \ge \mathrm{L}_{n,d,\varepsilon,\delta},
\end{aligned} \tag{5.10}$$

where

$$\mathrm{L}_{n,d,\varepsilon,\delta} := 1 - 4\left(\frac{2}{\delta}+1\right)^d \mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} - 2\left(\frac{2}{\delta}+1\right)^d \sqrt{M_{n,d}}\sqrt{10A_n} - 5\left(\frac{2}{\delta}+1\right)^d M_{n,d}.$$

The claim follows optimizing the relation (5.10) over $\delta \in (0, 1/2)$. □

6 Proof of Theorem 2.1

By Theorem 5.2 and (3.15) (which holds since $n \geq 3$), for any $\varepsilon \in (0,1)$ and $\delta \in (0,1/2)$, we have

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \leq s_{\min}(\mathbf{F}_{n,d}/\sqrt{n}) \leq s_{\max}(\mathbf{F}_{n,d}/\sqrt{n}) \leq 1+\varepsilon\right) \\
&\quad \geq \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \leq \left\|\frac{1}{\sqrt{n}}\mathbf{F}_{n,d}\mathbf{u}\right\| \leq (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d\right) \\
&\quad \geq \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_{n,d})^\top \mathbf{F}_{n,d} - \mathbf{I}_d\right\|_{op} \leq \varepsilon\right) \\
&\quad \geq 1 - 5\left(\frac{2}{\delta}+1\right)^d \left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{M_{n,d}} + M_{n,d}\right), \qquad (6.1)
\end{aligned}$$

where the quantity $M_{n,d}$ is defined by (5.2). Exploiting (2.1), (2.2) and (2.3), we have

$$\begin{aligned}
M_{n,d} &\leq nd\varphi_{1,n,d} + \left(\sum_{i=1}^n q_i!\right) d(d-1)\varphi_{2,n,d} \\
&\qquad + \left(\sum_{i=1}^n q_i \sum_{r=1}^{q_i-1} \sqrt{(2q_i-2r)!}(r-1)!\binom{q_i-1}{r-1}^2\right)(d\varphi_{3,n,d} + d(d-1)\varphi_{3,n,d}^2).
\end{aligned}$$

By this inequality and the assumption (2.4), we have

$$\begin{aligned}
M_{n,d} &\leq nd\varphi_{1,n,d} + Q!nd(d-1)\varphi_{2,n,d} \\
&\qquad + \left(\sum_{i=1}^n q_i \sum_{r=1}^{q_i-1} \sqrt{(2q_i-2r)!}\frac{[(q_i-1)!]^2}{(r-1)![(q_i-r)!]^2}\right)(d\varphi_{3,n,d} + d(d-1)\varphi_{3,n,d}^2) \\
&\leq nd\varphi_{1,n,d} + Q!nd(d-1)\varphi_{2,n,d} \\
&\qquad + Q(Q-1)[(Q-1)!]^2\sqrt{(2Q-2)!}\, n(d\varphi_{3,n,d} + d(d-1)\varphi_{3,n,d}^2) \leq \Phi_{n,d}.
\end{aligned}$$

The claim follows combining this relation with (6.1).

7 Proof of Theorem 2.2

Let $\Phi_n := \Phi_{n,d_n}$, where $\Phi_{n,d}$ is defined in the statement of Theorem 2.1. We start noticing that the first derivative of the function $\Delta_n(\delta)$, defined by

$$\delta \in (0,1/2) \xrightarrow{\Delta_n} \left(\frac{2}{\delta}+1\right)^{d_n} \left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n} + \Phi_n\right), \qquad (7.1)$$

has the same sign of the mapping

$$\delta \in (0,1/2) \mapsto \mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2}\left(\frac{n\varepsilon^2 f(\delta)}{4d_n} - 1\right) - \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n} - \Phi_n,$$

and the sign of this latter function is positive if and only if

$$\left(\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n}\right)\mathrm{e}^{\frac{n}{8}\varepsilon^2(1-2\delta)^2} + 1 < \frac{\varepsilon^2 f(\delta)}{4(d_n/n)}.$$

We shall show later on that there exists an index $\widetilde{n}_\varepsilon \geq 2$ such that

$$\left(\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n}\right)\mathrm{e}^{\frac{n\varepsilon^2}{8}} \leq c_\varepsilon \frac{n}{d_n}, \quad \forall\, n > \widetilde{n}_\varepsilon. \qquad (7.2)$$

So

$$\left(\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}} \frac{1}{(n-2)^{1/4}} \sqrt{\Phi_n}\right) \mathrm{e}^{\frac{n}{8}\varepsilon^2(1-2\delta)^2} + 1 < \left(\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}} \frac{1}{(n-2)^{1/4}} \sqrt{\Phi_n}\right) \mathrm{e}^{\frac{n\varepsilon^2}{8}} + 1$$
$$\leq 1 + c_\varepsilon \frac{n}{d_n}, \quad \forall\, n > \widetilde{n}_\varepsilon. \tag{7.3}$$

By the choice of ε and the definition of c_ε, we have

$$\frac{4c_\varepsilon}{\varepsilon^2} < f(\delta^*). \tag{7.4}$$

So, since

$$\frac{4(1 + c_\varepsilon \frac{n}{d_n})\frac{d_n}{n}}{\varepsilon^2} \to \frac{4c_\varepsilon}{\varepsilon^2}, \quad \text{as } n \to \infty,$$

we have that there exists $\overline{n}_\varepsilon \geq 2$ such that

$$0 < \frac{4(1 + c_\varepsilon \frac{n}{d_n})\frac{d_n}{n}}{\varepsilon^2} < f(\delta^*), \quad \forall\, n > \overline{n}_\varepsilon. \tag{7.5}$$

For $n > \overline{n}_\varepsilon$, consider the equation of third degree in $\delta \in (0, 1/2)$:

$$\frac{4(1 + c_\varepsilon \frac{n}{d_n})\frac{d_n}{n}}{\varepsilon^2} = f(\delta).$$

Due to (7.5), we have that this equation has two distinct positive solutions (on the domain $(0, 1/2)$) given by Cardano's formulas. These solutions are (from the smallest to the biggest) $\delta_1(n, \varepsilon)$ (defined in the statement of the theorem) and

$$\delta_2(n,\varepsilon) := 2\sqrt{\frac{7}{12}} \cos\left(\frac{\pi}{3} - \frac{1}{3}\arctan\left(2\frac{\sqrt{\frac{343}{1728} - \frac{1}{4}\left(\frac{3}{4} + \frac{2(\frac{d_n}{n} + c_\varepsilon)}{\varepsilon^2}\right)^2}}{\frac{3}{4} + \frac{2(\frac{d_n}{n} + c_\varepsilon)}{\varepsilon^2}}\right)\right). \tag{7.6}$$

It is easily realized that, for $n > \overline{n}_\varepsilon$,

$$\frac{4(\frac{d_n}{n} + c_\varepsilon)}{\varepsilon^2} < f(\delta), \quad \forall\, \delta \in (\delta_1(n,\varepsilon), \delta_2(n,\varepsilon)) \subset (0, 1/2). \tag{7.7}$$

By (7.3) and (7.7) we have that, for any $n > \max\{\widetilde{n}_\varepsilon, \overline{n}_\varepsilon\}$,

$$\left(\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}} \frac{1}{(n-2)^{1/4}} \sqrt{\Phi_n}\right) \mathrm{e}^{\frac{n}{8}\varepsilon^2(1-2\delta)^2} + 1 < 1 + c_\varepsilon \frac{n}{d_n} = \frac{n}{d_n}\left(\frac{d_n}{n} + c_\varepsilon\right)$$
$$< \frac{\varepsilon^2 f(\delta)}{4(d_n/n)}, \quad \forall\, \delta \in (\delta_1(n,\varepsilon), \delta_2(n,\varepsilon)).$$

So, by the argument given at the beginning of the proof, for $n > \max\{\widetilde{n}_\varepsilon, \overline{n}_\varepsilon\}$, the function $\Delta_n(\cdot)$ increases on $(\delta_1(n,\varepsilon), \delta_2(n,\varepsilon))$. Thus by (2.5) (note that all the assumptions of

Theorem 2.1 are satisfied), $\forall\, n > n_\varepsilon := \max\{n'', \widetilde{n}_\varepsilon, \overline{n}_\varepsilon\}$ we get

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \le s_{\min}\left(\frac{\mathbf{F}_n}{\sqrt{n}}\right) \le s_{\max}\left(\frac{\mathbf{F}_n}{\sqrt{n}}\right) \le 1+\varepsilon\right)\\
&\quad\ge \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \le \left\|\frac{1}{\sqrt{n}}\mathbf{F}_n\mathbf{u}\right\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u}\in\mathbb{R}^d\right)\\
&\quad\ge \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_n)^\top\mathbf{F}_n - \mathbf{I}_d\right\|_{op} \le \varepsilon\right)\\
&\quad\ge 1-5\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)^{d_n}\left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2} + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n}+\Phi_n\right)\\
&\quad\ge 1-5\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)^{d_n}\left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2} + c_\varepsilon\frac{n}{d_n}\mathrm{e}^{-n\varepsilon^2/8}\right),
\end{aligned}$$

which gives (2.9). Here, the latter inequality follows by (7.2).

We now prove the relation (2.12). We preliminary note that $\delta_1(n,\varepsilon) \to \overline{\delta}_\varepsilon$, as $n \to \infty$. It turns out that $\overline{\delta}_\varepsilon \in (0, 1/2)$. Indeed, by (7.4) the equation in δ:

$$\frac{4c_\varepsilon}{\varepsilon^2} = f(\delta), \quad \delta \in (0, 1/2)$$

has two positive solutions on $(0, 1/2)$, and $\overline{\delta}_\varepsilon$ is the smallest one (as it follows again by Cardano's formulas). We have

$$\begin{aligned}
\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)^{d_n}\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2} &= \mathrm{e}^{d_n\log\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2}\\
&= \mathrm{e}^{n\left[\frac{d_n}{n}\log\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)-\frac{1}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2\right]}.
\end{aligned}$$

Since

$$\frac{d_n}{n}\log\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right) - \frac{1}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2 \to -\frac{1}{8}\varepsilon^2(1-2\overline{\delta}_\varepsilon)^2 < 0, \quad \text{as } n\to\infty$$

we clearly have

$$\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)^{d_n}\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2} = \mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\overline{\delta}_\varepsilon)^2(1+o(1))}, \quad \text{as } n\to\infty. \tag{7.8}$$

And similarly one has

$$\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)^{d_n}\frac{n}{d_n}\mathrm{e}^{-\frac{n}{8}\varepsilon^2} = \frac{n}{d_n}\mathrm{e}^{n\left[\frac{d_n}{n}\log\left(\frac{2}{\delta_1(n,\varepsilon)}+1\right)-\frac{1}{8}\varepsilon^2\right]} = \frac{n}{d_n}\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1+o(1))}, \quad \text{as } n\to\infty. \tag{7.9}$$

The relation (2.12) follows on combining (7.8) and (7.9).

It remains to prove (7.2). By (2.8) and the fact that $d_n \ge 1$, we have that $\varphi_{3,n}$ tends to zero, as $n \to \infty$. So

$$\exists\, n^* \ge 2 \text{ such that } \varphi_{3,n} \le 1, \text{ for any } n > n^*. \tag{7.10}$$

Therefore

$$\Phi_n \le \kappa_Q n d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}, \quad \forall\, n > n^*.$$

So, for any $n > \max\{n'_\varepsilon, n^*\}$,

$$\begin{aligned}
&\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n} \\
&\le \kappa_Q n d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\} + \frac{\sqrt{5}}{\pi^{1/4}}\left(\frac{n}{n-2}\right)^{1/4}\sqrt{\kappa_Q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}} \\
&= \sqrt{\kappa_Q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}}\Bigg(n^{1/2}\sqrt{\kappa_Q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}} \\
&\qquad + \frac{\sqrt{5}}{\pi^{1/4}}\left(\frac{n}{n-2}\right)^{1/4}\Bigg).
\end{aligned}$$

Therefore, using that $n \ge 3$, the assumption (2.8), that $d_n \ge 1$ and that the mapping $x \mapsto x^3 \mathrm{e}^{-\frac{x}{4}}\mathbf{1}\{x \ge 0\}$ has a point of maximum at $x = 12$, we have that, for any $n > \max\{n'_\varepsilon, n^*\}$,

$$\begin{aligned}
&\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n} \\
&\quad\le \sqrt{\kappa_Q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}}\left(d_n^{-1}\sqrt{\kappa_Q \varepsilon^6 n^3 \mathrm{e}^{-\frac{n\varepsilon^2}{4}}} + \sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right) \\
&\quad\le \sqrt{\kappa_Q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}}\left(\mathrm{e}^{-3/2}\sqrt{12\kappa_Q} + \sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right). \qquad (7.11)
\end{aligned}$$

So

$$\begin{aligned}
\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n} &\le \check{c}\sqrt{\kappa_Q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}} \\
&\le \check{c}\sqrt{\kappa_Q \varepsilon^6 n^2 d_n^{-2}}\mathrm{e}^{-\frac{n\varepsilon^2}{8}} = c_\varepsilon \frac{n}{d_n}\mathrm{e}^{-\frac{n\varepsilon^2}{8}}, \quad \forall\, n > \widetilde{n}_\varepsilon := \max\{n^*, n'_\varepsilon\},
\end{aligned} \qquad (7.12)$$

where we set

$$\check{c} := \mathrm{e}^{-3/2}\sqrt{12\kappa_Q} + \sqrt{5}\left(\frac{3}{\pi}\right)^{1/4},$$

and the latter inequality (7.12) still follows by (2.8). The relation (7.2) is proved, and the proof of Theorem 2.2 is completed.

Remark 7.1. The aim of this remark is to highlight the role played by the "smallness condition" (2.8) in the proof of Theorem 2.2. To provide a vanishing upper bound for the infimum in (2.5), we study the sign of the first derivative of the function $\delta \mapsto \Delta_n(\delta)$, $\delta \in (0, \frac{1}{2})$, defined in (7.1). If one assumes (7.2), some analytical considerations show that the first derivative of $\Delta_n(\cdot)$ is positive on $(\delta_1(n,\varepsilon), \delta_2(n,\varepsilon)) \subset (0, \frac{1}{2})$, for all n large enough, where $\delta_1(n,\varepsilon)$ and $\delta_2(n,\varepsilon)$ are defined by (2.10) and (7.6), respectively. This clearly yields the lower bound (2.9). The "smallness condition" (2.8) is sufficient for (7.2) and can be verified for concrete chaotic models, see e.g. Section 9. Although we cannot a priori exclude that the optimization problem in (2.5) can be efficiently treated in a different manner (e.g. for specific classes of chaotic entries), our approach provides a fine lower bound, which can be useful to analyze various families of chaotic matrices.

Remark 7.2. By a close inspection of the proof of Theorem 2.2, one realizes that the index n_ε in (2.9) is defined by

$$n_\varepsilon := \max\{n'', \widetilde{n}_\varepsilon, \overline{n}_\varepsilon\}.$$

Here: n'' is provided by the assumption (ii) of Theorem 2.2; $\widetilde{n}_\varepsilon := \max\{n'_\varepsilon, n^*\}$, where n'_ε is provided by the assumption (2.8) and n^* is defined by (7.10); finally, $\overline{n}_\varepsilon$ is defined by relation (7.5).

Remark 7.3. Note that in (7.11) we bounded the quantity $\varepsilon^6 n^3 \mathrm{e}^{-\frac{n\varepsilon^2}{4}}$ by $12\mathrm{e}^{-3}$. Clearly, by taking n large we could bound this term by an arbitrarily small constant. We did not follow this route since not only it would complicate the computation of the index n_ε (see Remark 7.2), but above all since it would not change the substance, indeed the rate of the whole upper bound is provided by the quantity $\sqrt{n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}}$.

8 Proof of Theorem 2.5

The proof of Theorem 2.5 uses the following lemma which is proved in the Appendix E.

Lemma 8.1. *Let $\{\mathbf{A}_n\}_{n\in\mathbb{N}}$, $\mathbf{A}_n := (a_{jk}(n))_{1\le j,k\le d_n}$, $d_n \in \mathbb{N}$, be a sequence of $d_n \times d_n$ real, symmetric and definite positive matrices. Suppose that there exists two sequences of non-negative numbers $\{b_n\}_{n\in\mathbb{N}}$ and $\{c_n\}_{n\in\mathbb{N}}$ such that: $b_n, d_n c_n \to 0$ as $n\to\infty$, and, for any $n\in\mathbb{N}$,*

$$|a_{jj}(n) - 1| \le b_n, \quad \forall\, j = 1,\dots,d_n \tag{8.1}$$

and

$$\text{If } d_n \ge 2, \text{ then } |a_{jk}(n)| \le c_n, \quad \forall\, j,k = 1,\dots,d_n,\ j\neq k. \tag{8.2}$$

Let $\widetilde{n}$ be such that $1 - b_n - (d_n-1)c_n > 0$, for any $n > \widetilde{n}$. Then, setting $\mathbf{A}_n^{-1/2} = (\widetilde{a}_{jk}(n))_{1\le i,j\le d_n}$, for any $n>\widetilde{n}$, we have:

$$|\widetilde{a}_{jj}(n) - 1| \le \max\left\{\Big|\frac{1}{\sqrt{1+b_n+(d_n-1)c_n}} - 1\Big|, \Big|\frac{1}{\sqrt{1-b_n-(d_n-1)c_n}} - 1\Big|\right\},$$
$$\forall\, j = 1,\dots,d_n \tag{8.3}$$

and, if $d_n \ge 2$, then

$$|\widetilde{a}_{jk}(n)| \le d_n \max\left\{\Big|\frac{1}{\sqrt{1+b_n+(d_n-1)c_n}} - 1\Big|, \Big|\frac{1}{\sqrt{1-b_n-(d_n-1)c_n}} - 1\Big|\right\}$$
$$+ d_n b_n + [1 + d_n(d_n-1)]c_n, \quad j,k = 1,\dots,d_n,\ j\neq k. \tag{8.4}$$

In particular, if for any $n>\widetilde{n}$ it holds $1/2 \ge b_n + (d_n-1)c_n$, then, for any $n > \widetilde{n}$, we have:

$$|\widetilde{a}_{jj}(n) - 1| \le b_n + (d_n-1)c_n, \quad \forall\, j = 1,\dots,d_n \tag{8.5}$$

and

$$\text{if } d_n \ge 2, \text{ then } |\widetilde{a}_{jk}(n)| \le 2d_n b_n + [1 + 2d_n(d_n-1)]c_n, \quad \forall\, j,k = 1,\dots,d_n,\ j\neq k. \tag{8.6}$$

For reasons of readability, we divide the proof of Theorem 2.5 in two parts.

Part 1. We denote by $n'''' \ge 2$ an index such that $\widetilde{\varphi}_{1,n} + q!(d_n-1)\widetilde{\varphi}_{2,n} \le 1/2$, for any $n > n''''$ (note that such an index n'''' exists since by (2.15) we have that $\widetilde{\varphi}_{1,n} + q!(d_n - 1)\widetilde{\varphi}_{2,n} \to 0$, as $n\to\infty$).

We set $\check{n}_\varepsilon := \max\{n'_\varepsilon, n'', n''', n''''\}$. Here, the indexes n'_ε, n'' and n''' are defined in the statement of the theorem.

For $n > \check{n}_\varepsilon$, define the d_n-dimensional (column) vectors $\mathbf{V}_i^{(n)} := (\mathbf{\Sigma}^{(n)})^{-1/2}(\mathbf{D}_i^{(n)})^\top$, $i = 1,\dots,n$, and let $\mathbf{F}_n$ be the $n\times d_n$ random matrix whose rows are $(\mathbf{V}_1^{(n)})^\top, \dots, (\mathbf{V}_n^{(n)})^\top$

(the quantities $\mathbf{\Sigma}^{(n)}$ and $\mathbf{D}_i^{(n)}$ are defined in Section 2.3, before the statement of Theorem 2.5). We have

$$\frac{1}{n}(\mathbf{F}_n)^\top \mathbf{F}_n - \mathbf{I}_{d_n} = \frac{1}{n}(\mathbf{V}_1^{(n)} \dots \mathbf{V}_n^{(n)}) \begin{pmatrix} (\mathbf{V}_1^{(n)})^\top \\ (\mathbf{V}_2^{(n)})^\top \\ \cdot \\ \cdot \\ \cdot \\ (\mathbf{V}_n^{(n)})^\top \end{pmatrix} - \mathbf{I}_{d_n}$$
$$= \frac{1}{n}\sum_{i=1}^n (\mathbf{\Sigma}^{(n)})^{-1/2}(\mathbf{D}_i^{(n)})^\top \mathbf{D}_i^{(n)} (\mathbf{\Sigma}^{(n)})^{-1/2} - \mathbf{I}_{d_n}.$$

Multiplying this relation on the left and on the right by $(\mathbf{\Sigma}^{(n)})^{1/2}$ yields

$$\widehat{\mathbf{\Sigma}}^{(n)} - \mathbf{\Sigma}^{(n)} = (\mathbf{\Sigma}^{(n)})^{1/2}\left(\frac{1}{n}(\mathbf{F}_n)^\top \mathbf{F}_n - \mathbf{I}_{d_n}\right)(\mathbf{\Sigma}^{(n)})^{1/2}.$$

Taking the operator norm we get

$$\|\widehat{\mathbf{\Sigma}}^{(n)} - \mathbf{\Sigma}^{(n)}\|_{op} = \left\|(\mathbf{\Sigma}^{(n)})^{1/2}\left(\frac{1}{n}(\mathbf{F}_n)^\top \mathbf{F}_n - \mathbf{I}_{d_n}\right)(\mathbf{\Sigma}^{(n)})^{1/2}\right\|_{op}$$
$$\leq \left\|\frac{1}{n}(\mathbf{F}_n)^\top \mathbf{F}_n - \mathbf{I}_{d_n}\right\|_{op} \|\mathbf{\Sigma}^{(n)}\|_{op},$$

i.e.,

$$R_n \leq \left\|\frac{1}{n}(\mathbf{F}_n)^\top \mathbf{F}_n - \mathbf{I}_{d_n}\right\|_{op},$$

and so, for ε as in the statement of the theorem and any $n > \check{n}_\varepsilon$, we have

$$\mathbb{P}(R_n \leq \varepsilon) \geq \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_n)^\top \mathbf{F}_n - \mathbf{I}_{d_n}\right\|_{op} \leq \varepsilon\right). \tag{8.7}$$

Letting $(\widetilde{\sigma}^{(n)})_{jk}$, $1 \leq j, k \leq d_n$, denote the jk-entry of $(\mathbf{\Sigma}^{(n)})^{-1/2}$, we have

$$\mathbf{F}_n = \begin{pmatrix} I_q^{X_1}\left(\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{1k} f_k^{(n)}\right) & \dots & I_q^{X_1}\left(\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{d_n k} f_k^{(n)}\right) \\ I_q^{X_2}\left(\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{1k} f_k^{(n)}\right) & \dots & I_q^{X_2}\left(\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{d_n k} f_k^{(n)}\right) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ I_q^{X_n}\left(\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{1k} f_k^{(n)}\right) & \dots & I_q^{X_n}\left(\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{d_n k} f_k^{(n)}\right) \end{pmatrix}, \tag{8.8}$$

where $X_1, \dots, X_n$ are independent copies of the isonormal Gaussian process X. So the idea is to apply Theorem 2.1 to lower bound the rightmost probability in (8.7). We start noticing that the condition (2.4) is satisfied with $Q = q$. Then we note that the kernel of the ij-entry of the random matrix $\mathbf{F}_{n,d}$ in (8.8) is $\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)}$. We shall show in the second part of the proof that, $\forall\, n > \check{n}_\varepsilon$, it holds:

$$\sup_{1 \leq j \leq d_n}\left|1 - q!\left\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)}\right\|_{\mathfrak{H}^{\otimes q}}^2\right| \leq \varphi_{1,n}, \tag{8.9}$$

$$\text{if } d_n \geq 2, \text{ then } \sup_{1 \leq j \neq j' \leq d_n}\left|\langle \sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)}, \sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{j'k} f_k^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}\right| \leq \varphi_{2,n} \tag{8.10}$$

and

$$\text{if } q \geq 2, \text{ then } \sup_{1\leq j\leq d_n} \sup_{1\leq r\leq q-1} \Big\| \sum_{k=1}^{d_n} (\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)} \otimes_r \sum_{k=1}^{d_n} (\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)} \Big\|_{\mathfrak{H}^{\otimes 2(q-r)}} \leq \varphi_{3,n}. \quad (8.11)$$

Here the quantities $\varphi_{h,n}$, $h = 1, 2, 3$, are defined by

$$\varphi_{1,n} := \mathbf{1}\{d_n = 1\} \frac{19}{4} \widetilde{\varphi}_{1,n} + \mathbf{1}\{d_n \geq 2\} C_1(q) d_n^6 \max\{\widetilde{\varphi}_{1,n}, \widetilde{\varphi}_{2,n}\}, \quad (8.12)$$

where $C_1(q)$ is defined by (2.19),

$$\varphi_{2,n} := \mathbf{1}\{d_n \geq 2\} C_2(q) d_n^6 \max\{\widetilde{\varphi}_{1,n}, \widetilde{\varphi}_{2,n}\}, \quad C_2(q) := \frac{151}{4} + \frac{39}{2} q! + 10(q!)^2, \quad (8.13)$$

and

$$\varphi_{3,n} := \mathbf{1}\{q \geq 2\} C_3(q) d_n^6 \widetilde{\varphi}_{3,n}, \quad C_3(q) := \frac{37}{4} + \frac{9}{2} q! + 10(q!)^2. \quad (8.14)$$

By (8.7) and (2.5) we have

$$\begin{aligned} &\mathbb{P}(R_n \leq \varepsilon) \\ &\geq 1 - 5 \inf_{\delta \in (0,1/2)} \left(\frac{2}{\delta} + 1\right)^{d_n} \left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta)^2} + \frac{\sqrt{5}}{\pi^{1/4}} \frac{1}{(n-2)^{1/4}} \sqrt{\Phi_n} + \Phi_n \right), \forall\, n > \check{n}_\varepsilon. \end{aligned} \quad (8.15)$$

Here Φ_n is defined as in the statement of Theorem 2.1 with $Q := q$, $d = d_n$ and $\varphi_{h,n}$, $h = 1, 2, 3$, defined by (8.12), (8.13) and (8.14), respectively. We shall show later on that there exists an index $\widetilde{n}_\varepsilon \geq 2$ such that

$$\left(\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}} \frac{1}{(n-2)^{1/4}} \sqrt{\Phi_n} \right) \mathrm{e}^{\frac{n\varepsilon^2}{8}} \leq c'_\varepsilon \frac{n}{d_n}, \quad \forall\, n > \widetilde{n}_\varepsilon. \quad (8.16)$$

So, reasoning exactly as in the proof of Theorem 2.2, but using the inequality (8.15) in place of the inequality (2.5), we have the inequality (2.20) with

$$n_\varepsilon := \max\{\check{n}_\varepsilon, \widetilde{n}_\varepsilon, \overline{n}_\varepsilon\}$$

where $\overline{n}_\varepsilon$ is defined by (7.5) with c'_ε in place of c_ε. One proves relation (2.12), with c'_ε in place of c_ε, exactly as in the proof of Theorem 2.2, and so the lower bound in (2.20) tends to 1 as $n \to \infty$.

To conclude this first part of the proof it remains to prove (8.16). By (2.15) we have that $\varphi_{3,n}$ tends to zero, as $n \to \infty$, and so there exists $n^* \geq 2$ defined by (7.10) with $\varphi_{3,n}$ given by (8.14). Consequently,

$$\Phi_n \leq \kappa_q n d_n \max\{\varphi_{1,n}, d_n \varphi_{2,n}, d_n \varphi_{3,n}\}, \quad \forall\, n > n^*.$$

So, for any $n > n^*$, arguing as in the proof of Theorem 2.2, we have

$$\begin{aligned} \Phi_n + \frac{\sqrt{5}}{\pi^{1/4}} \frac{1}{(n-2)^{1/4}} \sqrt{\Phi_n} &= \sqrt{\kappa_q n^{1/2} d_n \max\{\varphi_{1,n}, d_n \varphi_{2,n}, d_n \varphi_{3,n}\}} \\ &\times \left(n^{1/2} \sqrt{\kappa_q n^{1/2} d_n \max\{\varphi_{1,n}, d_n \varphi_{2,n}, d_n \varphi_{3,n}\}} + \sqrt{5} \left(\frac{3}{\pi}\right)^{1/4} \right). \end{aligned} \quad (8.17)$$

By the definition of the sequences $\varphi_{h,n}$, $h=1,2,3$, and (2.15) we have

$$n^{1/2}\sqrt{\kappa_q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}} \leq \sqrt{\kappa_q C_1(q) n^{3/2} d_n^8 \max\{\widetilde{\varphi}_{1,n}, \widetilde{\varphi}_{2,n}, \widetilde{\varphi}_{3,n}\}}$$
$$\leq d_n^{-1}\sqrt{\kappa_q n^3 \varepsilon^6 \mathrm{e}^{-\frac{n\varepsilon^2}{4}}}, \quad \forall\, n > n'_\varepsilon. \tag{8.18}$$

By (8.17) and this latter inequality, arguing as for (7.11), $\forall\, n > \max\{n'_\varepsilon, n^*\}$, we have

$$\Phi_n + \frac{\sqrt{5}}{\pi^{1/4}}\frac{1}{(n-2)^{1/4}}\sqrt{\Phi_n}$$
$$\leq \sqrt{\kappa_q n^{1/2} d_n \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\}}\left(\mathrm{e}^{-3/2}\sqrt{12\kappa_q} + \sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right).$$

The claim (8.16), with $\widetilde{n}_\varepsilon := \max\{n'_\varepsilon, n^*\}$, easily follows by this latter inequality and (8.18) (along similar computations as for (7.12)).

Part 2. In this second part of the proof we show the relations (8.9), (8.10) and (8.11). We remark here that, although there is room to improve the bounds $\varphi_{h,n}$, $h=1,2,3$, (as can be easily realized by the derivations), they have the merit of being quite simple. This allows to put forward the computations, as we have seen in the first part of the proof of Theorem 2.5.

We divide this second part of the proof in four steps. In the first one we give some preliminary inequalities, in the other three steps we prove (8.9), (8.10) and (8.11), respectively.

Step 1: Preliminary inequalities.
For $n > \check{n}_\varepsilon$, set $\mathbf{A}_n := \mathbf{\Sigma}^{(n)} = (q!\langle f_j^{(n)}, f_k^{(n)}\rangle_{\mathfrak{H}^{\otimes q}})_{1\leq j,k\leq d_n}$, $b_n := \widetilde{\varphi}_{1,n}$ and $c_n := q!\widetilde{\varphi}_{2,n}$. By the assumption (2.15) we have that $b_n, d_n c_n \to 0$, as $n\to\infty$, and by the assumptions (2.16) and (2.17) we have that the conditions (8.1) and (8.2) hold. Note also that by the definition of the index $\check{n}_\varepsilon$ we have that $1/2 \geq b_n + (d_n - 1)c_n$ for any $n > \check{n}_\varepsilon$. Therefore, applying Lemma 8.1 we have that, $\forall\, n > \check{n}_\varepsilon$,

$$|(\widetilde{\sigma}^{(n)})_{jj} - 1| \leq \widetilde{\varphi}_{1,n} + q!(d_n - 1)\widetilde{\varphi}_{2,n}, \quad \forall\, j = 1, \ldots, d_n \tag{8.19}$$

and

$$\text{If } d_n \geq 2, \text{ then } |(\widetilde{\sigma}^{(n)})_{jk}| \leq 2d_n\widetilde{\varphi}_{1,n} + q![1 + 2d_n(d_n - 1)]\widetilde{\varphi}_{2,n},\ \forall\, j,k = 1,\ldots,d_n,\ j\neq k. \tag{8.20}$$

Note that by (8.19) and the definition of $\check{n}_\varepsilon$, it follows that, $\forall\, n > \check{n}_\varepsilon$,

$$|(\widetilde{\sigma}^{(n)})_{jj}| \leq 3/2, \quad \forall\, j = 1, \ldots, d_n \tag{8.21}$$

and

$$|(\widetilde{\sigma}^{(n)})^2_{jj} - 1| = |(\widetilde{\sigma}^{(n)})_{jj} + 1||(\widetilde{\sigma}^{(n)})_{jj} - 1|$$
$$\leq (|(\widetilde{\sigma}^{(n)})_{jj}| + 1)[\widetilde{\varphi}_{1,n} + q!(d_n - 1)\widetilde{\varphi}_{2,n}]$$
$$\leq \frac{5}{2}[\widetilde{\varphi}_{1,n} + q!(d_n - 1)\widetilde{\varphi}_{2,n}], \quad \forall\, j = 1, \ldots, d_n. \tag{8.22}$$

Note also that (8.20) (together with the elementary inequality $(a+b)^2 \leq 2a^2 + 2b^2$, $a, b \in \mathbb{R}$) gives, $\forall\, n > \check{n}_\varepsilon$,

$$(\widetilde{\sigma}^{(n)})^2_{jk} \leq 8d_n^2\widetilde{\varphi}^2_{1,n} + 4(q!)^2(1 + 4d_n^2(d_n - 1)^2)\widetilde{\varphi}^2_{2,n}, \tag{8.23}$$

for $d_n \geq 2$ and $j,k = 1,\dots,d_n$, $j \neq k$.
Step 2: Proof of (8.9).
For $n > \check{n}_\varepsilon$ and $j = 1,\dots,d_n$, we have

$$q!\Big\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\Big\|^2_{\mathfrak{H}^{\otimes q}} = q!(\widetilde{\sigma}^{(n)})^2_{jj}\|f_j^{(n)}\|^2_{\mathfrak{H}^{\otimes q}} + q!\mathbf{1}\{d_n \geq 2\}\sum_{k:\,k\neq j}^{1,d_n}(\widetilde{\sigma}^{(n)})^2_{jk}\|f_k^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ q!\mathbf{1}\{d_n \geq 2\}\sum_{k_1\neq k_2}^{1,d_n}(\widetilde{\sigma}^{(n)})_{jk_1}(\widetilde{\sigma}^{(n)})_{jk_2}\langle f_{k_1}^{(n)}, f_{k_2}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}},$$

and so, for $n > \check{n}_\varepsilon$ and $j = 1,\dots,d_n$,

$$\Big|1-q!\Big\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\Big\|^2_{\mathfrak{H}^{\otimes q}}\Big| \leq |1-q!(\widetilde{\sigma}^{(n)})^2_{jj}\|f_j^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}| + q!\mathbf{1}\{d_n \geq 2\}\sum_{k:\ k\neq j}^{1,d_n}(\widetilde{\sigma}^{(n)})^2_{jk}\|f_k^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ q!\mathbf{1}\{d_n \geq 2\}\sum_{k_1\neq k_2}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{jk_1}||(\widetilde{\sigma}^{(n)})_{jk_2}||\langle f_{k_1}^{(n)}, f_{k_2}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}|$$
$$= q!|(\widetilde{\sigma}^{(n)})^2_{jj} - 1|\|f_j^{(n)}\|^2_{\mathfrak{H}^{\otimes q}} + |1 - q!\|f_j^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}| + q!\mathbf{1}\{d_n \geq 2\}\sum_{k:\ k\neq j}^{1,d_n}(\widetilde{\sigma}^{(n)})^2_{jk}\|f_k^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ 2q!\mathbf{1}\{d_n \geq 2\}|(\widetilde{\sigma}^{(n)})_{jj}|\sum_{k_2:\ k_2\neq j}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{jk_2}||\langle f_j^{(n)}, f_{k_2}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}|$$
$$+ q!\mathbf{1}\{d_n \geq 3\}\sum_{k_2\neq k_1,\,k_1,k_2\neq j}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{jk_1}||(\widetilde{\sigma}^{(n)})_{jk_2}||\langle f_{k_1}^{(n)}, f_{k_2}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}|.$$

By this relation, (8.20), (8.21), (8.22), (8.23) and the assumptions (2.16) and (2.17), for any $n > \check{n}_\varepsilon$, it follows

$$\sup_{1\leq j\leq d_n}\Big|1 - q!\Big\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\Big\|^2_{\mathfrak{H}^{\otimes q}}\Big| \leq \frac{5}{2}(1+\widetilde{\varphi}_{1,n})[\widetilde{\varphi}_{1,n} + q!(d_n-1)\widetilde{\varphi}_{2,n}] + \widetilde{\varphi}_{1,n}$$
$$+ \mathbf{1}\{d_n \geq 2\}(d_n-1)(1+\widetilde{\varphi}_{1,n})[8d_n^2\widetilde{\varphi}^2_{1,n} + 4(q!)^2(1+4d_n^2(d_n-1)^2)\widetilde{\varphi}^2_{2,n}]$$
$$+ 3\mathbf{1}\{d \geq 2\}(d_n-1)q!\widetilde{\varphi}_{2,n}(2d_n\widetilde{\varphi}_{1,n} + q![1+2d_n(d_n-1)]\widetilde{\varphi}_{2,n})$$
$$+ q!\mathbf{1}\{d_n \geq 3\}(d_n-1)(d_n-2)\widetilde{\varphi}_{2,n}(8d_n^2\widetilde{\varphi}^2_{1,n} + 4(q!)^2(1+4d_n^2(d_n-1)^2)\widetilde{\varphi}^2_{2,n}).$$

By the definition of $\check{n}_\varepsilon$, we have that for any $n > \check{n}_\varepsilon$

$$\widetilde{\varphi}_{1,n} \leq 1/2 \quad\text{and}\quad \text{if } d_n \geq 2\text{, then } \widetilde{\varphi}_{2,n} \leq 1/2. \tag{8.24}$$

Therefore, for all $n > \check{n}_\varepsilon$, we have

$$\sup_{1\leq j\leq d_n}\Big|1 - q!\Big\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\Big\|^2_{\mathfrak{H}^{\otimes q}}\Big| \leq \frac{15}{4}[\widetilde{\varphi}_{1,n} + q!(d_n-1)\widetilde{\varphi}_{2,n}] + \widetilde{\varphi}_{1,n}$$
$$+ 6\mathbf{1}\{d_n \geq 2\}(d_n-1)[2d_n^2\widetilde{\varphi}_{1,n} + (q!)^2(1+4d_n^2(d_n-1)^2)\widetilde{\varphi}_{2,n}]$$
$$+ \frac{3}{2}\mathbf{1}\{d_n \geq 2\}(d_n-1)q!(2d_n\widetilde{\varphi}_{1,n} + q![1+2d_n(d_n-1)]\widetilde{\varphi}_{2,n})$$
$$+ 2q!\mathbf{1}\{d_n \geq 3\}(d_n-1)(d_n-2)(2d_n^2\widetilde{\varphi}_{1,n} + (q!)^2[1+4d_n^2(d_n-1)^2]\widetilde{\varphi}_{2,n}).$$

So, $\forall\, n > \check{n}_\varepsilon$, if $d_n = 1$ we have (see (8.12)) $\left|1 - q!\|(\widetilde{\sigma}^{(n)})_{11} f_1^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}\right| \le \frac{19}{4}\widetilde{\varphi}_{1,n} = \varphi_{1,n}$, and if $d_n \ge 2$ we have

$$\sup_{1\le j\le d_n}\left|1 - q!\Big\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)}\Big\|^2_{\mathfrak{H}^{\otimes q}}\right|$$
$$\le \Bigg\{\frac{19}{4} + \left(\frac{15}{4}q! + \frac{15}{2}(q!)^2\right) d_n + \left(3q! + 2(q!)^3\right) d_n^2 + (3q! + 12) d_n^3$$
$$+ 4q! d_n^4 + 24(q!)^2 d_n^5 + 8(q!)^3 d_n^6\Bigg\} \max\{\widetilde{\varphi}_{1,n}, \widetilde{\varphi}_{2,n}\} \le \varphi_{1,n},$$

where the latter inequality follows by the definition of $\varphi_{1,n}$ (see again (8.12)).
Step 3: Proof of (8.10).
For $n > \check{n}_\varepsilon$, $d_n \ge 2$ and $j, j' = 1, \dots, d_n$, $j \ne j'$, we have

$$\langle\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)}, \sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{j'k} f_k^{(n)}\rangle_{\mathfrak{H}^{\otimes q}} = \sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}(\widetilde{\sigma}^{(n)})_{j'k}\|f_k^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ \sum_{k\ne k'}^{1,d_n}(\widetilde{\sigma}^{(n)})_{jk}(\widetilde{\sigma}^{(n)})_{j'k'}\langle f_k^{(n)}, f_{k'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}$$
$$= (\widetilde{\sigma}^{(n)})_{jj}(\widetilde{\sigma}^{(n)})_{j'j}\|f_j^{(n)}\|^2_{\mathfrak{H}^{\otimes q}} + (\widetilde{\sigma}^{(n)})_{jj'}(\widetilde{\sigma}^{(n)})_{j'j'}\|f_{j'}^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ \mathbf{1}\{d_n \ge 3\}\sum_{k:\, k\ne j,\, k\ne j'}^{1,d_n}(\widetilde{\sigma}^{(n)})_{jk}(\widetilde{\sigma}^{(n)})_{j'k}\|f_k^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ (\widetilde{\sigma}^{(n)})_{j'j}\sum_{k:\, k\ne j}^{1,d_n}(\widetilde{\sigma}^{(n)})_{jk}\langle f_k^{(n)}, f_j^{(n)}\rangle_{\mathfrak{H}^{\otimes q}} + (\widetilde{\sigma}^{(n)})_{jj}(\widetilde{\sigma}^{(n)})_{j'j'}\langle f_j^{(n)}, f_{j'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}$$
$$+ (\widetilde{\sigma}^{(n)})_{jj}\mathbf{1}\{d_n \ge 3\}\sum_{k':\, k'\ne j,\, k'\ne j'}^{1,d_n}(\widetilde{\sigma}^{(n)})_{j'k'}\langle f_j^{(n)}, f_{k'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}$$
$$+ (\widetilde{\sigma}^{(n)})_{j'j'}\mathbf{1}\{d_n \ge 3\}\sum_{k:\, k\ne j',\, k\ne j}^{1,d_n}(\widetilde{\sigma}^{(n)})_{jk}\langle f_k^{(n)}, f_{j'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}$$
$$+ \mathbf{1}\{d_n \ge 3\}\sum_{k\ne k',\, k,k'\ne j\ k'\ne j'}^{1,d_n}(\widetilde{\sigma}^{(n)})_{jk}(\widetilde{\sigma}^{(n)})_{j'k'}\langle f_k^{(n)}, f_{k'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}},$$

and so, for $n > \check{n}_\varepsilon$, $d_n \ge 2$ and $j, j' = 1, \dots, d_n$, $j \ne j'$, we have

$$\left|\langle\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)}, \sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{j'k} f_k^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}\right|$$
$$\le |(\widetilde{\sigma}^{(n)})_{jj}||(\widetilde{\sigma}^{(n)})_{j'j}|\|f_j^{(n)}\|^2_{\mathfrak{H}^{\otimes q}} + |(\widetilde{\sigma}^{(n)})_{jj'}||(\widetilde{\sigma}^{(n)})_{j'j'}|\|f_{j'}^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ \mathbf{1}\{d_n \ge 3\}\sum_{k:\, k\ne j,\, k\ne j'}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{jk}||(\widetilde{\sigma}^{(n)})_{j'k}|\|f_k^{(n)}\|^2_{\mathfrak{H}^{\otimes q}}$$
$$+ |(\widetilde{\sigma}^{(n)})_{j'j}|\sum_{k:\, k\ne j}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{jk}||\langle f_k^{(n)}, f_j^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}| + |(\widetilde{\sigma}^{(n)})_{jj}|(\widetilde{\sigma}^{(n)})_{j'j'}||\langle f_j^{(n)}, f_{j'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}|$$
$$+ \mathbf{1}\{d_n \ge 3\}|(\widetilde{\sigma}^{(n)})_{jj}|\sum_{k':\, k'\ne j,\, k'\ne j'}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{j'k'}||\langle f_j^{(n)}, f_{k'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}|$$

$$+\mathbf{1}\{d_n\geq 3\}|(\widetilde{\sigma}^{(n)})_{j'j'}|\sum_{k:\,k\neq j',\,k\neq j}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{jk}||\langle f_k^{(n)},f_{j'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}|$$
$$+\mathbf{1}\{d_n\geq 3\}\sum_{k\neq k',\,k,k'\neq j\,k'\neq j'}^{1,d_n}|(\widetilde{\sigma}^{(n)})_{jk}||(\widetilde{\sigma}^{(n)})_{j'k'}||\langle f_k^{(n)},f_{k'}^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}|.$$

By this relation, (8.20), (8.21), (8.22), (8.23), (8.24) and the assumptions (2.16) and (2.17), for $n>\check{n}_\varepsilon$ and $d_n\geq 2$, it follows

$$\begin{aligned}
&\sup_{1\leq j\neq j'\leq d_n}\Big|\langle\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)},\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{j'k}f_k^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}\Big|\\
&\leq\Bigg\{\frac{9}{2q!}(2d+q![1+2d_n(d_n-1)])+\mathbf{1}\{d_n\geq 3\}\frac{3(d_n-2)}{q!}[2d_n^2+(q!)^2(1+4d_n^2(d_n-1)^2)]\\
&\qquad+(d_n-1)[2d_n^2+(q!)^2(1+4d_n^2(d_n-1)^2)]+\frac{9}{4}\\
&\qquad+\mathbf{1}\{d_n\geq 3\}\frac{3}{2}(d_n-2)(2d_n+q![1+2d_n(d_n-1)])\\
&\qquad+\mathbf{1}\{d_n\geq 3\}(d_n-1)(d_n-2)[2d_n^2+(q!)^2(1+4d_n^2(d_n-1)^2)]\Bigg\}\max\{\widetilde{\varphi}_{1,n},\widetilde{\varphi}_{2,n}\}\\
&\leq\Bigg\{27/4+(9/q!+(9q!)/2+(q!)^2)d_n+((q!)^2+12)d_n^2+(6/q!+2+3q!)d_n^3\\
&\qquad+2d_n^4+(12q!+4(q!)^2)d_n^5+4(q!)^2d_n^6\Bigg\}\max\{\widetilde{\varphi}_{1,n},\widetilde{\varphi}_{2,n}\}\leq\varphi_{2,n},
\end{aligned}$$

where the latter inequality follows by the definition of $\varphi_{2,n}$ (see (8.13)).
Step 4: Proof of (8.11).
By the definition of the contraction $\otimes_r$ and the triangular inequality, for any $n>\check{n}_\varepsilon$, if $q\geq 2$, we have

$$\begin{aligned}
&\Big\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\otimes_r\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\Big\|_{\mathfrak{H}^{\otimes 2(q-r)}}\\
&\quad\leq\sum_{k=1}^{d_n}\|f_k^{(n)}\otimes_r f_k^{(n)}\|_{\mathfrak{H}^{\otimes 2(q-r)}}(\widetilde{\sigma}^{(n)})_{jk}^2\\
&\qquad+\mathbf{1}\{d_n\geq 2\}\sum_{k\neq k'}^{1,d_n}\|f_k^{(n)}\otimes_r f_{k'}^{(n)}\|_{\mathfrak{H}^{\otimes 2(q-r)}}|(\widetilde{\sigma}^{(n)})_{jk}(\widetilde{\sigma}^{(n)})_{jk'}|,
\end{aligned}\tag{8.25}$$

for all $j=1,\dots,d_n$ and $r=1,\dots,q-1$. Note that, for $k\neq k'$, one has $\|f_k^{(n)}\otimes_r f_{k'}^{(n)}\|_{\mathfrak{H}^{\otimes 2(q-r)}}^2=\langle f_k^{(n)}\otimes_{q-r}f_k^{(n)},f_{k'}^{(n)}\otimes_{q-r}f_{k'}^{(n)}\rangle_{\mathfrak{H}^{\otimes 2r}}$ (see e.g. [20] p. 93). So by the Cauchy-Schwarz inequality $\|f_k^{(n)}\otimes_r f_{k'}^{(n)}\|_{\mathfrak{H}^{\otimes 2(q-r)}}^2\leq\|f_k^{(n)}\otimes_{q-r}f_k^{(n)}\|_{\mathfrak{H}^{\otimes 2r}}\|f_{k'}^{(n)}\otimes_{q-r}f_{k'}^{(n)}\|_{\mathfrak{H}^{\otimes 2r}}$. Therefore, combining this latter inequality with (8.25), for any $n>\check{n}_\varepsilon$, if $q\geq 2$, we have

$$\begin{aligned}
&\Big\|\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\otimes_r\sum_{k=1}^{d_n}(\widetilde{\sigma}^{(n)})_{jk}f_k^{(n)}\Big\|_{\mathfrak{H}^{\otimes 2(q-r)}}\\
&\quad\leq\|f_j^{(n)}\otimes_r f_j^{(n)}\|_{\mathfrak{H}^{\otimes 2(q-r)}}(\widetilde{\sigma}^{(n)})_{jj}^2+\mathbf{1}\{d_n\geq 2\}\sum_{k:\,k\neq j}^{1,d_n}\|f_k^{(n)}\otimes_r f_k^{(n)}\|_{\mathfrak{H}^{\otimes 2(q-r)}}(\widetilde{\sigma}^{(n)})_{jk}^2
\end{aligned}$$

$$+ 2|(\widetilde{\sigma}^{(n)})_{jj}|\|f_j^{(n)} \otimes_{q-r} f_j^{(n)}\|_{\mathfrak{H}^{\otimes 2r}}^{1/2}\mathbf{1}\{d_n \geq 2\} \sum_{k:\, k\neq j}^{1,d_n} \|f_k^{(n)} \otimes_{q-r} f_k^{(n)}\|_{\mathfrak{H}^{\otimes 2r}}^{1/2}|(\widetilde{\sigma}^{(n)})_{jk}|$$

$$+ \mathbf{1}\{d_n \geq 3\} \sum_{k\neq k',\, k,k'\neq j}^{1,d_n} \|f_k^{(n)} \otimes_{q-r} f_k^{(n)}\|_{\mathfrak{H}^{\otimes 2r}}^{1/2}\|f_{k'}^{(n)} \otimes_{q-r} f_{k'}^{(n)}\|_{\mathfrak{H}^{\otimes 2r}}^{1/2}|(\widetilde{\sigma}^{(n)})_{jk}(\widetilde{\sigma}^{(n)})_{jk'}|,$$

for all $j = 1,\dots,d_n$ and $r = 1,\dots,q-1$. By this relation, (8.20), (8.21), (8.22), (8.23), (8.24) and the assumption (2.18), $\forall\, n > \check{n}_\varepsilon$, if $q \geq 2$ we have

$$\begin{aligned}
&\sup_{1\leq j\leq d_n} \sup_{1\leq r\leq q-1} \Big\| \sum_{k=1}^{d_n} (\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)} \otimes_r \sum_{k=1}^{d_n} (\widetilde{\sigma}^{(n)})_{jk} f_k^{(n)} \Big\|_{\mathfrak{H}^{\otimes 2(q-r)}} \\
&\quad\leq \frac{9}{4}\widetilde{\varphi}_{3,n} + \mathbf{1}\{d_n \geq 2\}(d_n-1)[8d_n^2\widetilde{\varphi}_{1,n}^2 + 4(q!)^2(1+4d_n^2(d_n-1)^2)\widetilde{\varphi}_{2,n}^2]\widetilde{\varphi}_{3,n} \\
&\qquad + 3\mathbf{1}\{d_n \geq 2\}(d_n-1)\widetilde{\varphi}_{3,n}(2d_n\widetilde{\varphi}_{1,n} + q![1+2d_n(d_n-1)]\widetilde{\varphi}_{2,n}) \\
&\qquad + \mathbf{1}\{d_n \geq 3\}(d_n-1)(d_n-2)\widetilde{\varphi}_{3,n}[8d_n^2\widetilde{\varphi}_{1,n}^2 + 4(q!)^2(1+4d_n^2(d_n-1)^2)\widetilde{\varphi}_{2,n}^2] \\
&\quad\leq \Bigg\{ \frac{9}{4} + \mathbf{1}\{d_n \geq 2\}(d_n-1)[2d_n^2 + (q!)^2(1+4d_n^2(d_n-1)^2)] \\
&\qquad + 3\mathbf{1}\{d_n \geq 2\}(d_n-1)\left(d_n + \frac{1}{2}q![1+2d_n(d_n-1)]\right) \\
&\qquad + \mathbf{1}\{d_n \geq 3\}(d_n-1)(d_n-2)[2d_n^2 + (q!)^2(1+4d_n^2(d_n-1)^2)] \Bigg\}\widetilde{\varphi}_{3,n} \\
&\quad\leq \left\{ \frac{9}{4} + \left((q!)^2 + \frac{3}{2}q!\right)d_n + (3+(q!)^2)d_n^2 + (2+3q!)d_n^3 + 2d_n^4 + 4(q!)^2d_n^5 + 4(q!)^2d_n^6 \right\}\widetilde{\varphi}_{3,n} \\
&\quad\leq \varphi_{3,n},
\end{aligned}$$

where the latter inequality follows by the definition of $\varphi_{3,n}$ (see (8.14)).

Remark 8.2. By a close inspection of the proof of Theorem 2.5, one realizes that the index n_ε in (2.20) is defined by

$$n_\varepsilon := \max\{\check{n}_\varepsilon, \widetilde{n}_\varepsilon, \overline{n}_\varepsilon\}$$

where: $\check{n}_\varepsilon := \max\{n'_\varepsilon, n'', n''', n''''\}$, here the indexes n'_ε, n'' and n''' are defined in the statement of Theorem 2.5 and the index $n'''' \geq 2$ is such that $\widetilde{\varphi}_{1,n} + q!(d_n-1)\widetilde{\varphi}_{2,n} \leq 1/2$, for any $n > n''''$; $\widetilde{n}_\varepsilon := \max\{n'_\varepsilon, n^*\}$, here the index n^* is defined by (7.10) with $\varphi_{3,n}$ given by (8.14); $\overline{n}_\varepsilon$ is defined by (7.5) with c'_ε in place of c_ε.

9 Stochastic integrals of Ornstein-Uhlenbeck processes

We recall that, letting $B = \{B_t\}_{t\geq 0}$ denote a standard Brownian motion, the Ornstein-Uhlenbeck process $Y = \{Y_t\}_{t\geq 0}$ is defined as the unique solution of the Langevin equation

$$Y_t = -\frac{1}{2}\int_0^t Y_s \mathrm{d}s + B_t, \quad t \in [0,T], \tag{9.1}$$

where $T > 0$ is a finite positive horizon (see e.g. [26] for more insight into the Ornstein-Uhlenbeck process).

9.1 Approximate isometry

The following theorem holds.

Theorem 9.1. *Let $n \in \mathbb{N} \setminus \{1,2\}$, $\{B^{(i)}\}_{1\le i\le n}$ independent standard Brownian motions, $\{Y^{(i)}\}_{1\le i\le n}$ Ornstein-Uhlenbeck processes, $Y^{(i)}$ constructed via $B^{(i)}$ as in* (9.1)*, and $\{s_i\}_{1\le i\le n}$ positive numbers such that*

$$\exists\, \mathfrak{s} > 0 \text{ (not depending on } n\text{) such that } \max_{1\le i\le n} s_i > \mathfrak{s}. \tag{9.2}$$

Let

$$0 < \varepsilon < \frac{f(\delta^*)}{4\sqrt{3+2\sqrt{2}}\left(\mathrm{e}^{-3/2}\sqrt{12(3+2\sqrt{2})}+\sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right)} < 1$$

be arbitrarily fixed (here $f(\delta^*)$ is given by (2.7)*) and let $\mathbf{F}_n$ denote the $n \times d_n$, $d_n \in \mathbb{N}$, random matrix $\mathbf{F}_n := (F_{ij}^{(n)})_{1\le i\le n,\, 1\le j\le d_n}$ with*

$$F_{ij}^{(n)} := \frac{1}{\sqrt{s_i(m_n)^{j+1}}}\int_0^{s_i(m_n)^{j+1}} Y_t^{(i)}\mathrm{d}B_t^{(i)}, \quad \text{where } \{m_n\}_{n>2} \text{ is a sequence diverging to } +\infty. \tag{9.3}$$

Assume that $d_n = o(n)$, as $n \to \infty$, and that there exists an index $n_\varepsilon^\bullet \ge 2$ such that

$$\frac{d_n^4}{\sqrt{m_n}} \le \varepsilon^6 n^{3/2}\mathrm{e}^{-\frac{n\varepsilon^2}{4}}, \quad \forall\, n > n_\varepsilon^\bullet. \tag{9.4}$$

Then there exists $n_\varepsilon \ge 2$ such that

$$\begin{aligned}
&\mathbb{P}\left(1-\varepsilon \le s_{\min}\left(\frac{\mathbf{F}_n}{\sqrt{n}}\right) \le s_{\max}\left(\frac{\mathbf{F}_n}{\sqrt{n}}\right) \le 1+\varepsilon\right)\\
&\quad\ge \mathbb{P}\left((1-\varepsilon)\|\mathbf{u}\| \le \left\|\frac{1}{\sqrt{n}}\mathbf{F}_n\mathbf{u}\right\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u}\in\mathbb{R}^d\right)\\
&\quad\ge \mathbb{P}\left(\left\|\frac{1}{n}(\mathbf{F}_n)^\top\mathbf{F}_n - \mathbf{I}_d\right\|_{op} \le \varepsilon\right)\\
&\quad\ge 1-5\left(\frac{2}{\delta_1''(n,\varepsilon)}+1\right)^{d_n}\left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1''(n,\varepsilon))^2}+c_\varepsilon''\frac{n}{d_n}\mathrm{e}^{-n\varepsilon^2/8}\right), \quad \forall\, n > n_\varepsilon.
\end{aligned} \tag{9.5}$$

Here the quantities $\delta_1''(n,\varepsilon)$ and c_ε'' are defined, respectively, as $\delta_1(n,\varepsilon)$ in (2.10) *and c_ε in* (2.11) *with $Q=2$. It turns out that relation* (2.12) *holds with $\delta_1''(n,\varepsilon)$ and c_ε'' in place of $\delta_1(n,\varepsilon)$ and c_ε, respectively, and therefore the latter term in* (9.5) *tends to 1, as $n\to\infty$.*

Remark 9.2. Take ε as in the statement of Theorem 9.1 and chose m_n so that $m_n \to \infty$ and $m_n = o\left(n^{-\frac{1}{3}}\mathrm{e}^{8\varepsilon^2 n}\right)$, as $n\to\infty$. Then an easy computation shows that the "smallness" condition (9.4) implies $d_n^3 = o(n)$, as $n \to \infty$. As discussed in the Introduction and Remark 2.3, this is exactly the regime under which renormalized Wishart matrices, based on suitable classes of chaotic random variables, converge weakly to random matrices belonging to suitable Gaussian Ensembles, see [4, 6, 13, 24, 25].

Proof. (Theorem 9.1) The idea is to apply Theorem 2.2. We start noticing that $B^{(i)}$, $i = 1,\dots,n$, is an isonormal Gaussian process over the real separable Hilbert space $\mathfrak{H}_i = \mathfrak{H} = L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \mathrm{Leb})$, where $\mathcal{B}(\mathbb{R}_+)$ is the Borel σ-field on $\mathbb{R}_+$ and Leb is the Lebesgue measure, see e.g. Example 2.1.4 in [21]. Setting

$$f_{ij}^{(n)}(s,t) := \frac{1}{2\sqrt{s_i(m_n)^{j+1}}}\mathrm{e}^{-\frac{1}{2}|t-s|}\mathbf{1}_{[0,s_i(m_n)^{j+1}]}(t)\mathbf{1}_{[0,s_i(m_n)^{j+1}]}(s), \quad 1\le i\le n,\ 1\le j\le d_n,$$

by e.g. Exercise 5.4.4(8) in [21] we have

$$I_2^{B^{(i)}}(f_{ij}^{(n)}) = \frac{1}{\sqrt{s_i(m_n)^{j+1}}}\int_0^{s_i(m_n)^{j+1}} Y_t^{(i)}\mathrm{d}B_t^{(i)}.$$

We are therefore in the framework of Theorem 2.2 with $X_i = B^{(i)}$ and $q_i = 2$, $i = 1, \dots, n$. We note that $\kappa_Q = \kappa_2 = 3 + 2\sqrt{2}$.

A straightforward computation gives, for any $i = 1, \dots, n$ and $j_1, j_2 = 1, \dots, d_n$,

$$\begin{aligned}
&\langle f_{ij_1}^{(n)}, f_{ij_2}^{(n)} \rangle_{L^2(\mathrm{Leb}^2)} \\
&= \int_0^\infty \int_0^\infty f_{ij_1}^{(n)}(s,t) f_{ij_2}^{(n)}(s,t) \mathrm{d}s\mathrm{d}t \\
&= \frac{1}{2s_i\sqrt{(m_n)^{j_1+j_2+2}}} \\
&\quad \times \int_0^{s_i \min\{(m_n)^{j_1+1},(m_n)^{j_2+1}\}} \mathrm{d}t \int_0^{s_i \min\{(m_n)^{j_1+1},(m_n)^{j_2+1}\}} \mathbf{1}\{t \geq s\} \mathrm{e}^{-(t-s)} \mathrm{d}s \\
&= \frac{1}{2s_i\sqrt{(m_n)^{j_1+j_2+2}}} \int_0^{s_i \min\{(m_n)^{j_1+1},(m_n)^{j_2+1}\}} (1 - \mathrm{e}^{-t}) \mathrm{d}t \\
&= \frac{1}{2} \left(\frac{\min\{(m_n)^{j_1},(m_n)^{j_2}\}}{\sqrt{(m_n)^{j_1+j_2}}} - \frac{(1 - \mathrm{e}^{-s_i \min\{(m_n)^{j_1+1},(m_n)^{j_2+1}\}})}{s_i\sqrt{(m_n)^{j_1+j_2+2}}} \right). \qquad (9.6)
\end{aligned}$$

Therefore, (taking $j_1 = j_2 = j$) an easy computation gives

$$\begin{aligned}
\sup_{1\leq i\leq n,\, 1\leq j\leq d_n} |1 - 2\|f_{ij}^{(n)}\|^2_{L^2(\mathrm{Leb}^2)}| &= \sup_{1\leq i\leq n,\, 1\leq j\leq d_n} \frac{1}{s_i(m_n)^{j+1}} \left(1 - \mathrm{e}^{-s_i(m_n)^{j+1}}\right) \\
&\leq \frac{1}{\mathfrak{s}(m_n)^2} =: \varphi_{1,n}, \quad \forall\, n > n^\sharp \qquad (9.7)
\end{aligned}$$

where we used the assumption (9.2) and the index $n^\sharp$ is chosen so that $n^\sharp \geq 2$ and $m_n \geq 1$ $\forall\, n > n^\sharp$ (note that such an index $n^\sharp$ exists since $m_n \to \infty$, as $n \to \infty$). By (9.6), for any $i = 1, \dots, n$, $d_n \geq 2$ and $j_1 < j_2$, $j_1, j_2 \in \{1, \dots, d_n\}$, we have

$$|\langle f_{ij_1}^{(n)}, f_{ij_2}^{(n)} \rangle_{L^2(\mathrm{Leb}^2)}| \leq \frac{1}{2} \left(\frac{(m_n)^{j_1}}{\sqrt{(m_n)^{j_1+j_2}}} + \frac{(1 - \mathrm{e}^{-s_i(m_n)^{j_1+1}})}{s_i\sqrt{(m_n)^{j_1+j_2+2}}} \right) \leq \frac{(m_n)^{j_1}}{\sqrt{(m_n)^{j_1+j_2}}},$$

where the latter inequality follows by the elementary relation $1 - \mathrm{e}^{-x} \leq x$, $x \geq 0$. Consequently,

$$\forall\, n > n^\sharp, \text{ if } d_n \geq 2, \text{ then } \sup_{1\leq i\leq n} \sup_{1\leq j_1\neq j_2\leq d_n} |\langle f_{ij_1}^{(n)}, f_{ij_2}^{(n)} \rangle_{L^2(\mathrm{Leb}^2)}| \leq \frac{1}{\sqrt{m_n}} =: \varphi_{2,n}. \qquad (9.8)$$

Here the inequality follows noticing that if e.g. $j_1 < j_2$, i.e., $j_2 \geq j_1 + 1$, then the ratio $\frac{(m_n)^{j_1}}{\sqrt{(m_n)^{j_1+j_2}}}$ is less than $\frac{1}{\sqrt{m_n}}$ for $n > n^\sharp$, since for all these n it holds $m_n \geq 1$.

For any $i = 1, \dots, n$, $j = 1, \dots, d_n$, relation (A.4) yields

$$f_{ij}^{(n)} \otimes_1 f_{ij}^{(n)}(t_1, t_2) = \int_0^\infty f_{ij}^{(n)}(t_1, s) f_{ij}^{(n)}(t_2, s) \mathrm{d}s,$$

and so

$$\begin{aligned}
&\|f_{ij}^{(n)} \otimes_1 f_{ij}^{(n)}\|_{L^2(\mathrm{Leb}^2)} \\
&= \sqrt{\int_0^{s_i(m_n)^{j+1}} \int_0^{s_i(m_n)^{j+1}} \left(\int_0^{s_i(m_n)^{j+1}} f_{ij}^{(n)}(t_1, s) f_{ij}^{(n)}(t_2, s) \mathrm{d}s \right)^2 \mathrm{d}t_1 \mathrm{d}t_2} \\
&= \frac{1}{2\sqrt{2}s_i(m_n)^{j+1}} \\
&\quad \times \sqrt{\int_0^{s_i(m_n)^{j+1}} \int_0^{s_i(m_n)^{j+1}} \mathbf{1}\{t_2 \geq t_1\} \left(\int_0^{s_i(m_n)^{j+1}} \mathrm{e}^{-\frac{1}{2}(|t_1-s|+|t_2-s|)} \mathrm{d}s \right)^2 \mathrm{d}t_1 \mathrm{d}t_2}
\end{aligned}$$

$$= \frac{1}{2\sqrt{2}s_i(m_n)^{j+1}}$$
$$\times \sqrt{\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \left(\int_0^{s_i(m_n)^{j+1}} \mathrm{e}^{-\frac{1}{2}(|t_1-s|+|t_2-s|)} \mathrm{d}s \right)^2 \mathrm{d}t_1}. \tag{9.9}$$

A simple computation shows, for $0 \leq t_1 \leq t_2 \leq s_i(m_n)^{j+1}$,

$$\int_0^{s_i(m_n)^{j+1}} \mathrm{e}^{-\frac{1}{2}(|t_1-s|+|t_2-s|)} \mathrm{d}s$$
$$= \int_0^{t_1} \mathrm{e}^{-\frac{1}{2}(|t_1-s|+|t_2-s|)} \mathrm{d}s + \int_{t_1}^{t_2} \mathrm{e}^{-\frac{1}{2}(|t_1-s|+|t_2-s|)} \mathrm{d}s + \int_{t_2}^{s_i(m_n)^{j+1}} \mathrm{e}^{-\frac{1}{2}(|t_1-s|+|t_2-s|)} \mathrm{d}s$$
$$= \mathrm{e}^{-\frac{1}{2}(t_1+t_2)}(\mathrm{e}^{t_1}-1) + \mathrm{e}^{-\frac{1}{2}(t_2-t_1)}(t_2-t_1) + \mathrm{e}^{\frac{1}{2}(t_1+t_2)}(\mathrm{e}^{-t_2} - \mathrm{e}^{-s_i(m_n)^{j+1}}).$$

Inserting this expression into (9.9) and using Minkowski's inequality we have

$$\|f_{ij}^{(n)} \otimes_1 f_{ij}^{(n)}\|_{L^2(\mathrm{Leb}^2)} \leq \frac{1}{2\sqrt{2}s_i(m_n)^{j+1}} \left(\sqrt{\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{-(t_1+t_2)}(\mathrm{e}^{t_1}-1)^2 \mathrm{d}t_1} \right.$$
$$+ \sqrt{\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{-(t_2-t_1)}(t_2-t_1)^2 \mathrm{d}t_1}$$
$$\left. + \sqrt{\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{t_1+t_2}(\mathrm{e}^{-t_2} - \mathrm{e}^{-s_i(m_n)^{j+1}})^2 \mathrm{d}t_1} \right). \tag{9.10}$$

A straightforward computation yields

$$\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{-(t_1+t_2)}(\mathrm{e}^{t_1}-1)^2 \mathrm{d}t_1$$
$$= s_i(m_n)^{j+1} - \frac{5}{2} + 2s_i(m_n)^{j+1}\mathrm{e}^{-s_i(m_n)^{j+1}} + 2\mathrm{e}^{-s_i(m_n)^{j+1}} + \frac{1}{2}\mathrm{e}^{-2s_i(m_n)^{j+1}}, \tag{9.11}$$

$$\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{-(t_2-t_1)}(t_2-t_1)^2 \mathrm{d}t_1$$
$$= 2s_i(m_n)^{j+1} - 2 + (s_i(m_n)^{j+1})^2\mathrm{e}^{-s_i(m_n)^{j+1}} + 4s_i(m_n)^{j+1}\mathrm{e}^{-s_i(m_n)^{j+1}} + 2\mathrm{e}^{-s_i(m_n)^{j+1}}$$

and

$$\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{t_1+t_2}(\mathrm{e}^{-t_2} - \mathrm{e}^{-s_i(m_n)^{j+1}})^2 \mathrm{d}t_1$$
$$= s_i(m_n)^{j+1} - \frac{5}{2} + 2s_i(m_n)^{j+1}\mathrm{e}^{-s_i(m_n)^{j+1}} + 2\mathrm{e}^{-s_i(m_n)^{j+1}} + \frac{1}{2}\mathrm{e}^{-2s_i(m_n)^{j+1}}. \tag{9.12}$$

In particular, note that the quantities in (9.11) and (9.12) are equal. Therefore,

$$\frac{1}{s_i(m_n)^{j+1}} \sqrt{\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{-(t_1+t_2)}(\mathrm{e}^{t_1}-1)^2 \mathrm{d}t_1}$$
$$= \sqrt{\frac{1}{s_i(m_n)^{j+1}} - \frac{5}{2(s_i)^2(m_n)^{2(j+1)}} + \frac{2\mathrm{e}^{-s_i(m_n)^{j+1}}}{s_i(m_n)^{j+1}} + \frac{2\mathrm{e}^{-s_i(m_n)^{j+1}}}{(s_i)^2(m_n)^{2(j+1)}} + \frac{\mathrm{e}^{-2s_i(m_n)^{j+1}}}{2(s_i)^2(m_n)^{2(j+1)}}}$$
$$\leq \sqrt{\frac{3}{s_i(m_n)^{j+1}}} \leq \left(\frac{3}{\mathfrak{s}}\right)^{1/2} \frac{1}{m_n}, \quad \forall\, n > n^\sharp \tag{9.13}$$

where the latter inequality follows by the assumption (9.2) and the fact that $m_n \geq 1$ for any $n > n^\sharp$; further

$$
\begin{aligned}
&\frac{1}{s_i(m_n)^{j+1}}\sqrt{\int_0^{s_i(m_n)^{j+1}} \mathrm{d}t_2 \int_0^{t_2} \mathrm{e}^{-(t_2-t_1)}(t_2-t_1)^2 \mathrm{d}t_1} \\
&\quad = \sqrt{\frac{2}{s_i(m_n)^{j+1}} - \frac{2}{(s_i)^2(m_n)^{2(j+1)}} + \mathrm{e}^{-s_i(m_n)^{j+1}} + \frac{4\mathrm{e}^{-s_i(m_n)^{j+1}}}{s_i(m_n)^{j+1}} + \frac{2\mathrm{e}^{-s_i(m_n)^{j+1}}}{(s_i)^2(m_n)^{2(j+1)}}} \\
&\quad \leq \sqrt{\frac{6}{s_i(m_n)^{j+1}} + \mathrm{e}^{-s_i(m_n)^{j+1}}} \\
&\quad \leq \sqrt{\frac{7}{s_i(m_n)^{j+1}}} \leq \left(\frac{7}{\mathfrak{s}}\right)^{1/2}\frac{1}{m_n}, \quad \forall\, n > n^\sharp
\end{aligned} \tag{9.14}
$$

where the first inequality in (9.14) follows by the relation $\mathrm{e}^x > x$, $x \geq 0$, and the latter inequality in (9.14) is a consequence of the assumption (9.2) and the fact that $m_n \geq 1$ for any $n > n^\sharp$. Combining (9.13) and (9.14) with (9.10), we have

$$
\begin{aligned}
\sup_{1\leq i\leq n}\sup_{1\leq j\leq d_n}\|f_{ij}^{(n)}\otimes_1 f_{ij}^{(n)}\|_{L^2(\mathrm{Leb}^2)} &\leq \frac{1}{2\sqrt{2}}\left[2\left(\frac{3}{\mathfrak{s}}\right)^{1/2}+\left(\frac{7}{\mathfrak{s}}\right)^{1/2}\right]\frac{1}{m_n} \\
&\leq \left(\frac{6}{\mathfrak{s}}\right)^{1/2}\frac{1}{m_n} =: \varphi_{3,n}, \quad \forall\, n > n^\sharp.
\end{aligned} \tag{9.15}
$$

By relations (9.7), (9.8) and (9.15), we have that condition (ii) of Theorem 2.2 is satisfied with $n'' := n^\sharp$. Let n° be such that $n^\circ \geq 2$ and

$$
\frac{1}{\mathfrak{s}(m_n)^{3/2}} \leq 1 \quad \text{and} \quad \left(\frac{6}{\mathfrak{s}}\right)^{1/2}\frac{1}{\sqrt{m_n}} \leq 1, \quad \forall\, n > n^\circ \tag{9.16}
$$

(such an index n° exists since $m_n \to \infty$, as $n \to \infty$). Then

$$
\begin{aligned}
d_n^3 \max\{\varphi_{1,n}, d_n\varphi_{2,n}, d_n\varphi_{3,n}\} &= d_n^3 \max\left\{\frac{1}{\mathfrak{s}(m_n)^2}, \frac{d_n}{\sqrt{m_n}}, \left(\frac{6}{\mathfrak{s}}\right)^{1/2}\frac{d_n}{m_n}\right\} \\
&\leq d_n^4 \max\left\{\frac{1}{\mathfrak{s}(m_n)^2}, \frac{1}{\sqrt{m_n}}, \left(\frac{6}{\mathfrak{s}}\right)^{1/2}\frac{1}{m_n}\right\} \qquad (9.17) \\
&\leq \frac{d_n^4}{\sqrt{m_n}} \leq \varepsilon^6 n^{3/2}\mathrm{e}^{-\frac{n\varepsilon^2}{4}}, \quad \forall\, n > n'_\varepsilon := \max\{n_\varepsilon^\bullet, n^\circ\}
\end{aligned}
$$

where the inequality (9.17) follows by (9.16) and the latter inequality follows by the assumption (9.4). Therefore the assumption (i) of Theorem 2.2 is also satisfied. Consequently, this theorem can be applied and the claim follows. □

Remark 9.3. By Remark 7.2 and a close inspection of the proof of Theorem 9.1 we have that the index n_ε involved in the bound (9.5) can be chosen as

$$
n_\varepsilon := \max\{n^\sharp, n_\varepsilon^\bullet, n^\circ, \overline{n}_\varepsilon\}, \tag{9.18}
$$

where: $n^\sharp$ is such that $n^\sharp \geq 2$ and $m_n \geq 1\ \forall\, n > n^\sharp$; $n_\varepsilon^\bullet$ is provided by the assumption (9.4); n° is defined by (9.16) and $\overline{n}_\varepsilon$ is given by relation (7.5) with c''_ε in place of c_ε. Indeed, by Remark 7.2 and a close inspection of the proof of Theorem 9.1 we have that

$$
n_\varepsilon := \max\{n^\sharp, \widetilde{n}_\varepsilon, \overline{n}_\varepsilon\}
$$

where: $\widetilde{n}_\varepsilon := \max\{n_\varepsilon^\bullet, n^\circ, n^*\}$ and n^* is defined by (7.10) with $\varphi_{3,n} := (6/\mathfrak{s})^{1/2}m_n^{-1}$. Since if $n > \max\{n^\sharp, n^\circ\}$ then $\varphi_{3,n} \leq 1$, we finally have (9.18).

9.2 Covariance estimation

Let $\mathbf{I}_B^{(n)}(d_n,\theta)$ be the d_n-dimensional (row) random vector with j-th component defined by

$$F_j^{(n)}(\theta) := \frac{1}{\sqrt{\theta(m_n)^{j+1}}}\int_0^{\theta(m_n)^{j+1}} Y_t \mathrm{d}B_t, \quad 1 \le j \le d_n.$$

Here $n \in \mathbb{N}\setminus\{1,2\}$, $d_n \in \mathbb{N}$, $Y=\{Y_t\}_{t\ge 0}$ is the Ornstein-Uhlenbeck process defined by (9.1), θ is an unknown parameter such that $\theta \ge \mathfrak{s} > 0$, for a known positive constant $\mathfrak{s}$, and $\{m_n\}$ is a sequence diverging to $+\infty$. By (2.13) and (9.6) we have that the covariance matrix of the random vector $\mathbf{I}_B^{(n)}(d_n,\theta)$ is the $d_n \times d_n$ matrix $\mathbf{\Sigma}^{(n)}(\theta)$ with jk-entry

$$\frac{\min\{(m_n)^j,(m_n)^k\}}{\sqrt{(m_n)^{j+k}}} - \frac{1-\mathrm{e}^{-\theta\min\{(m_n)^{j+1},(m_n)^{k+1}\}}}{\theta\sqrt{(m_n)^{j+k+2}}}.$$

Let $\widehat{\mathbf{\Sigma}}^{(n)}$ be the sample covariance based on n data $\mathbf{D}_1^{(n)},\dots,\mathbf{D}_n^{(n)}$ (i.e., n independent samples from $\mathbf{I}_B^{(n)}(d_n,\theta)$) and let $R_n(\theta)$ be the relative error concerning the approximation of $\mathbf{\Sigma}^{(n)}(\theta)$ by the sample covariance $\widehat{\mathbf{\Sigma}}^{(n)}$ (see (2.14)).

The following theorem holds.

Theorem 9.4. *Let the foregoing notation and conditions prevail, and let*

$$0<\varepsilon<\frac{f(\delta^*)}{4\sqrt{3+2\sqrt{2}}\left(\mathrm{e}^{-3/2}\sqrt{12(3+2\sqrt{2})}+\sqrt{5}\left(\frac{3}{\pi}\right)^{1/4}\right)}<1$$

be arbitrarily fixed (here $f(\delta^*)$ is given by (2.7)). Assume that $d_n=o(n)$, as $n\to\infty$, and that there exists an index $n_\varepsilon^\bullet \ge 2$ such that

$$\frac{d_n^{10}}{\sqrt{m_n}} \le \frac{4\varepsilon^6}{1001} n^{3/2}\mathrm{e}^{-\frac{n\varepsilon^2}{4}}, \quad \forall\, n > n_\varepsilon^\bullet. \tag{9.19}$$

Then there exists $n_\varepsilon \ge 2$ such that

$$\mathbb{P}(R_n(\theta)\le\varepsilon) \ge 1-5\left(\frac{2}{\delta_1''(n,\varepsilon)}+1\right)^{d_n}\left(\mathrm{e}^{-\frac{n}{8}\varepsilon^2(1-2\delta_1(n,\varepsilon))^2}+c_\varepsilon''\frac{n}{d_n}\mathrm{e}^{-n\varepsilon^2/8}\right), \quad \forall\, n>n_\varepsilon \tag{9.20}$$

where $\delta_1''(n,\varepsilon)$ and c_ε'' are defined in the statement of Theorem 9.1. It turns out that relation (2.12) *holds with $\delta_1''(n,\varepsilon)$ and c_ε'' in place of $\delta_1(n,\varepsilon)$ and c_ε, respectively, and therefore the rightmost term in* (9.20) *tends to* 1*, as $n\to\infty$.*

Proof. The idea is to apply Theorem 2.5. By the computations in the proof of Theorem 9.1 we have that the condition (ii) of Theorem 2.5 is satisfied with $n'' := n^\sharp$,

$$\widetilde{\varphi}_{1,n} := \frac{1}{\mathfrak{s}(m_n)^2}, \quad \widetilde{\varphi}_{2,n} := \frac{1}{\sqrt{m_n}}, \quad \widetilde{\varphi}_{3,n} := \left(\frac{6}{\mathfrak{s}}\right)^{1/2}\frac{1}{m_n}$$

and $q=2$. Let n° be defined by (9.16). Then

$$\max\{\widetilde{\varphi}_{1,n},\widetilde{\varphi}_{2,n},\widetilde{\varphi}_{3,n}\} \le \frac{1}{\sqrt{m_n}}, \quad \forall\, n>n^\circ.$$

Using this relation and the assumption (9.19) one easily realizes that condition (i) of Theorem 2.5 (with $q=2$) holds taking $n_\varepsilon' := \max\{n^\circ, n_\varepsilon^\bullet\}$. Let $\overline{n}\ge 2$ be such that

$$m_n \ge 1 \quad \text{and} \quad 1-\frac{1}{\mathfrak{s}(m_n)^2}>0, \quad \forall\, n>\overline{n}$$

(such an index exists since $m_n \to \infty$, as $n \to \infty$). Let $n''' \geq \overline{n}$ be such that

$$\frac{d_n^2}{\sqrt{m_n}} < 1 - \frac{1}{\mathfrak{s}(m_n)^2}, \quad \forall\, n > n''' \tag{9.21}$$

(the existence of such an index n''' it follows by the fact that $d_n^2/\sqrt{m_n} \to 0$, as $n \to \infty$, which, in turn, follows by the assumption (9.19) and the fact that $d_n \geq 1$). The claim of the theorem follows by Theorem 2.5 if we check that

$$\forall\, n > n''' \text{ the matrices } \mathbf{\Sigma}^{(n)}(\theta) \text{ are strictly diagonally dominant,}$$

indeed in such a case they are positive definite, and therefore invertible. By the definition of $\mathbf{\Sigma}^{(n)}(\theta)$, we have that $\forall\, n > n'''$ this matrix is strictly diagonally dominant if and only if $\quad \forall\, n > n'''$ and $j = 1, \dots, d_n$,

$$1 - \frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\theta m_n^{j+1}} > 2 \sum_{j<k}^{1,d_n} \Big| \frac{m_n^j}{\sqrt{m_n^{j+k}}} - \frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\theta \sqrt{m_n^{j+k+2}}} \Big|. \tag{9.22}$$

To check this, note that, on one hand, since $\theta \geq \mathfrak{s} > 0$, for $n > n'''$ and $j = 1, \dots, d_n$, we have

$$1 - \frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\theta m_n^{j+1}} \geq 1 - \frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\mathfrak{s} m_n^{j+1}} > 1 - \frac{1}{\mathfrak{s} m_n^{j+1}} \geq 1 - \frac{1}{\mathfrak{s} m_n^2} > 0, \tag{9.23}$$

where the strict positivity follows by the definition of n'''. On the other hand, for $n > n'''$ and $j = 1, \dots, d_n$, we have

$$2 \sum_{j<k}^{1,d_n} \Big| \frac{m_n^j}{\sqrt{m_n^{j+k}}} - \frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\theta \sqrt{m_n^{j+k+2}}} \Big| \leq 2 \sum_{j<k}^{1,d_n} \max\left\{ \frac{m_n^j}{\sqrt{m_n^{j+k}}}, \frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\theta \sqrt{m_n^{j+k+2}}} \right\} \tag{9.24}$$

$$\leq 2 \sum_{j<k}^{1,d_n} \max\left\{ \frac{m_n^j}{\sqrt{m_n^{2j+1}}}, \frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\theta \sqrt{m_n^{2(j+1)+1}}} \right\} \tag{9.25}$$

$$\leq 2 \frac{d_n(d_n - 1)}{2} \frac{1}{\sqrt{m_n}} \tag{9.26}$$

$$\leq \frac{d_n^2}{\sqrt{m_n}} < 1 - \frac{1}{\mathfrak{s}(m_n)^2}. \tag{9.27}$$

Here: in (9.24) we used the elementary inequality $|a - b| \leq \max\{a, b\}$, $a, b \geq 0$; in (9.25) we majorized by taking $k = j + 1$ (which is the minimum possible value of k being $k > j$); in (9.26) we used that $\frac{1 - \mathrm{e}^{-\theta m_n^{j+1}}}{\theta m_n^{j+1}} \leq 1$; the strict inequality in (9.27) follows by the definition of n'''. Finally, relation (9.22) follows by (9.27) and (9.23). □

Remark 9.5. By Remark 8.2 and a close inspection of the proof of Theorem 9.4 we have that the index n_ε involved in the bound (9.20) can be chosen as

$$n_\varepsilon := \max\{n^\circ, n_\varepsilon^\bullet, n^\sharp, n''', n'''', \overline{n}_\varepsilon\},$$

where: n° is defined by (9.16); $n_\varepsilon^\bullet$ is defined by (9.19); $n^\sharp$ is such that $n^\sharp \geq 2$ and $m_n \geq 1$ $\forall\, n > n^\sharp$; n''' is such that (9.21) holds; n'''' is such that $n'''' \geq 2$ and

$$\frac{1}{\mathfrak{s}(m_n)^2} + \frac{2(d_n - 1)}{\sqrt{m_n}} \leq \frac{1}{2}, \quad \forall\, n > n''''$$

(such an index exists since $m_n \to \infty$ and $d_n/\sqrt{m_n} \to 0$, as $n \to \infty$); finally, $\overline{n}_\varepsilon$ is given by relation (7.5) with c_ε'' in place of c_ε. We omit the details of this verification.

A Elements of Malliavin calculus on the Wiener space

Hereon, we present the basic elements of Gaussian analysis and Malliavin calculus that are used in this paper, and refer the reader to [26] for any unexplained definition or result.

Let $\mathfrak{H}$ be a real separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_{\mathfrak{H}}$ and norm $\|x\|_{\mathfrak{H}} := \langle x,x\rangle_{\mathfrak{H}}^{1/2}$, $x\in\mathfrak{H}$. For $q\in\mathbb{N}$, let $\mathfrak{H}^{\otimes q}$ be the q-th tensor product of $\mathfrak{H}$ and let $\mathfrak{H}^{\odot q}$ be the associated q-th symmetric tensor product. We denote by $X=\{X(u)\}_{u\in\mathfrak{H}}$ an isonormal Gaussian process, defined on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, over $\mathfrak{H}$, i.e., X is a centered Gaussian family, defined on a complete probability space $(\Omega,\mathcal{F},\mathbb{P})$, indexed by elements of $\mathfrak{H}$ and such that, for any $u,v\in\mathfrak{H}$, $\mathbb{E}X(u)X(v)=\langle u,v\rangle_{\mathfrak{H}}$. Throughout this appendix, we take $\mathcal{F}=\sigma(X)$.

For any $q\in\mathbb{N}$, let $\mathcal{H}_q$ be the q-th Wiener chaos of X, that is, the closed linear subspace of $L^2(\Omega,\mathcal{F},\mathbb{P})$ generated by random variables of the type $\{H_q(X(u)):\ u\in\mathfrak{H},\ \|u\|_{\mathfrak{H}}=1\}$, where H_q is the q-th Hermite polynomial. By convention we write $\mathcal{H}_0=\mathbb{R}$. For any $q\in\mathbb{N}$, the mapping $I_q(u^{\otimes q})=q!H_q(X(u))$ can be extended to a linear isometry between $\mathfrak{H}^{\odot q}$, equipped with the modified norm $\sqrt{q!}\|\cdot\|_{\mathfrak{H}^{\odot q}}$ and the q-th Wiener chaos $\mathcal{H}_q$. For $q=0$, we write $I_0(c)=c$, $c\in\mathbb{R}$. In particular, the following isometry formula holds for any $f\in\mathfrak{H}^{\odot p}$ and $g\in\mathfrak{H}^{\odot q}$:

$$\mathbb{E}I_p(f)I_q(g)=\mathbf{1}\{p=q\}q!\langle f,g\rangle_{\mathfrak{H}^{\odot q}}. \tag{A.1}$$

It is well-known that any random variable $F\in L^2(\Omega,\mathcal{F},\mathbb{P})$ admits a unique chaotic expansion

$$F=\sum_{q=0}^{\infty} I_q(f_q), \tag{A.2}$$

Here, the series (A.2) converges in $L^2(\Omega,\mathcal{F},\mathbb{P})$, $I_0(f_0)=\mathbb{E}F$ and the $f_q\in\mathfrak{H}^{\odot q}$, $q\in\mathbb{N}$, are uniquely determined by F. For any $q\in\mathbb{N}\cup\{0\}$, we denote by J_q the projection operator on the q-th Wiener chaos. In particular, if F is as in (A.2), then $J_q(F)=I_q(f_q)$, $q\in\mathbb{N}\cup\{0\}$.

Let $\{e_k\}_{k\geq 1}$ be a complete orthonormal system in $\mathfrak{H}$. Given $f\in\mathfrak{H}^{\odot p}$ and $g\in\mathfrak{H}^{\odot q}$, for any $r=0,\dots,\min\{p,q\}$, $p,q\in\mathbb{N}$, we define the contraction of f and g of order r as the element of $\mathfrak{H}^{\otimes(p+q-2r)}$ given by

$$f\otimes_r g=\sum_{i_1,\dots,i_r=1}^{\infty}\langle f,e_{i_1}\otimes\cdots\otimes e_{i_r}\rangle_{\mathfrak{H}^{\otimes r}}\otimes\langle g,e_{i_1}\otimes\cdots\otimes e_{i_r}\rangle_{\mathfrak{H}^{\otimes r}}. \tag{A.3}$$

For $r=0$, $f\otimes_0 g:=f\otimes g$, i.e., the tensor product of f and g, and, for $p=q=r$, $f\otimes_q g=\langle f,g\rangle_{\mathfrak{H}^{\otimes q}}$.

We remark that if $\mathfrak{H}=L^2(A,\mathcal{A},\alpha)$, with α a non-atomic measure, then $\mathfrak{H}^{\otimes q}=L^2(A^q,\mathcal{A}^q,\alpha^q)$, $\mathfrak{H}^{\odot q}=L^2_s(A^q,\mathcal{A}^q,\alpha^q)$, $q\geq 1$, where $L^2_s(A^q,\mathcal{A}^q,\alpha^q)$ denotes the set of symmetric functions in $L^2(A^q,\mathcal{A}^q,\alpha^q)$. Moreover, for any $r=0,\dots,\min\{p,q\}$, $p,q\in\mathbb{N}$, $f\in L^2_s(A^p,\mathcal{A}^p,\alpha^p)$ and $g\in L^2_s(A^q,\mathcal{A}^q,\alpha^q)$, $f\otimes_r g$ is the function of $L^2(A^{p+q-2r},\mathcal{A}^{p+q-2r},\alpha^{p+q-2r})$ given by

$$\begin{aligned}&f\otimes_r g(t_1,\dots,t_{p-r},t_{p-r+1},\dots,t_{p+q-2r})\\&\quad:=\int_{A^r} f(t_1,\dots,t_{p-r},s_1,\dots,s_r)g(t_{p-r+1},\dots,t_{p+q-2r},s_1,\dots,s_r)\alpha^r(\mathrm{d}s_1,\dots,\mathrm{d}s_r).\end{aligned} \tag{A.4}$$

Let $\mathcal{S}$ be the set of all smooth cylindrical r.v.'s of the form $F=g(X(u_1),\dots,X(u_n))$, where $n\in\mathbb{N}$, $g:\mathbb{R}^n\to\mathbb{R}$ is an infinitely differentiable function with compact support and $u_i \in \mathfrak{H}$, for any $i = 1, \ldots, n$. We define the Malliavin derivative of F as the random element of $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathfrak{H})$:

$$D^1F = DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(X(u_1), \ldots, X(u_n))u_i.$$

By iteration, we define

$$D^2F = \sum_{i,j=1}^{n} \frac{\partial^2 g}{\partial x_i \partial x_j}(X(u_1), \ldots, X(u_n))u_i \otimes u_j,$$

which is a random element of $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathfrak{H}^{\circ 2})$. We denote by $\mathbb{D}^{1,2}$ the closure of $\mathcal{S}$ w.r.t. the norm $\|\cdot\|_{1,2}$ defined by

$$\|F\|_{1,2}^2 := \mathbb{E}F^2 + \mathbb{E}\|DF\|_{\mathfrak{H}}^2.$$

The next integration by parts formula (see e.g. [21] p. 53) plays a fundamental role in the theory. For any $F \in \mathbb{D}^{1,2}$ centered we have

$$\mathbb{E}Fg(F) = \mathbb{E}g'(F)\langle -DL^{-1}F, DF\rangle_{\mathcal{H}}, \quad g \in \mathcal{C}_b^1. \tag{A.5}$$

where $\mathcal{C}_b^1$ is defined by (3.3) and L^{-1} is the inverse of the Ornstein-Uhlenbeck operator, which is defined as follows. For any $F \in L^2(\mathbb{P})$ centered with chaotic expansion (A.2), we set

$$L^{-1}F := -\sum_{q\geq 1} \frac{1}{q} I_q(f_q). \tag{A.6}$$

Note also that,

$$\text{For any } f \in \mathfrak{H}^{\circ q},\ q \in \mathbb{N},\ DI_q(f) = qI_{q-1}(f). \tag{A.7}$$

B Proof of Proposition 2.7

We distinguish the cases $d = 1$ and $d \geq 2$.

Case $d = 1$. We first suppose $q \geq 2$. By the isometry formula (A.1) and the assumptions (2.21) and (2.22) it follows that, as $m \to \infty$,

$$\mathbb{E}I_q^X(g_1^{(m)})^2 = q!\|g_1^{(m)}\|_{\mathfrak{H}^{\otimes q}}^2 \to 1, \qquad \|g_1^{(m)} \otimes_r g_1^{(m)}\|_{\mathfrak{H}^{\otimes 2(q-r)}} \to 0, \quad \forall\, r = 1, \ldots, q-1.$$

The claim follows by Theorem 5.2.7 (Fourth-moment theorem) in [21]. Now assume $q = 1$. By formulas (A.6) and (A.7) we have $\langle DI_1^X(g_1^{(m)}), -DL^{-1}I_1^X(g_1^{(m)})\rangle_{\mathfrak{H}} = \|g_1^{(m)}\|_{\mathfrak{H}}^2$. The claim then follows by the assumption (2.21) and formula (3.35) of Theorem 3.1 in [20].

Case $d \geq 2$. We preliminary note that the isometry formula (A.1) and the assumption (2.21) yield, for any $j, k = 1, \ldots, d$ and $q \in \mathbb{N}$,

$$\mathbb{E}I_q^X(g_j^{(m)})I_q^X(g_k^{(m)}) = q!\langle g_j^{(m)}, g_k^{(m)}\rangle_{\mathfrak{H}^{\otimes q}} \to \delta_{jk}, \quad \text{as } m \to \infty.$$

We continue the proof distinguishing two cases: $q \geq 2$ and $q = 1$. We first suppose $q \geq 2$. Taking $j = k$ in the assumption (2.21) and reasoning exactly as in the Case $d = 1$ (with $q \geq 2$), we have that, for any $j = 1, \ldots, d$, $I_q^X(g_j^{(m)})$ converges in law to a random variable distributed according to the standard Gaussian law, as $m \to \infty$. The claim then follows by e.g. Proposition 3.9 in [23]. Now assume $q = 1$. Taking $j = k$ in the assumption (2.21) and arguing as in the Case $d = 1$ (with $q = 1$) we have that, for any $j = 1, \ldots, d$, $I_1^X(g_j^{(m)})$ converges in distribution to a random variable distributed according to the standard Gaussian law. The claim then follows by e.g. Proposition 3.9 in [23].

C Proof of the measurability of the event $\mathcal{E}$ in (4.5)

Without loss of generality, we set

$$(\Omega, \mathcal{F}, \mathbb{P}) := (\mathbb{R}^{n\times d}, \mathcal{B}(\mathbb{R}^{n\times d}), \mathbb{P}_{\mathbf{F}_{n,d}}),$$

where $\mathcal{B}(\mathbb{R}^{n\times d})$ is the Borel σ-field on $\mathbb{R}^{n\times d}$ and $\mathbb{P}_{\mathbf{F}_{n,d}}$ is the law of $\mathbf{F}_{n,d}$, and define $\mathbf{F}_{n,d}((y_{ij})) := (y_{ij})$, for any $(y_{ij}) \in \mathbb{R}^{n\times d}$. We have to check that

$$\{(y_{ij}) \in \mathbb{R}^{n\times d} : \ (1-\varepsilon)\|\mathbf{u}\| \le \|((y_{ij})/\sqrt{n})\mathbf{u}\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d\} \in \mathcal{B}(\mathbb{R}^{n\times d}).$$

To this aim, note that

$$\begin{aligned}
&\{(y_{ij}) \in \mathbb{R}^{n\times d} : \ (1-\varepsilon)\|\mathbf{u}\| \le \|((y_{ij})/\sqrt{n})\mathbf{u}\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d\} \\
&= \{(y_{ij}) \in \mathbb{R}^{n\times d} : \ (1-\varepsilon)\|\mathbf{u}\| \le \|((y_{ij})/\sqrt{n})\mathbf{u}\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d \setminus \{\mathbf{0}\}\} \\
&= \bigcap_{\mathbf{u}\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \Psi_{\mathbf{u}}^{-1}([1-\varepsilon, 1+\varepsilon])\},
\end{aligned}$$

where

$$\text{for any } \mathbf{u} \in \mathbb{R}^d \setminus \{\mathbf{0}\},\ \Psi_{\mathbf{u}}((y_{ij})) := \left\| \frac{(y_{ij})}{\sqrt{n}} \frac{\mathbf{u}}{\|\mathbf{u}\|} \right\|.$$

Note that, for any $\mathbf{u} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, the function $\Psi_{\mathbf{u}}(\cdot)$ is continuous, therefore, for any $\mathbf{u} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, $\Psi_{\mathbf{u}}^{-1}([1-\varepsilon, 1+\varepsilon])$ is a closed set and so $\bigcap_{\mathbf{u}\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \Psi_{\mathbf{u}}^{-1}([1-\varepsilon, 1+\varepsilon])\}$ is a closed set, which implies the Borel measurability of

$$\{(y_{ij}) : \ (1-\varepsilon)\|\mathbf{u}\| \le \|((y_{ij})/\sqrt{n})\mathbf{u}\| \le (1+\varepsilon)\|\mathbf{u}\|, \quad \text{for all } \mathbf{u} \in \mathbb{R}^d\}.$$

D Proof of Lemma 4.6

By e.g. Example 2.11 p. 29 in [43] we have the concentration inequality

$$\mathbb{P}\left(\left| \frac{Z_n}{n} - 1 \right| \ge \varepsilon \right) \le 2\mathrm{e}^{-\frac{n\varepsilon^2}{8}}.$$

The claim follows noticing that

$$\mathbb{P}(Z_n < (1-\varepsilon)n) + \mathbb{P}(Z_n > (1+\varepsilon)n) + \mathbb{P}\left(\left| \frac{Z_n}{n} - 1 \right| < \varepsilon \right) = 1.$$

E Proof of Lemma 8.1

By the Spectral Decomposition Theorem (see e.g. [15]) there exists an orthonormal basis of $\mathbb{R}^{d_n}$ formed by (row) eigenvectors of $\mathbf{A}_n$, say $\mathbf{v}_\ell(\mathbf{A}_n)$, $\ell = 1, \ldots, d_n$, with corresponding (positive) eigenvalues $\lambda_\ell(\mathbf{A}_n)$ such that

$$\mathbf{A}_n = \sum_{\ell=1}^{d_n} \lambda_\ell(\mathbf{A}_n) \mathbf{v}_\ell(\mathbf{A}_n)^\top \mathbf{v}_\ell(\mathbf{A}_n), \tag{E.1}$$

i.e., $\mathbf{A}_n = \mathbf{O}_n^\top \mathbf{\Lambda}_n \mathbf{O}_n$, where $\mathbf{\Lambda}_n$ is the diagonal $d_n \times d_n$ matrix with diagonal elements $\lambda_1(\mathbf{A}_n), \ldots, \lambda_{d_n}(\mathbf{A}_n)$, and $\mathbf{O}_n$ is the orthogonal $d_n \times d_n$ matrix with ℓ-th row equal to $\mathbf{v}_\ell(\mathbf{A}_n)$, $\ell = 1, \ldots, d_n$. Therefore, by the orthogonality of $\mathbf{O}_n$ it follows $\mathbf{A}_n^{-1} = \mathbf{O}_n^{-1} \mathbf{\Lambda}_n^{-1} \mathbf{O}_n$, and, by the definition of square root of a matrix we have $\mathbf{A}_n^{-1/2} = \mathbf{O}_n^{-1} \mathbf{\Lambda}_n^{-1/2} \mathbf{O}_n$, i.e.,

$$\mathbf{A}_n^{-1/2} = \sum_{\ell=1}^{d_n} \lambda_\ell(\mathbf{A}_n)^{-1/2} \mathbf{v}_\ell(\mathbf{A}_n)^\top \mathbf{v}_\ell(\mathbf{A}_n). \tag{E.2}$$

We continue dividing the proof in three steps. In the first step we prove (8.3), in the second step we prove (8.4) and in the third step we prove (8.5) and (8.6).

Step 1: Proof of (8.3).

By (E.2) (using an obvious notation) it follows

$$\begin{aligned}\widetilde{a}_{jj}(n)-1&=\sum_{\ell=1}^{d_n}\lambda_\ell(\mathbf{A}_n)^{-1/2}(\mathbf{v}_\ell(\mathbf{A}_n)^\top\mathbf{v}_\ell(\mathbf{A}_n))_{jj}-1\\&=\sum_{\ell=1}^{d_n}(\lambda_\ell(\mathbf{A}_n)^{-1/2}-1)(\mathbf{v}_\ell(\mathbf{A}_n))_j^2,\quad j=1,\dots,d_n,\end{aligned}\tag{E.3}$$

where we used that the diagonal elements of the matrix $\mathbf{v}_\ell(\mathbf{A}_n)^\top\mathbf{v}_\ell(\mathbf{A}_n)$ are given by $(\mathbf{v}_\ell(\mathbf{A}_n))_1^2,\dots,(\mathbf{v}_\ell(\mathbf{A}_n))_{d_n}^2$, and that their sum is equal to 1 since $\{\mathbf{v}_\ell(\mathbf{A}_n)\}_{\ell=1,\dots,d}$ is an orthonormal basis of $\mathbb{R}^{d_n}$. By Gershgorin's Circle Theorem (see e.g. [15]) we have that, for each $\ell\in\{1,\dots,d_n\}$ there exists $j_\ell\in\{1,\dots,d_n\}$ such that

$$|\lambda_\ell(\mathbf{A}_n)-a_{j_\ell j_\ell}(n)|\le\mathbf{1}\{d_n\ge 2\}\sum_{k:\,k\neq j_\ell}^{1,d_n}|a_{j_\ell k}(n)|.$$

By this inequality and the assumption (8.2), we have

$$|\lambda_\ell(\mathbf{A}_n)-a_{j_\ell j_\ell}(n)|\le(d_n-1)c_n,\quad\text{for every } n,d_n\in\mathbb{N}\text{ and }\ell=1,\dots,d_n.$$

In turn, by this relation and the elementary inequality $|a-b|\ge|a|-|b|$, $a,b\in\mathbb{R}$, it follows $|\lambda_\ell(\mathbf{A}_n)-1|-|a_{j_\ell j_\ell}(n)-1|\le(d_n-1)c_n$. Combining this inequality with the assumption (8.1), we have

$$|\lambda_\ell(\mathbf{A}_n)-1|\le b_n+(d_n-1)c_n,\quad\forall\, n,d_n\in\mathbb{N}\text{ and }\ell=1,\dots,d_n.\tag{E.4}$$

So, for any $n>\widetilde{n}$ and $\ell=1,\dots,d_n$,

$$\frac{1}{\sqrt{1+b_n+(d_n-1)c_n}}-1\le\lambda_\ell(\mathbf{A}_n)^{-1/2}-1\le\frac{1}{\sqrt{1-b_n-(d_n-1)c_n}}-1.\tag{E.5}$$

Using the elementary inequality $|c|\le\max\{|a|,|b|\}$, for any $a,b,c\in\mathbb{R}$ such that $a\le c\le b$, it follows that (E.5) implies

$$|\lambda_\ell(\mathbf{A}_n)^{-1/2}-1|\le\max\left\{\Big|\frac{1}{\sqrt{1+b_n+(d_n-1)c_n}}-1\Big|,\Big|\frac{1}{\sqrt{1-b_n-(d_n-1)c_n}}-1\Big|\right\},\tag{E.6}$$

for any $n>\widetilde{n}$ and $\ell=1,\dots,d_n$. Combining this latter relation with (E.3) (and the fact that the sum of the square of the components of the orthonormal eigenvectors $\mathbf{v}_\ell(\mathbf{A}_n)$ is equal to 1), we finally have (8.3).

Step 2: Proof of (8.4).

Let $d_n\ge 2$ and $j,k=1,\dots,d_n$, $j\neq k$. By (E.2) (using an obvious notation) we have

$$\begin{aligned}\widetilde{a}_{jk}(n)&=\sum_{\ell=1}^{d_n}\lambda_\ell(\mathbf{A}_n)^{-1/2}(\mathbf{v}_\ell(\mathbf{A}_n)^\top\mathbf{v}_\ell(\mathbf{A}_n))_{jk}\\&=\sum_{\ell=1}^{d_n}(\lambda_\ell(\mathbf{A}_n)^{-1/2}-1)(\mathbf{v}_\ell(\mathbf{A}_n)^\top\mathbf{v}_\ell(\mathbf{A}_n))_{jk}+\sum_{\ell=1}^{d_n}(\mathbf{v}_\ell(\mathbf{A}_n)^\top\mathbf{v}_\ell(\mathbf{A}_n))_{jk}.\end{aligned}$$

By this relation and (E.6), since the absolute values of the components of the orthonormal eigenvectors $\mathbf{v}_\ell(\mathbf{A}_n)$ are less than or equal to 1, for any $n > \widetilde{n}$ and $j \neq k$, we have

$$|\widetilde{a}_{jk}(n)| \leq d_n \max\left\{\Big|\frac{1}{\sqrt{1+b_n+(d_n-1)c_n}}-1\Big|, \Big|\frac{1}{\sqrt{1-b_n-(d_n-1)c_n}}-1\Big|\right\} + \Big|\sum_{\ell=1}^{d_n}(\mathbf{v}_\ell(\mathbf{A}_n)^\top \mathbf{v}_\ell(\mathbf{A}_n))_{jk}\Big|. \tag{E.7}$$

By (E.1) we have

$$a_{jk}(n) = \sum_{\ell=1}^{d_n}(\lambda_\ell(\mathbf{A}_n)-1)(\mathbf{v}_\ell(\mathbf{A}_n)^\top \mathbf{v}_\ell(\mathbf{A}_n))_{jk} + \sum_{\ell=1}^{d_n}(\mathbf{v}_\ell(\mathbf{A}_n)^\top \mathbf{v}_\ell(\mathbf{A}_n))_{jk},$$

and so

$$\begin{aligned}\Big|\sum_{\ell=1}^{d_n}(\mathbf{v}_\ell(\mathbf{A}_n)^\top \mathbf{v}_\ell(\mathbf{A}_n))_{jk}\Big| &\leq |a_{jk}(n)| + \Big|\sum_{\ell=1}^{d_n}(\lambda_\ell(\mathbf{A}_n)-1)(\mathbf{v}_\ell(\mathbf{A}_n)^\top \mathbf{v}_\ell(\mathbf{A}_n))_{jk}\Big| \\ &\leq c_n + d_n(b_n+(d_n-1)c_n) = d_n b_n + (1+d_n(d_n-1))c_n, \end{aligned} \tag{E.8}$$

where the latter inequality follows by the assumption (8.2), (E.4) and again the fact that the absolute values of the components of the orthonormal eigenvectors $\mathbf{v}_\ell(\mathbf{A}_n)$ are less than or equal to 1. By (E.8) and (E.7) we finally have (8.4).

Step 3: Proof of (8.5) *and* (8.6).

Note that, setting $a_n := b_n + (d_n-1)c_n$, it holds

$$\Big|\frac{1}{\sqrt{1+a_n}}-1\Big| = \frac{a_n}{\sqrt{1+a_n}(1+\sqrt{1+a_n})} \leq a_n, \quad \forall\, n \in \mathbb{N}$$

and

$$\Big|\frac{1}{\sqrt{1-a_n}}-1\Big| = \frac{a_n}{\sqrt{1-a_n}(1+\sqrt{1-a_n})} \leq a_n, \quad \forall\, n > \widetilde{n}$$

where this latter inequality is a consequence of the choice of $\widetilde{n}$. The claims then follow combining these relations with (8.3) and (8.4).

References

[1] Bai, Z. and Zhou, W. (2008). Large Sample Covariance Matrices Without Independence Structures in Columns. *Statistica Sinica*, **18** 425–442. MR2411613

[2] Bao, Z. (2012). Strong Convergence of ESD for the Generalized Sample Covariance Matrices When $p/n \to 0$. *Statistics & Probability Letters*, **82** 894–901. MR2910035

[3] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley and Sons, New York, US. MR0233396

[4] Bourguin, S., Diez, C.P. and Tudor, C.A. (2021). Limiting Behavior of Large Correlated Wishart Matrices with Chaotic Entries. *Bernoulli*, **27** 1077–1102. MR4255227

[5] Bourguin, S. and Dhang, T. (2022). High Dimensional Regimes of Non-Stationary Gaussian Correlated Wishart Matrices. *Random Matrices. Theory and Applications*, **1**, Paper No. 2250006, 43 pp. MR4391829

[6] Bubeck, S., Ding, J., Eldan, R. and Rácz, M.Z. (2016). Testing High-Dimensional Geometry in Random Graphs. *Random Structures and Algorithms*, **49** 503–532. MR3545825

[7] Bubeck, S. and Ganguly, S. (2018). Entropic CLT and Phase Transition in High-Dimensional Wishart Matrices. *International Mathematics Research Notices*, **2** 588–606. MR3801440

[8] Clausel, M., Roueff, F., Taqqu, M.S. and Tudor, C. (2014). Asymptotic Behavior of the Quadratic Variation of the Sum of Two Hermite Processes of Consecutive Orders. *Stochastic Processes and their Applications*, **124** 2517–2541. MR3192505

[9] Chronopoulou, A., Tudor, C.A. and Viens, F.G. (2011). Self-Similarity Parameter Estimation and Reproduction Property for Non-Gaussian Hermite Processes. *Communications on Stochastic Analysis*, **5**, 161–185. MR2808541

[10] Flint, I. and Torrisi, G.L. (2023). An Integration by Parts Formula for Functionals of the Dirichlet-Ferguson Measure, and Applications. *Potential Analysis*, 58 703–730. MR4568879

[11] Gaunt, E.R., Pickett, A. and Reinert, G. (2017). Chi-Square Approximation by Stein's Method with Application to Pearson's Statistic. *The Annals of Applied Probability*, **27** 720–756. MR3655852

[12] Gaunt, E.R. and Li, S. (2023). Bounding Kolmogorov Distances trhough Wasserstein and Related Probability Metrics. *Journal of Mathematical Analysis and Applications*, **522** 1–24. MR4533915

[13] Jiang, T. and Li, D. (2015). Approximation of Rectangular Beta-Laguerre Ensembles and Large Deviations. *Journal of Theoretical Probability*, **28** 804–847. MR3413957

[14] Johnstone, I.M. (2006). *High Dimensional Statistical Inference and Random Matrices.* Proceedings of the International Congress of Mathematicians, Madrid, Spain. MR2334195

[15] Lancaster, P. and Tismenetsky, M. (1985). *The Theory of Matrices.* Academic Press, New York, US. MR0792300

[16] Last, G., Peccati, G. and Schulte, M. (2016). Normal Approximation on Poisson spaces: Mehler's Formula, Second Order Poincar é Inequalities and Stabilization. *Probability Theory and Related Fields*, **165** 667–723. MR3520016

[17] Luk, H. (1994). *Stein's Method for the Gamma Distribution and Related Statistical Applications.* PhD thesis, University of Southern California, US. MR2693204

[18] Krokowski, K., Reichenbachs, A. and Thaele, C. (2017). Discrete Malliavin-Stein Method: Berry-Esseen Bounds for Random Graphs and Percolation. *The Annals of Probabability*, **45** 1071–1109. MR3630293

[19] Marchenko, V.A. and Pastur, L.A. (1967). Distribution of Eigenvalues for some Sets of Random Matrices. *Mathematics of the USSR-Sbornik*, **72** 507–536. MR0208649

[20] Nourdin, I. and Peccati, G. (2009). Stein's Method on Wiener Chaos. *Probability Theory and Related Fields*, **145** 75–118. MR2520122

[21] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus.* Cambridge University Press, Cambridge, UK. MR2962301

[22] Nourdin, I., Peccati, G. and Reinert, G. (2010). Stein's Method and Stochastic Analysis of Rademacher Functionals. *Electronic Journal of Probability*, **15** 1703–1742. MR2735379

[23] Nourdin, I., Peccati, G. and Réveillac, A. (2010). Multivariate Normal Approximation using Stein's Method and Malliavin Calculus. *Annales de l'Institut Henri Poincaré Probabilité et Statistiques*, **46** 45–58. MR2641769

[24] Nourdin, I. and Pu, F. (2022). Gaussian Fluctuation for Gaussian Wishart Matrices of Overall Correlation. *Statistics and Probability Letters*, **181** 11 pp. MR4329126

[25] Nourdin, I. and Zheng, G. (2022). Asymptotic Behavior of Large Gaussian Correlated Wishart Matrices. *Journal of Theoretical Probability*, **35** 2239–2268. MR4509068

[26] Nualart, D. (2006). *The Malliavin Calculus and Related Topics.* Springer, Heidelberg, Germany. MR2200233

[27] Nualart, D. and Peccati, G. (2005). Central Limit Theorems for Sequences of Multiple Stochastic Integrals. *he Annals of Probability*, **33** 177–193. MR2118863

[28] Nualart, D. and Ortiz-Latorre, S. (2008). Central Limit Theorem for Multiple Stochastic Integrals and Malliavin Calculus. *Stochastic Processes and Their Applications*, **118** 614–628. MR2394845

[29] Pajor, A. and Pastur L. (2009). On the Limiting Empirical Measure of Eigenvalues of the Sum of Rank One Matrices with Log-Concave Distribution. *Studia Mathematica*, **195** 11–29. MR2539559

[30] Peccati, G., Solé, J.L., Taqqu, M. S. and Utzet, F. (2010). Stein's Method and Normal approximation of Poisson Functionals. *The Annals of Probability*, **38** 443–478. MR2642882

[31] Peccati, G. and Tudor, C.A. (2005). Gaussian Limits for Vector-Valued Multiple Stochastic Integrals. *Seminaire de Probabilités*, **XXXVIII** 247–262. MR2126978

[32] Pipiras, V. and Taqqu, M.S. (2017). *Long-range Dependence and Self-Similarity.* Cambridge Series in Statistical and Probabilistic Mathematics, [45], Cambridge University Press, Cambridge, UK. MR3729426

[33] Privault, N. and Torrisi, G.L. (2013). Probability Approximation by Clark-Ocone Covariance Representation. *Electronic Journal of Probability*, **18** 25 pp. MR3126574

[34] Privault, N. and Torrisi, G. L. (2015). The Stein and Chen–Stein methods for Functionals of Non-Symmetric Bernoulli Processes. *ALEA: Latin American Journal of Probability and Mathematical Statistics*, **12** 309–356. MR3357130

[35] Qiu, J., Li, Z. and Yao, J. (2023). Asymptotic Normality for Eigenvalue Statistics of a General Sample Covariance Matrix When $p/n \to \infty$ and Applications. *Annals of Statistics*, **51** 1427–1451. MR4630955

[36] Rácz, M.Z. and Richey, J. (2019). A Smooth Transition from Wishart to GOE. *Journal of Theoretical Probability*, **32** 898–906. MR3959632

[37] Torrisi, G.L. (2016). Gaussian Approximation of Nonlinear Hawkes Processes. *The Annals of Applied Probability*, **26** 2106–2140. MR3543891

[38] Torrisi, G.L. (2017). Probability Approximation of Point Processes with Papangelou Conditional Intensity. *Bernoulli*, **23** 2210–2256. MR3648030

[39] Torrisi, G.L. (2017). Poisson Approximation of Point Processes with Stochastic Intensity, and Application to Nonlinear Hawkes Processes. *Annales de l'Institut Henri Poincaré Probabilité et Statistiques*, **53** 679–700. MR3634270

[40] Torrisi, G.L. (2023). Quantitative Multidimensional Central Limit Theorems for Means of the Dirichlet-Ferguson Measure. *ALEA: Latin American Journal of Probability and Mathematical Statistics*, **20** 825–860. MR4615655

[41] Tudor, C.A. (2013). *Analysis of Variations for Self-Similar Processes.* Probability and its Applications, Springer, New York, US. MR3112799

[42] Vershynin, R. (2018) *High-Dimensional Probability.* Cambridge University Press, Cambridge, UK. MR3837109

[43] Wainwright, M.J. (2019). *High-Dimensional Statistics.* Cambridge University Press, Cambridge, UK. MR3967104

[44] Wishart, J. (1928). The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population. *Biometrika*, **20A** 32–52.

[45] Yaskov, P. (2016). Necessary and Sufficient Conditions for the Marchenko-Pastur Theorem. *Electronic Communications in Probability*, **21** 1–8. MR3568347