Fluctuations of the traces of
complex-valued iid random matrices

by Salim Noreddine*
Université Paris 6

This version: June 15th, 2011

Abstract: In this paper, we extend to the complex-valued case a recent result by Nourdin and Peccati on the gaussian fluctuations of traces of i.i.d. matrices.

Keywords: Central limit theorems; Nualart-Peccati criterion of asymptotic normality; Invariance principles; i.i.d. random matrices; Normal approximation.

1 Introduction

Let X be a centered random variable with unit variance, taking its values in $\mathbb{C}$ and admitting moments of all orders. Let $\{X_{i,j}\}_{i,j\geqslant 1}$ be a family of independent and identically distributed copies of X. We denote by X_n the random matrix defined as

$$X_n = \left\{ \frac{X_{i,j}}{\sqrt{n}} \right\}_{1\leqslant i,j\leqslant n}.$$

In this paper, we aim to find the limit (in law) of

$$\text{trace}(X_n^d) - E\big[\text{trace}(X_n^d)\big], \tag{1.1}$$

where X_n^d denotes the d-th power of X_n. To achieve this goal, we will make use of the following identity:

$$\text{trace}(X_n^d) = n^{-\frac{d}{2}} \sum_{i_1,\ldots,i_d=1}^{n} X_{i_1,i_2} X_{i_2,i_3} \ldots X_{i_d,i_1}.$$

In the case where X is real-valued, the problem was solved by Nourdin and Peccati [1]. The present study extends [1] to the more general case of a complex-valued random variable X.

When $d = 1$, the expression (1.1) is very simple; we have indeed

$$\text{trace}(X_n) - E\big[\text{trace}(X_n)\big] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{i,i}.$$

As a consequence, if X is real-valued then a straightforward application of the standard cental limit theorem (CLT) yields the convergence in law to $N(0,1)$. The case where X

*Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 6, Boîte courrier 188, 4 Place Jussieu, 75252 Paris Cedex 5, France. Email: `salim.noreddine@polytechnique.org`

is complex-valued is not much more difficult, as one only needs to use the bidimensional CLT to get the convergence in law to $Z = Z_1 + iZ_2$, where (Z_1, Z_2) is a gaussian vector with the same covariance matrix as that of $(\mathrm{Re}(X), \mathrm{Im}(X))$.

When $d \geqslant 2$, we have:

$$\mathrm{trace}(X_n^d) - E\big[\mathrm{trace}(X_n^d)\big] = n^{-\frac{d}{2}} \sum_{i_1,\dots,i_d=1}^{n} \big(X_{i_1,i_2}....X_{i_d,i_1} - E[X_{i_1,i_2}....X_{i_d,i_1}]\big). \qquad (1.2)$$

If X is real-valued, it is shown in [1] that there is convergence in law of (1.2) to the centered normal law with variance d. The idea behind the proof is to separate the sum in the right-hand side of (1.2) into two parts: a first part consisting of the sum over the diagonal terms, i.e. the terms with indices $i_1, \dots, i_d$ such that there is at least two distinct integers p and q satisfying $(i_p, i_{p+1}) = (i_q, i_{q+1})$; and a second part consisting of the sum over non-diagonal terms, i.e. the sum over the remaining indices. Using combinatorial arguments, it is possible to show that the sum over diagonal terms converges to 0 in L^2. Thus, the contribution to the limit comes from the non-diagonal terms only. In order to tackle the corresponding sum, the idea [1] is to focus first on the particular case where the entries $X_{i,j}$ are gaussian. Indeed, in this context calculations are much simpler because we then deal with a quantity belonging to the d-th Wiener chaos, so that the Nualart-Peccati [3] criterion of asymptotic normality may be applied. Then, we conclude in the general case (that is, when the entries are no longer supposed to be gaussian) by extending the invariance principle of Nourdin, Peccati and Reinert [2], so to deduce that it was actually not a loss of generality to have assumed that the entries were gaussian.

In this paper we study the more general case of complex-valued entries. As we will see, the obtained limit law is now that of a random variable $Z = Z_1 + iZ_2$, where (Z_1, Z_2) is a gaussian vector whose covariance matrix is expressed by means of the limits of the expectations of the square of (1.1), as well as the modulus of the square of (1.1). To show our result, our strategy consists to adapt, to the complex case, the same method used in the real case. Specifically, we show the following theorem.

Theorem 1.1 *Let $\{X_{ij}\}_{i,j\geqslant 1}$ be a family of centered, complex-valued, independent and identically distributed random variables, with unit variance and admitting moments of all orders. Set*

$$X_n = \left\{ \frac{X_{i,j}}{\sqrt{n}} \right\}_{1\leqslant i,j\leqslant n}.$$

Then, for any integer $k \geqslant 1$,

$$\big\{\mathrm{trace}(X_n^d) - E\big[\mathrm{trace}(X_n^d)\big]\big\}_{1\leqslant d\leqslant k} \xrightarrow{law} \{Z_d\}_{1\leqslant d\leqslant k}.$$

The limit vector $\{Z_d\}_{1\leqslant d\leqslant k}$ takes its values in $\mathbb{C}^k$, and is characterized as follows: the random variables $Z_1, \dots, Z_k$ are independent and, for any $1 \leqslant d \leqslant k$, we have $Z_d =$

$Z_d^1 + iZ_d^2$, where (Z_d^1, Z_d^2) denotes a gaussian vector with covariance matrix equal to

$$\sqrt{d}\begin{pmatrix} a & c \\ c & b \end{pmatrix},$$

with $a+b=1$ and $a-b+i2c = E(X_{1,1}^2)^d$.

The closest result to ours in the existing literature, other than the previously quoted reference by Nourdin and Peccati [1], is due to Rider and Silverstein [4]. At this stage of the exposition, we would like to stress that Theorem 1.1 already appears in the paper [4], but under the following additional assumption on the law of $X_{1,1}$: $\mathrm{Re}(X_{1,1})$ and $\mathrm{Im}(X_{1,1})$ must have a joint density with respect to Lebesgue measure, this density must be bounded, and there exists a positive α such that $E((X_{1,1})^k) \leqslant k^{\alpha k}$ for every $k > 2$. These assumptions can sometimes be too restrictive, typically when one wants to deal with discrete laws. Nevertheless, it is fair to mention that Rider and Silverstein focus more generally on gaussian fluctuations of $\mathrm{trace}\big(f(X_n)\big) - E[\mathrm{trace}\big(f(X_n)\big)]$, when $f:\mathbb{C}\to\mathbb{C}$ is holomorphic and satisfies some additional technical assumptions (whereas, in our paper, we 'only' discuss the polynomial case $f \in \mathbb{C}[X]$).

The rest of the paper is devoted to the proof of Theorem 1.1. To be in position to do so in Section 4, we need to establish some preliminary combinatorial results in Section 2, as well as some results related to the gaussian approximation in Section 3.

2 Some preliminary combinatorial results

As we said, before giving the proof of theorem 1.1 we need to present some combinatorial results. In what follows, we assume that $d \geqslant 2$ is fixed. Note that the pairs $(i_1, i_2), \ldots, (i_d, i_1)$ appearing in formula 1.2 are completely determined by the d-tuple $(i_1, \ldots, i_d)$. Indeed, it is straightforward that the set C_n of elements $((i_1, i_2), .., (i_d, i_1))$ in $\big([1,n]^2\big)^d$ is in bijection with $[1,n]^d$ via the application $((i_1, i_2), .., (i_d, i_1)) \mapsto (i_1, \ldots, i_d)$. The cardinality of C_n is therefore equal to n^d. We denote by D_n the set of *diagonal* terms of C_n, i.e. the set of elements of C_n such that there exist (at least) two distinct integers j and k such that $(i_j, i_{j+1}) = (i_k, i_{k+1})$, with the convention that $i_{d+1} = i_1$. We denote by $ND_n = C_n \setminus D_n$ the set of *non-diagonal* terms. If $i_1, \ldots, i_d$ are pairwise distinct, then $((i_1, i_2), .., (i_d, i_1))$ belongs to ND_n. Thus, the cardinality of ND_n is greater or equal to $n(n-1)\ldots(n-d+1)$, or, equivalently, the cardinality of D_n is less or equal to $n^d - n(n-1)\ldots(n-d+1)$. In particular, the cardinality of D_n is $O(n^{d-1})$.

For any integer $p \geqslant 1$, any integers $\alpha, \beta \in [1,n]$ and any element I_p having the following form

$$I_p = \big((x_1, y_1), \ldots, (x_p, y_p)\big) \in \big([1,n]^2\big)^p, \tag{2.3}$$

we denote by $mc_{I_p}(\alpha, \beta)$ the number of times that the pair (α, β) appears in (2.3). Furthermore, we denote by $mp_{I_p}(\alpha)$ the number of occurrences of α in $\{x_i, y_j\}_{1\leqslant i,j \leqslant p}$. For example, if $I_4 = ((1,3),(3,4),(1,3),(5,7))$ then $mp_{I_4}(3) = 3$, $mp_{I_4}(1) = 2$, $mc_{I_4}(1,3) = 2$ and $mc_{I_4}(3,4) = 1$. For r elements

$$J_k = \big((i_1^{(k)}, i_2^{(k)}), \ldots, (i_d^{(k)}, i_1^{(k)})\big) \in C_n, \quad k = 1, \ldots, r,$$

we define the concatenation $J_1 \sqcup \ldots \sqcup J_r$ as being

$$\Big((i_1^{(1)}, i_2^{(1)}), \ldots, (i_d^{(1)}, i_1^{(1)}), (i_1^{(2)}, i_2^{(2)}), \ldots, (i_d^{(2)}, i_1^{(2)}), \ldots, (i_1^{(r)}, i_2^{(r)}), \ldots, (i_d^{(r)}, i_1^{(r)})\Big).$$

As such, $J_1 \sqcup \ldots \sqcup J_r$ is an element of $(C_n)^r$.

From now on, we denote by $\#A$ the cardinality of a finite set A. The following technical lemma will allow us to estimate the moments of (1.2). More precisely, $(i), (ii), (iii), (iv)$ will imply that the variance of the sum of the diagonal terms converges in L^2 to 0, (v) and (vi) will allow us to show that the variance of the sum of the non-diagonal terms converges to d, and (vii) and $(viii)$ will be used in the computation of the fourth moment of that sum.

Lemma 2.1 *Let the notations previously introduced prevail, and consider the following sets:*

$$\begin{aligned}
A_n &= \big\{I_{2d} = J_1 \sqcup J_2 \in (D_n)^2 : mc_{I_{2d}}(\alpha,\beta) \neq 1 \textit{ for every } \alpha,\beta \in [1,n]\big\} \\
B_n &= \big\{I_{2d} \in A_n : mp_{I_{2d}}(\alpha) \in \{0,4\} \textit{ for every } \alpha \in [1,n]\big\} \\
&= \big\{I_{2d} \in A_n : mc_{I_{2d}}(\alpha,\beta) \in \{0,2\} \textit{ for every } \alpha,\beta \in [1,n]\big\} \\
E_n &= \big\{I_d \in D_n : mc_{I_d}(\alpha,\beta) \neq 1 \textit{ for every } \alpha,\beta \in [1,n]\big\} \\
F_n &= \{I_d \in E_n : mp_{I_d}(\alpha) \in \{0,4\} \textit{ for every } \alpha \in [1,n]\} \\
&= \{I_d \in E_n : mc_{I_d}(\alpha,\beta) \in \{0,2\} \textit{ for every } \alpha,\beta \in [1,n]\} \\
G_n &= \big\{I_{2d} = J_1 \sqcup J_2 \in (ND_n)^2 : mc_{I_{2d}}(\alpha,\beta) \in \{0,2\} \textit{ for every } \alpha,\beta \in [1,n]\big\} \\
H_n &= \big\{I_{2d} \in G_n : mp_{I_{2d}}(\alpha) \in \{0,4\} \textit{ for every } \alpha \in [1,n]\big\} \\
&= \big\{I_{2d} \in G_n : mc_{I_{2d}}(\alpha,\beta) \in \{0,2\} \textit{ for every } \alpha,\beta \in [1,n]\big\} \\
K_n &= \big\{I_{4d} = J_1 \sqcup J_2 \sqcup J_3 \sqcup J_4 \in (ND_n)^4 : mc_{I_{4d}}(\alpha,\beta) \in \{0,2,4\} \textit{ for every } \alpha,\beta \in [1,n]\big\} \\
L_n &= \{I_{4d} \in H_n : mp_{I_{4d}}(\alpha) \in \{0,4\} \textit{ for every } \alpha \in [1,n]\} \\
&= \{I_{4d} \in H_n : mc_{I_{4d}}(\alpha,\beta) \in \{0,2\} \textit{ for every } \alpha,\beta \in [1,n]\}.
\end{aligned}$$

As $n \to \infty$, we have:

(i) $\#(A_n \setminus B_n) = O(n^{d-1})$.

(ii) If d is even, $\#B_n = n \ldots (n-d+1)$; if d is odd, $\#B_n = 0$.

(iii) $\#(E_n \setminus F_n) = O(n^{\frac{d-1}{2}})$.

(iv) If d is even, $\#F_n = n \ldots (n - \frac{d}{2} + 1)$; if d is odd, $\#F_n = 0$.

(v) $\#G_n \setminus H_n = O(n^{d-1})$.

(vi) $\#H_n = d \times n \dots (n-d+1)$.

(vii) $\#(K_n \setminus L_n) = O(n^{2d-1})$.

(viii) $\#L_n = 3d^2 \times n \dots (n-2d+1)$.

Proof.

(i) Let $I_{2d} = \big((i_1,i_2),\dots,(i_d,i_1),(i_{d+1},i_{d+2}),\dots,(i_{2d},i_{d+1})\big) \in A_n \setminus B_n$. By definition of A_n, we have $mp_{I_{2d}}(i_j) \geqslant 4$ for any $j = 1,\dots,2d$. Furthermore, the fact that $I_{2d} \notin B_n$ ensures the existence of at least one integer j_0 between 1 and $2d$ such that $mp_{I_{2d}}(i_{j_0}) > 4$. Let $\sigma : [1,2d] \to [1,2d]$ be defined by $j \mapsto \sigma(j) = \min\{k : \ i_k = i_j\}$. It is readily checked that

$$4d = \sum_{\alpha \in \mathrm{Im}(\sigma)} mp_{I_{2d}}(i_\alpha).$$

We conclude that $\#\mathrm{Im}(\sigma) < d$. Therefore, $\#\left(A_n \setminus B_n\right) = O(n^{d-1})$.

(ii) Assume that B_n is non-empty. Let

$$I_{2d} = \big((i_1,i_2),\dots,(i_d,i_1),(i_{d+1},i_{d+2}),\dots,(i_{2d},i_{d+1})\big) \in B_n.$$

For every integer $j \in [1,2d]$, we have $mp_{I_{2d}}(i_j) = 4$. Defining σ and proceeding as in point (i) above, we obtain that $\#\mathrm{Im}(\sigma) = d$. We set $m = \min\{l \in \mathrm{Im}(\sigma) \,|\, l+1 \notin \mathrm{Im}(\sigma)\}$. Since $\#\mathrm{Im}(\sigma) = d$, it follows that $m \leqslant d$. In fact $m \leqslant d-1$, otherwise the elements of the d-tuple $(i_1,\dots,i_d)$ would be all distinct, and $((i_1,i_2),\dots,(i_d,i_1))$ could not be in D_n, which would yield a contradiction. In the case $d = 2$, $I_{2d} = \big((i_1,i_2),(i_2,i_1),(i_3,i_4),(i_4,i_3)\big) \in B_n$ if and only if $i_1 = i_2$, $i_3 = i_4$ and $i_1 \neq i_3$. Thus, the cardinality of B_n is equal to $n(n-1)$.

In what follows we suppose that $d \geqslant 3$.

Let us show that d is even (which will prove that B_n is empty if d is odd) and that I_{2d} can be written as

$$\Big((l_1,l_2),\dots,(l_{\frac{d}{2}-1},l_{\frac{d}{2}}),(l_{\frac{d}{2}},l_1),(l_1,l_2),\dots,(l_{\frac{d}{2}-1},l_{\frac{d}{2}}),(l_{\frac{d}{2}},l_1),$$
$$(j_1,j_2)\dots,(j_{\frac{d}{2}-1},j_{\frac{d}{2}}),(j_{\frac{d}{2}},j_1),(j_1,j_2)\dots,(j_{\frac{d}{2}-1},j_{\frac{d}{2}}),(j_{\frac{d}{2}},j_1)\Big),$$

where $l_1,\dots,l_{\frac{d}{2}},j_1,\dots,j_{\frac{d}{2}}$ are pairwise distinct integers in $[1,n]$, which will prove that the formula for $\#B_n$ given in (ii) holds true. The proof is divided in several parts.

(a) Using a proof by contradiction, let us assume that there exists an integer q in $[m+1, d]$ such that i_q does not belong to $\{i_1, \ldots, i_m\}$. We denote by γ the smallest element verifying this. Note that $\gamma \geqslant m+2$ necessarily, and that there exists an integer $p \leqslant m$ such that $i_{\gamma-1} = i_p$. Therefore, $i_{\gamma-1}$ appears in the four pairs

$$(i_{p-1}, i_p),\ (i_p, i_{p+1}),\ (i_{\gamma-2}, i_{\gamma-1}),\ (i_{\gamma-1}, i_\gamma).$$

Note that for the four pairs above, it is possible that the two pairs in the middle are the same. By definition of B_n, we have $mp_{I_{2d}}(i_{\gamma-1}) = 4$ so these pairs are the only pairs of I_{2d} containing the integer $i_{\gamma-1}$. Moreover, by definition of A_n, we have $mc_{I_{2d}}(i_{\gamma-1}, i_\gamma) \geqslant 2$. Thus, we necessarily have either $(i_{\gamma-1}, i_\gamma) = (i_p, i_{p+1})$; or $(i_{\gamma-1}, i_\gamma) = (i_{\gamma-2}, i_{\gamma-1})$; or $(i_{\gamma-1}, i_\gamma) = (i_{p-1}, i_p)$. If we had $(i_{\gamma-1}, i_\gamma) = (i_{\gamma-2}, i_{\gamma-1})$, then we would have $i_{\gamma-2} = i_{\gamma-1} = i_\gamma$ and i_γ would appear at least six times in the writing of I_{2d}, which is not possible . Similarly, $(i_{\gamma-1}, i_\gamma) = (i_{p-1}, i_p)$ is impossible. Thus, it must hold that $(i_{\gamma-1}, i_\gamma) = (i_p, i_{p+1})$. We can therefore state that $i_\gamma = i_{p+1}$. Since we also have that $p+1 \leqslant m+1$ and $i_{m+1} \in \{i_1, \ldots, i_m\}$, we can conclude that $i_\gamma \in \{i_1, \ldots, i_m\}$, which yields the desired contradiction. Hence,

$$\{i_{m+1}, \ldots, i_d\} \subset \{i_1, \ldots, i_m\}. \tag{2.4}$$

(b) Let us show that if $l, k \leqslant d-1$ are two distinct integers satisfying $i_k = i_l$, then $(i_k, i_{k+1}) = (i_l, i_{l+1})$. Let $l, k \leqslant d-1$ be two integers such that $l \neq k$ et $i_k = i_l$. The integer i_l appears in the four pairs $\{(i_{l-1}, i_l), (i_l, i_{l+1}), (i_{k-1}, i_k), (i_k, i_{k+1})\}$ (or only in three pairs, if both pairs in the middle are the same, which happens whether $l = k-1$). As $mp_{I_{2d}}(i_k) = 4$, these pairs are the only pairs of I_{2d} containing the integer i_k. By definition of A_n, all pairs of I_{2d} must have at least two occurrences in I_{2d}. If we have $(i_k, i_{k+1}) = (i_{k-1}, i_k)$ then we have $i_k = i_{k+1} = i_{k-1}$ and i_k appears at least six times in I_{2d}, which cannot be true. Similarly, $(i_k, i_{k+1}) = (i_{l-1}, i_l)$ is impossible. Therefore, it must hold that $(i_k, i_{k+1}) = (i_l, i_{l+1})$.
(c) It follows from the definition of m that there exists an integer $r \in [1, m]$ satisfying $i_{m+1} = i_r$. Let us show that

$$(i_1, \ldots, i_d) = (i_1, \ldots, i_m, i_r, \ldots, i_{r+d-m-1}). \tag{2.5}$$

If $m = d-1$, then $(i_1, \ldots, i_d) = (i_1, \ldots, i_m, i_r)$ and (2.5) is verified. If $m \leqslant d-2$ then, being given that $i_{m+1} = i_r$ and that we already showed in (b) that if $l, k \leqslant d-1$ are two distinct integers satisfying $i_k = i_l$ then $i_{k+1} = i_{l+1}$, we can state that $i_{m+2} = i_{r+1}$. Thus, if $m = d-2$ then $(i_1, \ldots, i_d) = (i_1, \ldots, i_m, i_r, i_{r+1})$, and (2.5) is once again verified. Finally, if $m \leqslant d-3$, we iterate this process as many times as necessary until we get (2.5).

(d) Let us now prove that the elements of $(i_r, \ldots, i_{r+d-m-1})$ are all distinct. Once again, we use a proof by contradiction. Thus, let us assume that there exists an integer p in $[1,n]$ which appears at least twice in the uplet $(i_r,\dots,i_{r+d-m-1})$. We then have $\{i_r,\dots,i_{r+d-m-1}\} = \{i_{m+1},\dots,i_d\} \subset \{i_1,\dots,i_m\}$, see (2.4) and (2.5). Thus, p appears at least three times overall in the uplet $(i_1,\dots,i_d)$. This latter fact implies $mp_{I_{2d}}(p) \geqslant 6$, which contradicts the assumption $mp_{I_{2d}}(p) = 4$.

(e) Finally, let us establish that $2m = d$ and $r = 1$. The elements of $(i_r,\dots,i_{r+d-m-1})$ being all distinct, the couple $(i_{r+d-m-1},i_1) = (i_d,i_1)$ cannot belong to the set of pairs

$$\{(i_r,i_{r+1}),\dots,(i_{r+d-m-2},i_{r+d-m-1})\} = \{(i_{m+1},i_{m+2}),\dots,(i_{d-1},i_d)\}.$$

(because, by (d), no pair of this set can have $i_{r+d-m-1}$ as a first coordinate.) Moreover, since i_1 does not belong to $\{i_2,\dots,i_m\}$ then the pair $(i_{r+d-m-1},i_1)$ cannot belong to the set of pairs $\{(i_1,i_2),\dots,(i_{m-1},i_m)\}$ (because no pair of this set can have i_1 as a second coordinate). Also, the integer i_d appearing at least twice in the uplet $(i_1,\dots,i_d)$, it cannot belong to the uplet $(i_{d+1},\dots,i_{2d})$ (otherwise, i_d would appear at least six times in the pairs of I_{2d}). Thus, the only way for the occurrence of the pair (i_d,i_1) in I_{2d} to be greater or equal than 2 is that $(i_{r+d-m-1},i_1) = (i_m,i_{m+1})$. Therefore, $i_1 = i_{m+1} = i_r$. As $i_1,\dots,i_m$ are all distinct and $r \leqslant m$, it must hold that $r = 1$. Hence, $r+d-m-1 = d-m$ and $i_{d-m} = i_m$. Since $i_1,\dots,i_{d-m}$ are all distinct, see indeed (d), it must be true that $m \geqslant d-m$. Since $i_{d-m} = i_m$, we conclude that $d-m = m$, that is, $d = 2m$. As such, we establish that $(i_{d+1},\dots,i_{2d}) = (i_{d+1},\dots,i_{\frac{3d}{2}},i_{d+1},\dots,i_{\frac{3d}{2}})$. Let us finally note that $i_1,\dots,i_{\frac{d}{2}},i_{d+1},\dots,i_{\frac{3d}{2}}$ are necessarily distinct because $mp_{I_{2d}}(i_j) = 4$. This completes the proof of part (ii).

(iii) Consider $I = \big((i_1,i_2),\dots,(i_d,i_1)\big) \in E_n \setminus F_n$. Let $\xi : [1,d] \to [1,d]$ be defined by $j \mapsto \xi(j) = min\{k \,|\, i_k = i_j\}$. From the equation $2d = \sum_{\alpha\in Im(\xi)} mp_I(i_\alpha)$, and using the fact that all $mp_I(i_\alpha)$ are greater or equal than 4, as well as there exists an α in $\mathrm{Im}(\xi)$ satisfying $mp_I(i_\alpha) > 4$, we conclude that $\#\mathrm{Im}(\xi) < \frac{d}{2}$. Therefore, the cardinality of $E_n \setminus F_n$ is equal to $O(n^{\frac{d-1}{2}})$.

(iv) Let us assume that F_n is not empty. Consider $I_d = \big((i_1,i_2),\dots,(i_d,i_1)\big) \in F_n$. Proceeding as in point (ii) above, we conclude that F_n is empty in the case where d is odd, and that the elements of F_n have the following form when d is even:

$$\Big((l_1,l_2),\dots,(l_{\frac{d}{2}-1},l_{\frac{d}{2}})(l_{\frac{d}{2}},l_1),(l_1,l_2),\dots,(l_{\frac{d}{2}-1},l_{\frac{d}{2}})(l_{\frac{d}{2}},l_1)\Big).$$

Here, $l_1,\dots,l_{\frac{d}{2}}$ are pairwise distinct integers in $[1,n]$. The formula of $\#F_n$ given in (iv) follows directly from that.

(v) Consider $I = \big((i_1,i_2),\dots,(i_d,i_1),(i_{d+1},i_{d+2}),\dots,(i_{2d},i_{d+1})\big) \in G_n \setminus H_n$. Let $\zeta : [1,2d] \to [1,2d]$ be defined by $j \mapsto \zeta(j) = min\{k \,|\, i_k = i_j\}$. From the identity $4d = \sum_{\alpha\in Im(\zeta)} mp_I(i_\alpha)$, and using the fact that $mp_I(i_\alpha)$ are greater or equal than 4, as well as that there exists an α in $\mathrm{Im}(\zeta)$ satisfying $mp_I(i_\alpha) > 4$, we conclude as in (i) that $\#\mathrm{Im}(\zeta) < d$. Therefore, the cardinality of $G_n \setminus H_n$ is $O(n^{d-1})$.

(vi) Consider $I = \big((i_1, i_2), \ldots, (i_d, i_1), (i_{d+1}, i_{d+2}), \ldots, (i_{2d}, i_{d+1})\big) \in H_n$.
(a) By definition of ND_n, there is no redundancy neither among the pairs $(i_1, i_2), \ldots, (i_d, i_1)$ nor among the pairs $(i_{d+1}, i_{d+2}), \ldots, (i_{2d}, i_{d+1})$. Therefore, to satisfy the constraint defining H_n, it is necessary and sufficient that each couple of $(i_1, i_2), \ldots, (i_d, i_1)$ matches one and only one couple among $(i_{d+1}, i_{d+2}), (i_{d+2}, i_{d+3}), \ldots, (i_{2d}, i_{d+1})$.
(b) Using a proof by contradiction, let us show that the elements of $\{i_1, \ldots, i_d\}$ are pairwise distinct. If p and q were two distinct integers in $[1, d]$ such that $i_p = i_q$ then, according to (a), there would exist $k \in [d+1, 2d]$ satisfying $(i_p, i_{p+1}) = (i_k, i_{k+1})$, which would yield $i_p = i_q = i_k$ and, consequently, $mp_I(i_p) \geqslant 6$. This would contradict the fact that $mp_I(i_p) = 4$.
(c) Let us establish that, for every $p \in [1, d]$ and $q \in [d+1, 2d]$ such that $i_p = i_q$, we have $i_{p+1} = i_{q+1}$. Using a proof by contradiction, let us assume that there exists an integer $q' \in [d+1, 2d]$ different from q such that $(i_p, i_{p+1}) = (i_{q'}, i_{q'+1})$. Then it must hold that $i_p = i_q = i_{q'}$ and $mp_I(i_p) \geqslant 6$, which contradicts the fact that $mp_I(i_p) = 4$.

The results (a), (b) et (c) allow us to conclude that there exists an integer $k \in [1, d]$ satisfying $(i_1, \ldots, i_d) = (i_{d+k}, i_{d+k+1}, \ldots, i_{2d}, i_{d+1}, \ldots, i_{d+k-1})$. Thus, the elements of G_n are completely characterized by a given integer $k \in [1, d]$ and a given set $\{i_1, \ldots, i_d\}$ where i_j are pairwise distinct integers in $[1, n]$. We can therefore conclude that $\#H_n = d \times n \ldots (n - d + 1)$.

(vii) Consider $I = \big((i_1, i_2), \ldots, (i_d, i_1), \ldots, (i_{3d+1}, i_{3d+2}), \ldots, (i_{4d}, i_{3d+1})\big) \in K_n \setminus L_n$. Let $\eta : [1, 4d] \longrightarrow [1, 4d]$ be the application defined by $\eta(j) = min\{k |\, i_k = i_j\}$. From the identity $8d = \sum_{\alpha \in Im(\eta)} mp_I(i_\alpha)$, and using the fact that $mp_I(i_\alpha)$ are all greater or equal than 4, as well as for at least one $\alpha \in \mathrm{Im}(\eta)$ it must hold that $mp_I(i_\alpha) > 4$, we conclude that $\#\mathrm{Im}(\eta) < 2d$. Therefore, the cardinality of $K_n \setminus L_n$ is $O(n^{2d-1})$.

(viii) Consider $I = \big((i_1, i_2), \ldots, (i_d, i_1), \ldots, (i_{3d+1}, i_{3d+2}), \ldots, (i_{4d}, i_{3d+1})\big) \in L_n$. For every $j \leqslant 4d$, we have $mp_I(i_j) = 4$. Then $2d = \#\mathrm{Im}(\eta)$, with η as in point (vii).

(a) Using a proof by contradiction, let us show that, for every $k \in [0, 3]$, the integers $i_{kd+1}, \ldots, i_{(k+1)d}$ are all distinct. Assume that there exist two distinct integers l and h in $[1, d]$, as well as an integer $k \in [0, 3]$, satisfying $i_{kd+l} = i_{kd+h}$. By definition of the set L_n, we have $\big((i_{kd+1}, i_{kd+2}), \ldots, (i_{(k+1)d}, i_{kd+1})\big) \in ND_n$. Then, the pairs

$$\big\{(i_{kd+1}, i_{kd+2}), \ldots, (i_{(k+1)d}, i_{kd+1})\big\}$$

are all distinct, and we have $mc_I(i_{kd+h}, i_{kd+h+1}) = 2$, which implies that there exists $k' \in [0, 3]$, different from k, and $h' \in [1, d]$ satisfying $i_{kd+h} = i_{k'd+h'}$. It follows that i_{kd+h} appears at least six times in I, which contradicts the fact that $mp_I(i_{kd+h}) = 4$.
(b) For any $p = 0, \ldots, 3$, let us introduce $M_p = \{i_{pd+1}, \ldots, i_{(p+1)d}\}$. For any integers p, q in $[0, 3]$, we have either $M_p \bigcap M_q = \varnothing$ or $M_p = M_q$. Otherwise there would exist an integer j such that $i_{qd+j} \in M_p$ and $i_{qd+j+1} \notin M_p$ and, since $mc_I(i_{qd+j}, i_{qd+j+1}) = 2$, there would exist $q' \in [0, 3]$, different from p and q, and $j' \in [1, d]$ such that

$(i_{qd+j}, i_{qd+j+1}) = (i_{q'd+j'}, i_{q'd+j'+1})$; therefore i_{qd+j} would appear at least six times in I, which would yield a contradiction.
(c) If $M_p = M_q$, then proceeding as in point (vi), we show that there exists $j \in [1, d]$ such that

$$(i_{pd+1}, \ldots, i_{(p+1)d}) = (i_{qd+j}, \ldots, i_{(q+1)d}, i_{qd+1}, \ldots, i_{dq+j-1}).$$

The results (a), (b) et (c) allow us to conclude that a generic element of L_n is characterized by:
- the choice of one case among the following three cases: either $M_0 = M_1$ and $M_2 = M_3$; or $M_0 = M_2$ and $M_1 = M_3$; or $M_0 = M_3$ and $M_1 = M_2$. In what follows, we consider the case $M_0 = M_1$ and $M_2 = M_3$ (we can proceed similarly in the other two cases);
- the choice of $2d$ integers $i_1, \ldots, i_d, i_{2d+1}, \ldots, i_{3d}$ that are pairwise distinct in $[1, n]$;
- the choice of an integer $k \in [1, d]$ such that $(i_{d+1}, \ldots, i_{2d}) = (i_k, \ldots, i_d, i_1, \ldots, i_{k-1})$;
- the choice of an integer $k' \in [1, d]$ such that

$$(i_{3d+1}, \ldots, i_{4d}) = (i_{2d+k'}, \ldots, i_{3d}, i_{2d+1}, \ldots, i_{2d+k'-1}).$$

It is now easy to deduce that $\#L_n = 3d^2 n \ldots (n - 2d + 1)$.

■

3 Gaussian approximations

Let $X = \{X^i\}_{i \geqslant 1}$ be a family of centered independent random variables taking values in $\mathbb{R}^r$ and having pairwise uncorrelated components with unit variance. Let $G = \{G^i\}_{i \geqslant 1}$ be a family of independent standard gaussian random variables taking values in $\mathbb{R}^r$ and having independent components. Suppose also that X and G are independent, and set

$$X = (X_1^1, \ldots, X_r^1, X_1^2, \ldots, X_r^2, \ldots) = (X_1, \ldots, X_r, X_{r+1}, \ldots, X_{2r}, \ldots).$$

i.e., $X_{j+(i-1)r} = X_j^i$.

Consider integers $m \geqslant 1$, $d_m \geqslant \ldots \geqslant d_1 \geqslant 2$, $N_1, \ldots, N_m$, as well as real symmetric functions $f_1, \ldots, f_m$ such that each function f_i is defined on $[1, rN_i]^{d_i}$ and vanishes at the points $(i_1, ..., i_{d_i})$ such that $\exists j \neq k$ for which $\lceil i_j/r \rceil = \lceil i_k/r \rceil$ (we remind that $\lceil x \rceil$ means the unique integer k such that $k < x \leqslant k+1$). Let us define

$$Q^i(X) = Q_{d_i}(f_i, X) = \sum_{i_1, \ldots, i_d = 1}^{rN_i} f_i(i_1, \ldots, i_{d_i})\, X_{i_1} \ldots X_{i_{d_i}}.$$

In the case of complex-valued matrices, the real and imaginary parts of the entries $X_{i,j}$ are not necessarily independent. Therefore, we will need to modify the results used by Nourdin and Peccati in the paper [3]. The following lemma is a variant, weaker in terms of assumptions, of the hypercontractivity property.

Lemma 3.1 *Let the notations previously introduced prevail. Assume that* $\alpha = \sup_i E(|X_i|^4) < \infty$ *and set* $K = 36 \times 25^r \times (1 + 2\alpha^{\frac{3}{4}})^2$. *Then*

$$E(Q_d(X)^4) \leqslant K^d E(Q_d(X)^2)^2. \tag{3.6}$$

Proof. Set

$$\begin{cases} U = \sum_{\forall k:\, i_k \notin \{(N-1)r+1,\ldots,Nr\}} f(i_1,\ldots,i_d) X_{i_1}\ldots X_{i_d} \\ V_j = \sum_{\exists! k: i_k=(N-1)r+j} f(i_1,\ldots,i_d) X_{i_1}\ldots \widehat{X_{(N-1)r+j}} \ldots X_{i_d} \end{cases}$$

The notation $\widehat{X_{(N-1)r+j}}$ means that this term is removed from the product. Observe that $X_{(N-1)r+j} = X_j^N$ according to the notation that we adopted previously, and that the quantity $Q_d(X)$ is given by:

$$Q_d(X) = U + \sum_{j=1}^r X_j^N V_j$$

(as f vanishes at the previously specified points). Note that, for every $p \leqslant N$ and every $i, j \in [1, r]$, X_j^p is independent from U and V_i. Thus, by choosing $p = N$, we have

$$\begin{aligned} E(Q_d(X)^4) &= \sum_{s_0+\ldots+s_r=4} \frac{24}{s_0!\ldots s_r!} E(U^{s_0} \prod_{j=1}^r (V_j X_j^N)^{s_j}) \\ &= E(U^4) + \sum_{s_1+\ldots+s_r=2} \frac{12}{s_1!\ldots s_r!} E(U^2 \prod_{j=1}^r V_j^{s_j}) E(\prod_{j=1}^r (X_j^N)^{s_j}) \\ &+ \sum_{s_1+\ldots+s_r=3} \frac{24}{s_1!\ldots s_r!} E(U \prod_{j=1}^r V_j^{s_j}) E(\prod_{j=1}^r (X_j^N)^{s_j}) \\ &+ \sum_{s_1+\ldots+s_r=4} \frac{24}{s_1!\ldots s_r!} E(\prod_{j=1}^r V_j^{s_j}) E(\prod_{j=1}^r (X_j^N)^{s_j}). \end{aligned}$$

In the equation above, we used that $\sum_{s_1+\ldots+s_r=1} \frac{4}{s_1!\ldots s_r!} E(U^3 \prod_{j=1}^r (V_j X_j^N)^{s_j}) = 0$ since X_j^N are centered. By using the generalized Hölder inequality, we obtain:

$$E(U^{s_0} \prod_{j=1}^r V_j^{s_j}) \leqslant E\left(U^4\right)^{\frac{s_0}{4}} \prod_{j=1}^r E(V_j^4)^{\frac{s_j}{4}}.$$

Since the terms $E(V_j^4)^{\frac{s_j}{4}}$ are upper bounded by $\left(\sum_{j=1}^r E(V_j^4)^{\frac{1}{2}}\right)^{\frac{s_j}{2}}$, we obtain:

$$\sum_{s_1+\ldots+s_r=4-s_0} E(U^{s_0} \prod_{j=1}^r V_j^{s_j}) \leqslant 5^r E\left(U^4\right)^{\frac{s_0}{4}} \left(\sum_{j=1}^r E(V_j^4)^{\frac{1}{2}}\right)^{\frac{4-s_0}{2}}.$$

Using the generalized Hölder inequality again, we have $E(\prod_{j=1}^{r}(X_j^N)^{s_j}) \leqslant \prod_{j=1}^{r} E\left((X_j^N)^4\right)^{\frac{s_j}{4}} \leqslant \alpha^{\frac{\sum s_j}{4}}$. Therefore:

$$E(Q_d(X)^4) \leqslant E(U^4) + 12 \times 5^r E(U^4)^{\frac{1}{2}} \sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}} + 24 \times 5^r \alpha^{\frac{3}{4}} E(U^4)^{\frac{1}{4}} \left(\sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}}\right)^{\frac{3}{2}} + 24 \times 5^r \alpha \left(\sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}}\right)^2. \tag{3.7}$$

Note that α does not appear in the second term of the right-hand side of the inequality above because X_j^N are random variable with unit variance and zero covariance. Using the inequality $x^{\frac{1}{4}}y^{\frac{3}{2}} \leqslant x^{\frac{1}{2}}y + y^2$, obtained by separating the cases $x \leqslant y^2$ and $x \geqslant y^2$, we get:

$$E(U^4)^{\frac{1}{4}} \left(\sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}}\right)^{\frac{3}{2}} \leqslant E(U^4)^{\frac{1}{2}} \sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}} + \left(\sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}}\right)^2.$$

Then

$$E(Q_d(X)^4) \leqslant E(U^4) + 12 \times 5^r (1 + 2\alpha^{\frac{3}{4}}) E(U^4)^{\frac{1}{2}} \sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}} + 24 \times 5^r (\alpha^{\frac{3}{4}} + \alpha) \left(\sum_{j=1}^{r} E(V_j^4)^{\frac{1}{2}}\right)^2. \tag{3.8}$$

To prove the hypercontractivity property (3.6), we will use an induction on N. When $N = 1$, because f vanishes at the previously specified points, then the only case where the value taken by $Q_d(X)$ is not zero is when $d = 1$, that is, when $Q_d(X)$ has the form $\sum_{j=1}^{r} a_j X_j^1$. In this case, $U = 0$ and $V_j = a_j$. Thus, by (3.7), we have $E(Q_d(X)^4) \leqslant 24 \times 5^r \alpha \left(\sum_{j=1}^{r} a_j^2\right)^2$. It follows that $E(Q_d(X)^4) \leqslant K E(Q_d(X)^2)^2$. Let us now assume that the result holds for $N - 1$. Then, because U and V_j are functions of $X^1, \ldots, X^{N-1}$, we can apply the recursive hypothesis to $E(U^4)$ and $E(V_j^4)$, and obtain that:

$$\begin{aligned}
E(Q_d(X)^4) \leqslant\ & K^d \left[E(U^2)^2 + \frac{12 \times 5^r (1 + 2\alpha^{\frac{3}{4}})}{K^{\frac{1}{2}}} E(U^2) \sum_{j=1}^{r} E(V_j^2) \right] \\
& + K^d \frac{24 \times 5^r (\alpha + \alpha^{\frac{3}{4}})}{K} \left(\sum_{j=1}^{r} E(V_j^2) \right)^2 \\
& \leqslant K^d \left[E(U^2)^2 + 2E(U^2) \sum_{j=1}^{r} E(V_j^2) + \left(\sum_{j=1}^{r} E(V_j^2) \right)^2 \right] \\
& = K^d \left[E(U^2) + \sum_{j=1}^{r} E(V_j^2) \right]^2 .
\end{aligned}$$

Furthermore, since the X_j^N are centered, unit-variance and independent of U and of V_j, we have

$$\begin{aligned}
E(Q_d(X)^2) & = E((U + \sum_{j=1}^{r} X_j^N V_j)^2) \\
& = E(U^2) + 2 \sum_{j=1}^{r} E(UV_j) E(X_j^N) + \sum_{i,j=1,\ldots,r} E(V_i V_j) E(X_i^N X_j^N) \\
& = E(U^2) + \sum_{j=1}^{r} E(V_j^2),
\end{aligned}$$

which completes the proof. ■

The following two lemmas will be used to prove the convergence in law of the sum of the non-diagonal terms in (1.2), and to show that the limit does not depend on the common law of $X_{i,j}$.

Lemma 3.2 *Let $\{X^i\}_{i\geqslant 1}$, $\{G^i\}_{i\geqslant 1}$ and $Q^i(X)$ be as in the beginning of section 3. Let us assume that $\beta = \sup_i E(|X_i|^3) < \infty$, $E(Q^i(X)^2) = 1$, and that V is the symmetric matrix defined as $V(i,j) = E(Q^i(X)Q^j(X))$. Consider $Z_V = (Z_V^1, \ldots, Z_V^m) \sim N_m(0, V)$ (i.e. Z_V is a gaussian vector with a covariance matrix equal to V).*

1. *If $\varphi : \mathbb{R}^m \to \mathbb{R}$ is a function of class C^3 such that $\left\| \varphi''' \right\|_\infty < \infty$ then*

$$\left| E(\varphi(Q^1(X), \ldots, Q^m(X))) - E(\varphi(Q^1(G), \ldots, Q^m(G))) \right|$$

$$\leqslant \left\| \varphi''' \right\|_\infty \left(\beta + \sqrt{\frac{8}{\pi}} \right) K^{\frac{3}{4}(d_m - 1)} r^3 m^4 \frac{d_m!^3}{d_1!(d_1-1)!} \sqrt{\max_{1\leqslant k \leqslant m} \max_{1 \leqslant j \leqslant N_k} inf_j f_k},$$

where

$$inf_j f_k = \sum_{i_1,\dots,i_{d_k-1}=1}^{rN_k} f_k(j, i_1, \dots, i_{d_k-1})^2.$$

2. *If* $\varphi : \mathbb{R}^m \to \mathbb{R}$ *is a function of class* C^3 *such that* $\left\|\varphi'''\right\|_\infty < \infty$ *then*

$$\left|E(\varphi(Q^1(X), \dots, Q^m(X))) - E(\varphi(Z_V)\right| \leqslant \left\|\varphi''\right\|_\infty \left(\sum_{i=1}^{m} \Delta_{i,i} + 2 \sum_{1\leqslant i<j\leqslant m} \Delta_{i,j}\right)$$

$$+ \frac{\left\|\varphi'''\right\|_\infty m^4 d_m!^3}{d_1!(d_1-1)!} \left(\left(\beta + \sqrt{\frac{8}{\pi}}\right) K^{\frac{3}{4}(d_m-1)} r^3 + \sqrt{\frac{32}{\pi}} \left(\frac{64}{\pi}\right)^{d_m-1}\right) \sqrt{\max_{1\leqslant k\leqslant m} \max_{1\leqslant j\leqslant N_k} inf_j f_k}$$

where $inf_j f_k$ *as above and* $\Delta_{i,j}$ *given by*

$$d_j \sum_{s=1}^{d_i-1} (s-1)! \begin{pmatrix} d_i - 1 \\ s-1 \end{pmatrix} \begin{pmatrix} d_j - 1 \\ s-1 \end{pmatrix} \sqrt{(d_i + d_j - 2s)!} \left(\|f_i \star_{d_i-s} f_i\|_2 + \left\|f_j \star_{d_j-s} f_j\right\|_2\right)$$

$$+ 1_{d_i<d_j} \sqrt{d_j! \begin{pmatrix} d_j \\ d_i \end{pmatrix} \left\|f_j \star_{d_j-d_i} f_j\right\|_2},$$

with

$$f_j \star_r f_j(i_1, \dots, i_{2d_j-2r}) = \sum_{k_1,\dots,k_r=1}^{rN_j} f_j(k_1, \dots, k_r, i_1, \dots, i_{d_j-r}) f_j(k_1, \dots, k_r, i_{d_j-r+1}, \dots, i_{2d_j-2r}).$$

Proof. Set $Q(X) = (Q^1(X), \dots, Q^m(X))$ and, for any $1 \leqslant p \leqslant N+1$, consider

$$\begin{cases} Z^{(p)} = \left(G_1, \dots, G_{(p-1)r}, X_{(p-1)r+1}, \dots, X_{rN}\right) \\ U_p^{(i)} = \sum_{\forall k:\, i_k \notin \{(p-1)r+1,\dots,pr\}} f_i(i_1, \dots, i_d) Z_{i_1}^{(p)} \dots Z_{i_d}^{(p)} \\ V_{p,j}^{(i)} = \sum_{\exists! k: i_k = (p-1)r+j} f_i(i_1, \dots, i_d) Z_{i_1}^{(p)} \dots \widehat{Z_{(p-1)r+j}^{(p)}} \dots Z_{i_d}^{(p)} \end{cases}$$

The notation $\widehat{Z_{(p-1)r+j}^{(p)}}$ means that this term is removed from the product. Let us set $U_p = (U_p^{(1)}, \dots, U_p^{(m)})$ and $V_{p,j} = (V_{p,j}^{(1)}, \dots, V_{p,j}^{(m)})$. Note that $Q(Z^{(p)})$ can be written as

$$Q(Z^{(p)}) = U_p + \sum_{j=1}^{r} X_j^p V_{p,j}.$$

Similarly, we have: $Q(Z^{(p+1)}) = U_p + \sum_{j=1}^{r} G_j^p V_{p,j}$. For a vector $Y = (Y_1, \dots, Y_m)$ in $\mathbb{R}^m$ and a vector $s = (s_1, \dots, s_m)$ in $\mathbb{N}^m$, we set $Y^s = \prod_{i=1}^m Y_i^{s_i}$.

1. Let φ be a function of class C^3. The Taylor formula gives:

$$\left| E\left(\varphi(Q(Z^{(p)}))\right) - E\left(\sum_{|s|\leqslant 2}\frac{1}{s!}\partial^s\varphi(U_p)\left(\sum_{j=1}^{r}X_j^pV_{p,j}\right)^s\right)\right| \leqslant \left\|\varphi'''\right\|_\infty \left|E\left(\sum_{|s|=3}\left(\sum_{j=1}^{r}X_j^pV_{p,j}\right)^s\right)\right|.$$

Note that, for every p, X_j^p is independent from U_p and from $V_{p,i}$. Thus, we have:

$$\begin{aligned}\left|E\left(\textstyle\sum_{|s|=3}\left(\sum_{j=1}^{r}X_j^pV_{p,j}\right)^s\right)\right|) &= \left|E\left(\sum_{k,l,q=1}^{m}\sum_{j_1=1}^{r}X_{j_1}^pV_{p,j_1}^{(k)}\sum_{j_2=1}^{r}X_{j_2}^pV_{p,j_2}^{(l)}\sum_{j_3=1}^{r}X_{j_3}^pV_{p,j_3}^{(q)}\right)\right| \\ &= \left|\sum_{k,l,q=1}^{m}\sum_{j_1=1}^{r}\sum_{j_2=1}^{r}\sum_{j_3=1}^{r}E\left(X_{j_1}^pX_{j_2}^pX_{j_3}^p\right)E\left(V_{p,j_1}^{(k)}V_{p,j_2}^{(l)}V_{p,j_3}^{(q)}\right)\right|.\end{aligned}$$

The Hölder inequality ensures that:

$$\left|E\left(X_{j_1}^pX_{j_2}^pX_{j_3}^p\right)\right| \leqslant E\left(\left|X_{j_1}^p\right|^3\right)^{\frac{1}{3}}E\left(\left|(X_{j_2}^p)\right|^3\right)^{\frac{1}{3}}E\left(\left|(X_{j_3}^p)\right|^3\right)^{\frac{1}{3}} \leqslant \beta.$$

Using the Hölder inequality, as well as the lemma 3.1 and the relation $E\left((V_{p,n}^{(k)})^2\right) = d_k!^2 inf_{pr+n}f_k$, we obtain

$$\begin{aligned}\left|E\left(V_{p,j_1}^{(k)}V_{p,j_2}^{(l)}V_{p,j_3}^{(q)}\right)\right| &\leqslant E\left(\left|V_{p,j_1}^{(k)}\right|^4\right)^{\frac{1}{4}}E\left(\left|V_{p,j_2}^{(l)}\right|^4\right)^{\frac{1}{4}}E\left(\left|V_{p,j_3}^{(q)}\right|^4\right)^{\frac{1}{4}} \\ &\leqslant K^{\frac{3}{4}(d_m-1)}E\left(\left|V_{p,j_1}^{(k)}\right|^2\right)^{\frac{1}{2}}E\left(\left|V_{p,j_2}^{(l)}\right|^2\right)^{\frac{1}{2}}E\left(\left|V_{p,j_3}^{(q)}\right|^2\right)^{\frac{1}{2}} \\ &\leqslant K^{\frac{3}{4}(d_m-1)}\left(d_m!^2\max_{1\leqslant j\leqslant r}\max_{1\leqslant k\leqslant m}inf_{pr+j}f_k\right)^{\frac{3}{2}}.\end{aligned}$$

Then,

$$\left|E\left(\varphi(Q(Z^{(p)}))\right) - E\left(\sum_{|s|\leqslant 2}\frac{1}{s!}\partial^s\varphi(U_p)\left(\sum_{j=1}^{r}X_j^pV_{p,j}\right)^s\right)\right|$$

$$\leqslant \left\|\varphi'''\right\|_\infty \beta K^{\frac{3}{4}(d_m-1)}(r\,m)^3\left(d_m!^2\max_{1\leqslant j\leqslant r}\max_{1\leqslant k\leqslant m}inf_{pr+j}f_k\right)^{\frac{3}{2}}.$$

By writing the same formula for $Q(Z^{(p+1)})$ we obtain this time

$$\left|E\left(\varphi(Q(Z^{(p+1)}))\right) - E\left(\sum_{|s|\leqslant 2}\frac{1}{s!}\partial^s\varphi(U_p)\left(\sum_{j=1}^{r}G_j^pV_{p,j}\right)^s\right)\right|$$

$$\leqslant \left\|\varphi'''\right\|_\infty \sqrt{\frac{8}{\pi}}K^{\frac{3}{4}(d_m-1)}(r\,m)^3\left(d_m!^2\max_{1\leqslant j\leqslant r}\max_{1\leqslant k\leqslant m}inf_{pr+j}f_k\right)^{\frac{3}{2}}.$$

In the last inequality, the term $\sqrt{\frac{8}{\pi}}$ comes from the fact that G_j^p are standard gaussian which implies that $E\left(\left|G_j^p\right|^3\right) = \sqrt{\frac{8}{\pi}}$. Since the vectors X^p and G^p are centered, have the same covariance matrix and are independent from U_p and from V_j^p, then by putting the two inequalities together, we obtain:

$$\left|E\left(\varphi\left(Q(Z^{(p+1)})\right)\right) - E\left(\varphi\left(Q(Z^{(p)})\right)\right)\right|$$

$$\leqslant \left\|\varphi^{'''}\right\|_\infty \left(\beta + \sqrt{\frac{8}{\pi}}\right) K^{\frac{3}{4}(d_m-1)}(r\,m)^3 \left(d_m!^2 \max_{1\leqslant j\leqslant r} \max_{1\leqslant k\leqslant m} inf_{pr+j} f_k\right)^{\frac{3}{2}}. \tag{3.9}$$

Since $\sum_{j=1}^{r \max_i N_i} inf_j f_k = \frac{E\left((Q^k(X))^2\right)}{d_k!(d_k-1)!}$ then

$$\begin{aligned}
\sum_{p=1}^{\max_i N_i} \max_{1\leqslant j\leqslant r} \max_{1\leqslant k\leqslant m} inf_{pr+j} f_k &\leqslant \sum_{p=1}^{\max_i N_i} \sum_{j=1}^{r} \sum_{k=1}^{m} inf_{pr+j} f_k \\
&\leqslant \sum_{k=1}^{m} \sum_{j=1}^{r \max_i N_i} inf_j f_k \leqslant \sum_{k=1}^{m} \frac{E\left((Q^k(X))^2\right)}{d_k!(d_k-1)!} \leqslant \frac{m}{d_1!(d_1-1)!}.
\end{aligned}$$

By summing over p in (3.9), we finally obtain that:

$$\left|E\left(\varphi\left(Q(X)\right)\right) - E\left(\varphi\left(Q(G)\right)\right)\right|$$

$$\leqslant \left\|\varphi^{'''}\right\|_\infty \left(\beta + \sqrt{\frac{8}{\pi}}\right) K^{\frac{3}{4}(d_m-1)} r^3 m^4 \frac{d_m!^3}{d_1!(d_1-1)!} \sqrt{\max_{1\leqslant k\leqslant m} \max_{1\leqslant j\leqslant N_k} inf_j f_k}.$$

2. Let φ be a function of class C^3. We have

$$\left|E\left(\varphi(Q(X))\right) - E\left(\varphi(Z_V)\right)\right| \leqslant \left|E\left(\varphi(Q(X))\right) - E\left(\varphi(Q(G))\right)\right| + \left|E\left(\varphi(Q(G))\right) - E\left(\varphi(Z_V)\right)\right|.$$

For the first term we use the point 1 of lemma 3.2 to find an upper bound. For the second term we observe that the vector G have independent components, which allows us to use Theorem 7.2 in [2] to get the following inequality :

$$\left|E(\varphi(Q^1(X), \ldots, Q^m(X))) - E(\varphi(Z_V)\right| \leqslant \left\|\varphi^{''}\right\|_\infty \left(\sum_{i=1}^{m} \Delta_{i,i} + 2\sum_{1\leqslant i<j\leqslant m} \Delta_{i,j}\right)$$

$$+C\left\|\varphi^{'''}\right\|_\infty \sqrt{\frac{32}{\pi}} \left[\sum_{j=1}^{m} \left(\frac{64}{\pi}\right)^{\frac{d_j-1}{3}} d_j!\right]^3 \sqrt{\max_{1\leqslant k\leqslant m} \max_{1\leqslant j\leqslant N_k} inf_j f_k}.$$

The constant C is such that $\sum\limits_{i=1}^{\max\limits_k N_k} \max\limits_{1\leqslant j\leqslant m} inf_i f_j \leqslant C$ and since

$$\sum_{i=1}^{\max\limits_k N_k} \max_{1\leqslant j\leqslant m} inf_i f_j \leqslant \sum_{j=1}^{m}\sum_{i=1}^{\max\limits_k N_k} inf_i f_j \leqslant \sum_{j=1}^{m}\frac{E\left((Q^j(X))^2\right)}{d_j!(d_j-1)!} \leqslant \frac{m}{d_1!(d_1-1)!}$$

then we can choose the constant C equal to $\frac{m}{d_1!(d_1-1)!}$. Thus, we obtain

$$\left|E(\varphi(Q^1(X),\dots,Q^m(X))) - E(\varphi(Z_V)\right| \leqslant \left\|\varphi^{''}\right\|_\infty \left(\sum_{i=1}^{m}\Delta_{i,i} + 2\sum_{1\leqslant i<j\leqslant m}\Delta_{i,j}\right)$$

$$+\left\|\varphi^{'''}\right\|_\infty \sqrt{\frac{32}{\pi}}\left(\frac{64}{\pi}\right)^{d_m-1} m^4\frac{(d_m!)^3}{d_1!(d_1-1)!}\sqrt{\max_{1\leqslant k\leqslant m}\max_{1\leqslant j\leqslant N_k} inf_j f_k}.$$

■

Lemma 3.3 *Let the notations used in lemma 3.2 prevail. Consider the class $\mathcal{H}$ of indicator functions on measurable convex sets in $\mathbb{R}^m$. Let us define*

$$B_1 = \left(\sum_{i=1}^{m}\Delta_{i,i} + 2\sum_{1\leqslant i<j\leqslant m}\Delta_{i,j}\right)$$

$$B_2 = \frac{m^4 d_m!^3}{d_1!(d_1-1)!}\left(\left(\beta+\sqrt{\frac{8}{\pi}}\right)K^{\frac{3}{4}(d_m-1)}r^3 + \sqrt{\frac{32}{\pi}}\left(\frac{64}{\pi}\right)^{d_m-1}\right)\sqrt{\max_{1\leqslant k\leqslant m}\max_{1\leqslant j\leqslant N_k} inf_j f_k}$$

1. *Let us assume that the covariance matrix V is the m-dimensional identity matrix. Then*

$$\sup_{h\in\mathcal{H}(\mathbb{R}^m)}\left|E\left[h(Q^1(X),\dots,Q^m(X))\right] - E\left[h(Z_V)\right]\right| \leqslant \left(\frac{8}{3^{\frac{6}{7}}}+\frac{4}{3^{\frac{13}{7}}}\right)(5B_1+5B_2)^{\frac{1}{7}}m^{\frac{3}{7}}.$$

2. *Let us assume that the covariance matrix V is invertible and let $\Lambda = diag(\lambda_1,\dots,\lambda_k)$ be the diagonal matrix of the eigenvalues of V. Let B be an orthogonal matrix (i.e. $B^TB = I_m$ and $BB^T = I_m$) such that $V = B\Lambda B^T$, and let $b = max_{i,j}(\Lambda^{-\frac{1}{2}}B^T)$. Then*

$$\sup_{h\in\mathcal{H}(\mathbb{R}^m)}\left|E\left[h(Q^1(X),\dots,Q^m(X))\right] - E\left[h(Z_V)\right]\right| \leqslant \left(\frac{8}{3^{\frac{6}{7}}}+\frac{4}{3^{\frac{13}{7}}}\right)(5b^2B_1+5b^3B_2)^{\frac{1}{7}}m^{\frac{3}{7}}.$$

Proof. 1. Let us assume that the covariance matrix V is the m-dimensional identity matrix. Denote by Φ the standard normal distribution in $\mathbb{R}^m$, and by ϕ the corresponding density function. Consider $h \in \mathcal{H}(\mathbb{R}^m)$ and define the following function: $h_t(x) = \int_{\mathbb{R}^m} h(\sqrt{t}y+\sqrt{1-t}x)\Phi(dy)$, $\quad 0<t<1$. The key result is Lemma 2.11 in [5] which states that, for every probability measure Q on $\mathbb{R}^m$, every random variables $W \sim Q$ and $Z \sim \Phi$, and any $0 < t < 1$, we have

$$\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h(W)] - E[h(Z_V)]| \leqslant \frac{4}{3}\left[\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h_t(W)] - E[h_t(Z_V)]| + 2\sqrt{m}\sqrt{t}\right]. \tag{3.10}$$

Let us define $u(x,t,z) = (2\pi t)^{-\frac{m}{2}} exp\left(-\sum_{i=1}^{m} \frac{(z_i-\sqrt{1-t}x_i)^2}{2t}\right)$. Using the change of variable $z = \sqrt{t}y + \sqrt{1-t}x$ in $h_t(x)$ leads to

$$h_t(x) = \int_{\mathbb{R}^m} h(z)u(x,t,z)dz.$$

By the dominated convergence theorem, we may differentiate under the integral sign and obtain

$$\frac{\partial^2 h_t}{\partial x_i^2}(x) = -\frac{1-t}{t}\int_{\mathbb{R}^m} h(z)u(x,t,z)dz + \frac{1-t}{t^2}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)^2 u(x,t,z)dz.$$

Since $\|h\|_\infty \leqslant 1$ then we have

$$\left|\frac{\partial^2 h_t}{\partial x_i^2}(x)\right| \leqslant \frac{1-t}{t} + \frac{1-t}{t^2}\int_{\mathbb{R}^m} (z_i - \sqrt{1-t}x_i)^2 u(x,t,z)dz.$$

If $(Y_1,\ldots,Y_m)$ is a gaussian vector with covariance matrix tI_m then $\int_{\mathbb{R}^m}(z_i - \sqrt{1-t}x_i)u(x,t,z)dz = E(Y_i^2) = t$. Therefore, we have

$$\left|\frac{\partial^2 h_t}{\partial x_i^2}(x)\right| \leqslant 2\frac{1-t}{t}.$$

Furthermore, for $i \neq j$ we have

$$\frac{\partial^2 h_t}{\partial x_i \partial x_j}(x) = \frac{1-t}{t^2}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)(z_j - \sqrt{1-t}x_j)u(x,t,z)dz,$$

so that $\left|\frac{\partial^2 h_t}{\partial x_i \partial x_j}(x)\right| \leqslant \frac{1-t}{t^2}E(|Y_i|)E(|Y_i|) = \frac{2(1-t)}{\pi t}$. We conclude that $\|h''\|_\infty \leqslant \frac{2}{t} \leqslant \frac{5}{t^3}$. Similarly, for i, j, k in $[1, m]$ it holds that:

$$\left|\frac{\partial^3 h_t}{\partial x_i \partial x_j \partial x_k}(x)\right|$$
$$\leqslant \frac{(1-t)^{\frac{3}{2}}}{t^3} \max 3E(|Y_i|)t + E(|Y_i|^3); E(|Y_j|)t + E(|Y_i|^2)E(|Y_j|); E(|Y_i|)E(|Y_j|)E(|Y_k|).$$

Therefore $\|h'''\|_\infty \leqslant \frac{5}{t^3}$. Combining the latter inequality with the result (3.10) and point 2 of Lemma 3.2, we obtain

$$\begin{aligned}
&\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h(Q(X))] - E[h(Z_V)]| \\
&\leqslant \frac{4}{3}\left[\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h_t(Q(X))] - E[h_t(Z_V)]| + 2\sqrt{m}\sqrt{t}\right] \\
&\leqslant \frac{8}{3}\sqrt{m}\sqrt{t} + \frac{4}{3}(5B_1 + 5B_2)t^{-3}.
\end{aligned}$$

The function in the right-hand side of the inequality reaches its minimum at $t = \left(\frac{15(B_1+B_2)}{\sqrt{m}}\right)^{\frac{2}{7}}$, hence

$$\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h(Q(X))] - E[h(Z_V)]| \leqslant \left(\frac{8}{3^{\frac{6}{7}}} + \frac{4}{3^{\frac{13}{7}}}\right)(5B_1 + 5B_2)^{\frac{1}{7}} m^{\frac{3}{7}}.$$

2. Set $Q(X) = (Q^1(X), \ldots, Q^m(X))$. For any $h \in \mathcal{H}(\mathbb{R}^m)$, we have

$$E(h(Q(X))) - E(h(Z_v)) = E(h(B\Lambda^{\frac{1}{2}}\Lambda^{-\frac{1}{2}}B^T Q(X))) - E(h(B\Lambda^{\frac{1}{2}}\Lambda^{-\frac{1}{2}}B^T Z_v)).$$

Define $g(x) = h(B\Lambda^{\frac{1}{2}}x)$, $x \in \mathbb{R}^m$. Since $g \in \mathcal{H}(\mathbb{R}^m)$ then, using inequality (3.10), we get

$$\begin{aligned}
&\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h(Q(X))] - E[h(Z_V)]| \\
&\leqslant \sup_{g\in\mathcal{H}(\mathbb{R}^m)} \left|E\left[g(\Lambda^{-\frac{1}{2}}B^T Q(X))\right] - E\left[g(\Lambda^{-\frac{1}{2}}B^T Z_V)\right]\right| \\
&\leqslant \frac{4}{3}\left[\sup_{g\in\mathcal{H}(\mathbb{R}^m)} \left|E\left[g_t(\Lambda^{-\frac{1}{2}}B^T Q(X))\right] - E\left[g_t(\Lambda^{-\frac{1}{2}}B^T Z_V)\right]\right| + 2\sqrt{m}\sqrt{t}\right].
\end{aligned}$$

We can find an upper bound for the second and third derivatives of $f_t(x) = g_t(\Lambda^{-\frac{1}{2}}B^T x)$. Indeed, $\|f_t''\|_\infty \leqslant 5b^2t^{-3}$ and $\|f_t'''\|_\infty \leqslant 5b^3t^{-3}$. By using the same reasoning as in point 1 and replacing B_1 by b^2B_1 and B_2 by b^3B_2 in (3.10), we obtain the result. ∎

4 Proof of Theorem 1.1

We use hereafter the notation adopted in the beginning of Section 2. If we separate the diagonal terms from the non-diagonal terms in (1.2), we obtain

$$\begin{aligned}
\mathrm{trace}(X_n^d) - E(\mathrm{trace}(X_n^d)) &= \frac{1}{n^{\frac{d}{2}}} \sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} (X_{i_1,i_2}\ldots X_{i_d,i_1} - E(X_{i_1,i_2}\ldots X_{i_d,i_1})) \\
&+ \frac{1}{n^{\frac{d}{2}}} \sum_{((i_1,i_2),\ldots,(i_d,i_1))\in ND_n} X_{i_1,i_2}\ldots X_{i_d,i_1}.
\end{aligned}$$

The expectation in the second sum is equal to zero because the $X_{i,j}$ are independent and centered. The variance of the term containing the diagonal terms is upper bounded by $O\left(\frac{1}{\sqrt{n}}\right)$ and, therefore, goes to 0 as n goes to infinity. Indeed, if we set $M = \sup_{i,j} E(|X_{i,j}|^{2d})$, then

$$\mathrm{Var}\left(\frac{1}{n^{\frac{d}{2}}} \sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right)$$

$$= \frac{1}{n^d}\left[E\left(\left(\sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right)^2\right) - \left(E\left(\sum_{(i_1,i_2),\ldots,(i_d,i_1)\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right)\right)^2\right].$$

Keeping the notation introduced in lemma 2.1, we have:

$$E\left(\left(\sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right)^2\right)$$

$$= \sum_{((i_1,i_2),\ldots,(i_d,i_1),(i_{d+1},i_{d+2}),\ldots,(i_{2d},i_{d+1}))\in A_n} E\left(X_{i_1,i_2}\ldots X_{i_d,i_1}X_{i_{d+1},i_{d+2}}\ldots X_{i_{2d},i_{d+1}}\right).$$

Since $E\left(X_{i_1,i_2}\ldots X_{i_d,i_1}X_{i_{d+1},i_{d+2}}\ldots X_{i_{2d},i_{d+1}}\right)$ is equal to 1 over the subset B_n of A_n, and is upper bounded by M over the subset $A_n \setminus B_n$, then we can state that:

$$\left|E\left(\left(\sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right)^2\right) - \#B_n\right| \leqslant M\#(A_n \setminus B_n). \tag{4.11}$$

Furthermore, since the $X_{i,j}$ are centered and independent, then $E(X_{i_1,i_2}\ldots X_{i_d,i_1}) = 0$ if $((i_1,i_2),\ldots,(i_d,i_1)) \in D_n \setminus E_n$. Thus,

$$E\left(\sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right) = \sum_{((i_1,i_2),\ldots,(i_d,i_1))\in E_n} E(X_{i_1,i_2}\ldots X_{i_d,i_1}).$$

On the other hand, $E(X_{i_1,i_2}\ldots X_{i_d,i_1})$ is equal to 1 over the subset F_n of E_n, and bounded by $\sqrt{M}$ over $E_n \setminus F_n$. Then,

$$\left|E\left(\sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right) - \#F_n\right| \leqslant \sqrt{M}\#(E_n \setminus F_n). \tag{4.12}$$

Finally, by combining the estimations (4.11) and (4.12), and using points (i) to (iv) of Lemma 2.1 and the fact that $\#D_n = O(n^{d-1})$, we get the following result, with Z_n defined by $Z_n = \sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}$:

$$\begin{aligned}\mathrm{Var}\,(Z_n) &= E(Z_n^2) - E(Z_n)^2 \\ &= \#B_n + \big(E(Z_n^2) - \#B_n\big) - (\#F_n)^2 + \big((\#F_n)^2 - E(Z_n)^2 \\ &= \#B_n - (\#F_n)^2 + \big(E(Z_n^2) - \#B_n\big) + \big(\#F_n - E(Z_n)\big)\big(\#F_n + E(Z_n)\big).\end{aligned}$$

From points (ii) and (iv) of Lemma 2.1, it follows that $\#B_n - (\#F_n)^2 = O(n^{d-1})$. Using point (i) and the relation (4.11), we obtain the estimation $E(Z_n^2) - \#B_n = O(n^{d-1})$. Finally, using points $(iii) - (iv)$ and the relation (4.12), we get the following estimations: $E(Z_n) - \#F_n = O(n^{\frac{d-1}{2}})$ and $E(Z_n) + \#F_n = O(n^{\frac{d}{2}})$. From these estimations, we conclude that:

$$\mathrm{Var}\left(\frac{1}{n^{\frac{d}{2}}}\sum_{((i_1,i_2),\ldots,(i_d,i_1))\in D_n} X_{i_1,i_2}\ldots X_{i_d,i_1}\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

Consider now a bijection $\sigma : [1,n^2] \to [1,n]\times[1,n]$. Let us define $X_i = X_{\sigma(i)}$ and $R = E(Re(X_i)^2)$. When $R = 1$, the X_i are real-valued, which corresponds exactly to the result of Nourdin et Peccati [1] (there is then nothing more to prove). By contrast, when $R = 0$, the X_i are purely imaginary-valued; factoring out by i^d in the trace formula shows that the result in this case can be derived from the case $R = 1$. In what follows, we can then freely assume that $R \in (0,1)$. Set $\rho = \frac{E(Re(X_i)Im(X_i))}{R\sqrt{1-R}}$, and define:

$$\begin{cases} X_i^0 = Re(X_i) - \rho\sqrt{\frac{R}{1-R}}\,Im(X_i) \\ X_i^1 = Im(X_i) \\ X_{i,j}^0 = Re(X_{i,j}) - \rho\sqrt{\frac{R}{1-R}}\,Im(X_{i,j}) \\ X_{i,j}^1 = Im(X_{i,j}) \end{cases},$$

$$f_n(i_1,\ldots,i_d) = \frac{1}{n^{\frac{d}{2}}}\mathbf{1}_{\{(\sigma(i_1),\ldots,\sigma(i_d))\in ND_n\}},$$

and

$$Q_d(f_n,\,X) = \sum_{i_1,\ldots,i_d=1}^{n^2} f_n(i_1,\ldots,i_d)\,X_{i_1}\ldots X_{i_d}.$$

We have:

$$X_{i_1}\ldots X_{i_d} = \prod_{k=1}^{d}\left(X_{i_k}^0 + \rho\sqrt{\frac{R}{1-R}}X_{i_k}^1 + iX_{i_k}^1\right) = \prod_{k=1}^{d}\left(X_{i_k}^0 + \left(\rho\sqrt{\frac{R}{1-R}} + i\right)X_{i_k}^1\right).$$

Hence

$$X_{i_1}\dots X_{i_d}=\sum_{j_1,\dots,j_d\in\{0,1\}^d}\left(i+\rho\sqrt{\frac{R}{1-R}}\right)^{\sum j_k}X_{i_1}^{j_1}\dots X_{i_d}^{j_d},$$

which yields

$$\begin{aligned}Q_d(f_n,\,X) &=\sum_{k=0}^{d}\left(i+\rho\sqrt{\frac{R}{1-R}}\right)^{k}\sum_{\substack{(j_1,\dots,j_d)\in\{0,1\}^d\\ j_1+\dots+j_d=k}}\sum_{(i_1,\dots,i_d)\in[1,n^2]^d}f_n(i_1,\dots,i_d)X_{i_1}^{j_1}\dots X_{i_d}^{j_d}\\ &=\sum_{k=0}^{d}\left(i+\rho\sqrt{\frac{R}{1-R}}\right)^{k}\sum_{\substack{(j_1,\dots,j_d)\in\{0,1\}^d\\ j_1+\dots+j_d=k}}\sum_{\left((i_1,i_2),\dots,(i_d,i_1)\right)\in ND_n}\frac{1}{n^{\frac{d}{2}}}X_{i_1,i_2}^{j_1}\dots X_{i_d,i_1}^{j_d}.\end{aligned}$$

We define for any two elements $(i_1,\dots,i_d)$, $(j_1,\dots,j_d)$ of $[1,n]^d$, and $(p_1,\dots,p_d)\in\{0,1\}^d$ the quantity $g_n^k\left[((i_1,j_1),p_1),\dots((i_d,j_d),p_d)\right]$ as follows: $g_n^k\left[((i_1,j_1),p_1),\dots((i_d,j_d),p_d)\right]=\frac{1}{n^{\frac{d}{2}}}$ if $((i_1,j_1),\dots,(i_d,j_d))\in ND_n$ and $\sum_{l=1}^{d}p_l=k$, and $g_n^k\left[((i_1,j_1),p_1),\dots((i_d,j_d),p_d)\right]=0$ otherwise. Set $R_0=\sqrt{\mathrm{Var}(X_i^0)}$, $R_1=\sqrt{\mathrm{Var}(X_i^1)}$, and $Y=(Y_{i,j}^k)_{\substack{(i,j)\in[1,n]^2\\ k\in\{0,1\}}}$ a family of random variables defined by $Y_{i,j}^k=\frac{X_{i,j}^k}{R_k}$. Then

$$\begin{aligned}Q_d(g_n^k,Y) &=\sum_{\substack{(x_1,\dots,x_d)\in[1,n]^d\\ (y_1,\dots,y_d)\in[1,n]^d\\ (p_1,\dots,p_d)\in\{0,1\}^d}}g_n^k\left[((x_1,y_1),p_1),\dots,((x_d,y_d),p_d)\right]Y_{x_1,y_1}^{p_1}\dots Y_{x_d,y_d}^{p_d}\\ &=\frac{1}{(R_0)^{d-k}(R_1)^k}\sum_{\substack{(j_1,\dots,j_d)\in\{0,1\}^d\\ \sum j_p=k}}\sum_{\left((i_1,i_2),\dots,(i_d,i_1)\right)\in ND_n}\frac{1}{n^{\frac{d}{2}}}X_{i_1,i_2}^{j_1}\dots X_{i_d,i_1}^{j_d}.\end{aligned}$$

We can then conclude that

$$Q_d(f_n,\,X)=\sum_{k=0}^{d}\left(i+\rho\sqrt{\frac{R}{1-R}}\right)^{k}(R_0)^{d-k}(R_1)^kQ_d(g_n^k,Y).$$

If $\tilde{g_n^k}$ stands for the symmetrization of g_n^k then $Q_d(\tilde{g_n^k},Y)=Q_d(g_n^k,Y)$, where $\tilde{g_n^k}=\sum_{\sigma\in S_d}g_n^{k,\sigma}$ and $g_n^{k,\sigma}\left[((x_1,y_1),p_1),\dots,((x_d,y_d),p_d)\right]=g_n^k\left[((x_{\sigma(1)},y_{\sigma(1)}),p_{\sigma(1)}),\dots,((x_{\sigma(d)},y_{\sigma(d)}),p_{\sigma(d)})\right]$. To establish that $Q_d(f_n,\,X)$ converge in law to the variable $Z_d^1+iZ_d^2$ where $Z_d=(Z_d^1,Z_d^2)$ is a gaussian vector, it is sufficient to show that the $Q_d(g_n^k,Y)$, $k=0,\dots,d$, converge in law to a gaussian vector having independent components. Using part 2 of Lemma 3.3 (in the particular case $r=2$), we show that $Q_d(g_n^k,Y)$ converges in law to a gaussian vector whose covariance matrix V is given by $V(k,k') = \lim_{\infty} E(Q_d(g_n^k,Y)Q_d(g_n^{k'},Y))$. To do so, it is sufficient to check the assumptions of Lemma 3.3, that is, (i) $\max_{i=1,\dots,2N} inf_{(a,b),p}\, \tilde{g_N^k} \to 0$, (ii) for every $1 \leqslant s \leqslant d-1$, $\left\|\tilde{g_N^k} \star_s \tilde{g_N^k}\right\|_2 \to 0$, (iii) $E(Q_d(g_n^k,Y)Q_d(g_n^{k'},Y) \to \delta_{i,j}$ (with $\delta_{i,j}$ the Kronecker symbol), and (iv) $E(Q_d(g_n^k,Y)^2) \to \sigma^2$. We can rewrite $Q_d(g_n^k,Y)$ as

$$Q_d(g_n^k,Y) = \frac{1}{n^{\frac{d}{2}}} \sum_{\substack{(j_1,\dots,j_d)\in\{0,1\}^n \\ j_1+\dots+j_d=k}} \sum_{(i_1,\dots,i_d)\in ND_n} Y_{i_1,i_2}^{j_1} \dots Y_{i_d,i_1}^{j_d}.$$

The second-order moment of $Q_d(g_n^k,Y)$ is equal to

$$\frac{1}{n^d} \sum_{\substack{(j_1,\dots,j_{2d})\in\{0,1\}^n \\ j_1+\dots+j_d=j_{d+1}+\dots+j_{2d}=k}} \sum_{\substack{\left((i_1,i_2),\dots,(i_d,i_1)\right)\in ND_n \\ \left((i_{d+1},i_{d+2}),\dots,(i_{2d},i_{d+1})\right)\in ND_n}} E(Y_{i_1,i_2}^{j_1} \dots Y_{i_d,i_1}^{j_d} Y_{i_{d+1},i_{d+2}}^{j_{d+1}} \dots Y_{i_{2d},i_{d+1}}^{j_{2d}}). \tag{4.13}$$

For the expectation corresponding to the indices $i_1,\dots,i_{2d},j_1,\dots,j_{2d}$ in (4.13) to be different from zero, it must hold that $(i_1,\dots,i_{2d})$ belongs to G_n, where G_n has been defined in Lemma 2.1. Furthermore, since the subset $G_n \setminus H_n$ is of cardinality $O(n^{d-1})$, its contribution to the moment of order 2 of $Q_d(g_n^k,Y)$ is $O(\frac{1}{n})$. It remains then to see what happens when $(i_1,\dots,i_{2d})$ belongs to H_n. In this case, let us recall from the proof of point (vi) of Lemma 2.1 that the elements of the set H_n are completely characterized by d given pairwise distinct integers $i_1,\dots,i_d \in [1,n]$ and a given integer $k \in [1,d]$ such that $(i_{d+1},\dots,i_{2d}) = (i_k,\dots,i_d,i_1,\dots,i_{k-1})$. Moreover, if the expectation corresponding to the indices $i_1,\dots,i_{2d},j_1,\dots,j_{2d}$ in (4.13) is different from zero, then it must hold that $(j_{d+1},\dots,j_{2d}) = (j_k,\dots,j_d,j_1,\dots,j_{k-1})$ and this expectation is equal to 1. Thus,

$$\begin{aligned} E\left(Q_d(g_n^k,Y)^2\right) &= \frac{1}{n^d} \sum_{\substack{(j_1,\dots,j_d)\in\{0,1\}^n \\ j_1+\dots+j_d=k}} d \times n \dots \times (n-d+1) + O\left(\frac{1}{n}\right) \\ &= \frac{dC_d^k \times n \dots \times (n-d+1)}{n^d} + O(\frac{1}{n}), \end{aligned}$$

which yields $E\left(Q_d(g_n^k,Y)^2\right) \underset{n\to\infty}{\longrightarrow} dC_d^k$. Moreover, $E(Q_d(g_N^k,\,Y)Q_d(g_N^j,\,Y))$ is equal to

$$\frac{1}{n^d} \sum_{\substack{(j_1,\dots,j_{2d})\in\{0,1\}^n \\ j_1+\dots+j_d=k,j_{d+1}+\dots+j_{2d}=j}} \sum_{\substack{\left((i_1,i_2),\dots,(i_d,i_1)\right)\in ND_n \\ \left((i_{d+1},i_{d+2}),\dots,(i_{2d},i_{d+1})\right)\in ND_n}} E(Y_{i_1,i_2}^{j_1} \dots Y_{i_d,i_1}^{j_d} Y_{i_{d+1},i_{d+2}}^{j_{d+1}} \dots Y_{i_{2d},i_{d+1}}^{j_{2d}}).$$

Similarly to the computation of the second-order moment of $Q_d(g_n^k,Y)$, the set of elements for which the expectation in (4.14) is different from zero is the set G_n of Lemma

2.1. The subset $G_n \setminus H_n$ is of cardinality $O(n^{d-1})$, which implies that its contribution to $E(Q_d(g_N^k, Y)Q_d(g_N^j, Y))$ is $O(\frac{1}{n})$. Furthermore, the elements of the set H_n are characterized by d given pairwise distinct integers $i_1, \ldots, i_d \in [1, n]$ and a given integer $k \in [1, d]$ such that $(i_{d+1}, \ldots, i_{2d}) = (i_k, \ldots, i_d, i_1, \ldots, i_{k-1})$. Moreover, for $E(Y^{j_1}_{i_1,i_2} \ldots Y^{j_d}_{i_d,i_1} Y^{j_{d+1}}_{i_{d+1},i_{d+2}} \ldots Y^{j_{2d}}_{i_{2d},i_{d+1}})$ to be different from zero, it must hold that $(j_{d+1}, \ldots, j_{2d}) = (j_k, \ldots, j_d, j_1, \ldots, j_{k-1})$, which is impossible in the case $j \neq k$. We conclude that $E(Q_d(g_N^k, Y)Q_d(g_N^j, Y)) \to 0$ for every $j \neq k$.

From the definition of $\tilde{g_n^k}$, it is clear that $\left\|\tilde{g_n^k}\right\|_\infty \leqslant \left\|g_n^k\right\|_\infty \leqslant \frac{1}{n^{\frac{d}{2}}}$. Then,

$$\begin{aligned}
inf_{(a,b),p}\, \tilde{g_N^k} &= \sum_{\substack{(x_1,\ldots,x_{d-1})\in[1,n]^{d-1}\\(y_1,\ldots,y_{d-1})\in[1,n]^{d-1}\\(p_1,\ldots,p_{d-1})\in\{0,1\}^{d-1}}} \tilde{g_n^k}\,(((a,b),p),((x_1,y_1),p_1),\ldots,((x_{d-1},y_{d-1}),p_{d-1}))^2\\
&\leqslant \sum_{\substack{(i_1,\ldots,i_d)\in[1,n]^d\\(p_1,\ldots,p_d)\in\{0,1\}^d\\\sigma\in\mathfrak{S}_d}} g_n^k\left(((i_{\sigma(1)},i_{\sigma(1)+1}),p_{\sigma(1)}),\ldots,((i_{\sigma(d)},i_{\sigma(d)+1}),p_{\sigma(d)})\right)^2\\
&\times 1_{\{a=i_{\sigma(1)}\}} \times 1_{\{b=i_{\sigma(1)+1}\}} \times 1_{\{p=p_{\sigma(1)}\}}\\
&\leqslant 2^{d-1}(d)!n^{d-2}\left\|\tilde{g_n^k}\right\|_\infty^2 \leqslant \frac{2^{d-1}(d)!}{n^2}.
\end{aligned}$$

Therefore $\max\limits_{i=1,\ldots,2N} inf_{(a,b),p}\, \tilde{g_N^k} \leqslant \frac{2^{d-1}(d)!}{n^2} \to 0.$

Now, let $1 \leqslant s \leqslant d-1$ and $\sigma_1, \sigma_2 \in \mathfrak{S}_d$. Then

$$\begin{aligned}
&g_n^{k,\sigma_1} \star_s g_n^{k,\sigma_2}\left[((x_1,y_1),p_1),\ldots,((x_{d-s},y_{d-s}),p_{d-s}),((x'_1,y'_1),p'_1),\ldots,((x'_{d-s},y'_{d-s}),p'_{d-s})\right]\\
&= \sum_{\substack{(x_{d-s+1},\ldots,x_d)\in[1,n]^s\\(y_{d-s+1},\ldots,y_d)\in[1,n]^s\\(p_{d-s+1},\ldots,p_d)\in\{0,1\}^s}}\\
&g_n^{k,\sigma_1}\left[((x_1,y_1),p_1),\ldots,((x_{d-s},y_{d-s}),p_{d-s}),((x_{d-s+1},y_{d-s+1}),p_{d-s+1}),\ldots,((x_d,y_d),p_d)\right]\\
&\times g_n^{k,\sigma_2}\left[((x'_1,y'_1),p'_1),\ldots,((x'_{d-s},y'_{d-s}),p'_{d-s}),((x_{d-s+1},y_{d-s+1}),p_{d-s+1}),\ldots,((x_d,y_d),p_d)\right]
\end{aligned}$$

so that

$$\left\|\tilde{g_n^k}\star_s \tilde{g_n^k}\right\|_2^2$$

$$= \frac{1}{(d!)^4}\sum_{\sigma_1,\sigma_2,\sigma_3,\sigma_4\in\mathfrak{S}_d}\sum_{\substack{(x_1,\ldots,x_{d-s})\in[1,n]^{d-s}\\(y_1,\ldots,y_{d-s})\in[1,n]^{d-s}\\(p_1,\ldots,p_{d-s})\in\{0,1\}^{d-s}}}\sum_{\substack{(x'_1,\ldots,x'_{d-s})\in[1,n]^{d-s}\\(y'_1,\ldots,y'_{d-s})\in[1,n]^{d-s}\\(p'_1,\ldots,p'_{d-s})\in\{0,1\}^{d-s}}}\sum_{\substack{(x_{d-s+1},\ldots,x_d)\in[1,n]^s\\(y_{d-s+1},\ldots,y_d)\in[1,n]^s\\(p_{d-s+1},\ldots,p_d)\in\{0,1\}^s}}\sum_{\substack{(x'_{d-s+1},\ldots,x'_d)\in[1,n]^s\\(y'_{d-s+1},\ldots,y'_d)\in[1,n]^s\\(p'_{d-s+1},\ldots,p'_d)\in\{0,1\}^s}}$$

$$\times g_n^{k,\sigma_1}\left[((x_1,y_1),p_1),\ldots,((x_{d-s},y_{d-s}),p_{d-s}),((x_{d-s+1},y_{d-s+1}),p_{d-s+1}),\ldots,((x_d,y_d),p_d)\right]$$
$$\times g_n^{k,\sigma_2}\left[((x'_1,y'_1),p'_1),\ldots,((x'_{d-s},y'_{d-s}),p'_{d-s}),((x_{d-s+1},y_{d-s+1}),p_{d-s+1}),\ldots,((x_d,y_d),p_d)\right]$$
$$\times g_n^{k,\sigma_3}\left[((x_1,y_1),p_1),\ldots,((x_{d-s},y_{d-s}),p_{d-s}),((x'_{d-s+1},y'_{d-s+1}),p'_{d-s+1}),\ldots,((x'_d,y'_d),p'_d)\right]$$
$$\times g_n^{k,\sigma_4}\left[((x'_1,y'_1),p'_1),\ldots,((x'_{d-s},y'_{d-s}),p'_{d-s}),((x'_{d-s+1},y'_{d-s+1}),p'_{d-s+1}),\ldots,((x'_d,y'_d),p'_d)\right].$$

For the sake of notational simplicity and because this case is representative of the difficulty, in the rest of the proof we assume that $\sigma_1=\sigma_2=\sigma_3=\sigma_4=I_d$, where I_d stands for the identity permutation over $[1,d]$. Since g_n^k is equal to zero at point $[((x_1,y_1),p_1),\ldots,((x_d,y_d),p_d)]$ if $y_i\neq x_{i+1}$ or $y_d\neq x_1$, then

$$\sum_{\substack{(x_1,\ldots,x_{d-s})\in[1,n]^{d-s}\\(y_1,\ldots,y_{d-s})\in[1,n]^{d-s}\\(p_1,\ldots,p_{d-s})\in\{0,1\}^{d-s}}}\sum_{\substack{(x'_1,\ldots,x'_{d-s})\in[1,n]^{d-s}\\(y'_1,\ldots,y'_{d-s})\in[1,n]^{d-s}\\(p'_1,\ldots,p'_{d-s})\in\{0,1\}^{d-s}}}\sum_{\substack{(x_{d-s+1},\ldots,x_d)\in[1,n]^s\\(y_{d-s+1},\ldots,y_d)\in[1,n]^s\\(p_{d-s+1},\ldots,p_d)\in\{0,1\}^s}}\sum_{\substack{(x'_{d-s+1},\ldots,x'_d)\in[1,n]^s\\(y'_{d-s+1},\ldots,y'_d)\in[1,n]^s\\(p'_{d-s+1},\ldots,p'_d)\in\{0,1\}^s}}$$

$$\times g_n^{k,I_d}\left[((x_1,y_1),p_1),\ldots,((x_{d-s},y_{d-s}),p_{d-s}),((x_{d-s+1},y_{d-s+1}),p_{d-s+1}),\ldots,((x_d,y_d),p_d)\right]$$
$$\times g_n^{k,I_d}\left[((x'_1,y'_1),p'_1),\ldots,((x'_{d-s},y'_{d-s}),p'_{d-s}),((x_{d-s+1},y_{d-s+1}),p_{d-s+1}),\ldots,((x_d,y_d),p_d)\right]$$
$$\times g_n^{k,I_d}\left[((x_1,y_1),p_1),\ldots,((x_{d-s},y_{d-s}),p_{d-s}),((x'_{d-s+1},y'_{d-s+1}),p'_{d-s+1}),\ldots,((x'_d,y'_d),p'_d)\right]$$
$$\times g_n^{k,I_d}\left[((x'_1,y'_1),p'_1),\ldots,((x'_{d-s},y'_{d-s}),p'_{d-s}),((x'_{d-s+1},y'_{d-s+1}),p'_{d-s+1}),\ldots,((x'_d,y'_d),p'_d)\right]$$

$$=\sum_{\substack{(\alpha_1,\ldots,\alpha_{d-s+1})\in[1,n]^{d-s+1}\\(p_1,\ldots,p_{d-s})\in\{0,1\}^{d-s}}}\sum_{\substack{(\alpha'_1,\ldots,\alpha'_{d-s+1})\in[1,n]^{d-s+1}\\(p'_1,\ldots,p'_{d-s})\in\{0,1\}^{d-s}}}\sum_{\substack{(i_1,\ldots,i_{s-1})\in[1,n]^{s-1}\\(p_{d-s+1},\ldots,p_d)\in\{0,1\}^s}}\sum_{\substack{(i'_1,\ldots,i'_{s-1})\in[1,n]^{s-1}\\(p'_{d-s+1},\ldots,p'_d)\in\{0,1\}^s}}$$

$$\times 1_{\{\alpha_1=\alpha'_1\}}\times 1_{\{\alpha_{d-s+1}=\alpha'_{d-s+1}\}}$$
$$\times g_n^k\left[((\alpha_1,\alpha_2),p_1),\ldots,((\alpha_{d-s},\alpha_{d-s+1}),p_{d-s}),((\alpha_{d-s+1},i_1),p_{d-s+1}),\ldots,((i_{s-1},x_1),p_d)\right]$$
$$\times g_n^k\left[((\alpha'_1,\alpha'_2),p'_1),\ldots,((\alpha'_{d-s},\alpha'_{d-s+1}),p'_{d-s}),((\alpha'_{d-s+1},i_1),p_{d-s+1}),\ldots,((i_{s-1},x'_1),p_d)\right]$$
$$\times g_n^k\left[((\alpha_1,\alpha_2),p_1),\ldots,((\alpha_{d-s},\alpha_{d-s+1}),p_{d-s}),((\alpha_{d-s+1},i'_1),p'_{d-s+1}),\ldots,((i'_{s-1},x_1),p'_d)\right]$$
$$\times g_n^k\left[((\alpha'_1,\alpha'_2),p'_1),\ldots,((\alpha'_{d-s},\alpha'_{d-s+1}),p'_{d-s}),((\alpha'_{d-s+1},i'_1),p'_{d-s+1}),\ldots,((i'_{s-1},x'_1),p'_d)\right]$$
$$\leqslant 2^{2d}n^{2d-2}\left\|g_n^k\right\|_\infty^4\leqslant 2^{2d}n^{-2}.$$

We conclude that $\left\|\tilde{g_n^k}\star_s \tilde{g_n^k}\right\|_2^2\to 0$.

Then, all the assumptions of Lemma 3.3 are fulfilled by $Q_d(g_N^k,Y)$. Therefore,

$$Q_d(f_N,Y)\xrightarrow{law}\sum_{k=1}^d\left(i+\rho\sqrt{\frac{R}{1-R}}\right)^k R_0^k\,R_1^{d-k}\sqrt{dC_d^k}G_k,$$

where the G_k's are independent standard gaussian random variables. We can rewrite this result as:

$$Q_d(f_N, Y) \xrightarrow{law} Z_d^1 + iZ_d^2,$$

where $Z_d = (Z_d^1, Z_d^2)$ is a gaussian vector; its covariance matrix is $\begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix}$, with $d = \sigma_1^2 + \sigma_2^2$ and $dE(X_1^2)^d = \sigma_1^2 - \sigma_2^2 + i\,2\sigma_{1,2}$.

This completes the proof of Theorem 1.1. ■

Acknowledgments. This work is part of my forthcoming PhD dissertation. I am extremely grateful to my advisor Ivan Nourdin for suggesting this topic, as well as for many advices and encouragements.

References

[1] I. Nourdin and G. Peccati (2008). Universal Gaussian fluctuations of non-Hermitian matrix ensembles. *ALEA* **7**, 341-375.

[2] I. Nourdin, G. Peccati and G. Reinert (2010). Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *Ann. Probab.* **38** (5), 1947-1985.

[3] D. Nualart and G. Peccati (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (1), 177-193.

[4] B. Rider and J.W. Silverstein (2006). Gaussian Fluctuations for non-hemitian random matrix ensembles. *Ann. Probab.* **34** (6), 2118-2143.

[5] F. Goetze (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19**, 724-739.