HIGH DIMENSIONAL REGIMES OF NON-STATIONARY GAUSSIAN CORRELATED WISHART MATRICES

SOLESNE BOURGUIN† AND THANH DANG†

Abstract. We study the high-dimensional asymptotic regimes of correlated Wishart matrices $d^{-1}YY^T$, where $Y$ is a $n \times d$ Gaussian random matrix with correlated and non-stationary entries. We prove that under different normalizations, two distinct regimes emerge as both $n$ and $d$ grow to infinity. The first regime is the one of central convergence, where the law of the properly renormalized Wishart matrices becomes close in Wasserstein distance to that of a Gaussian orthogonal ensemble matrix. In the second regime, a non-central convergence happens, and the law of the normalized Wishart matrices becomes close in Wasserstein distance to that of the so-called Rosenblatt-Wishart matrix recently introduced by Nourdin and Zheng. We then proceed to show that the convergences stated above also hold in a functional setting, namely as weak convergence in $C([a, b]; M_n(\mathbb{R}))$. As an application of our main result (in the central convergence regime), we show that it can be used to prove convergence in expectation of the empirical spectral distributions of the Wishart matrices to the semicircular law. Our findings complement and extend a rich collection of results on the study of the fluctuations of Gaussian Wishart matrices, and we provide explicit examples based on Gaussian entries given by normalized increments of a bi-fractional or a sub-fractional Brownian motion.

1. Introduction and main results

Random matrix theory plays a fundamental role in many areas of mathematics, either theoretical ones such as non-commutative algebra, combinatorics, geometry or spectral analysis, or applied ones such as statistical physics, signal processing or multivariate analysis and statistical theory. In the latter, one type of random matrices are particularly important as they are used to model, for example, sample covariance matrices (see e.g., the surveys [9, 18, 26]), which in the era of data driven analysis have now a growing importance in practice. This type of matrices are called Wishart matrices and have been introduced by the statistician John Wishart in [30]. Given an underlying $n \times d$ random matrix $Y$, the associated Wishart matrix is given by $d^{-1}YY^T$ and is hence a symmetric $n \times n$ random matrix. In this paper, we are interested in the asymptotic behaviour of such large Wishart matrices as both dimensions grow to infinity. The case where both dimensions are allowed to grow in our context is fundamental, especially when one interprets the growth of the dimensions as underlying data sets becoming very large over time if the Wishart matrices are seen as being sample covariance matrices for instance. The case where $n$ is taken to be fixed and only $d$ grows to infinity (called the one-dimensional regime) is well understood via standard probabilistic results such as the laws of large numbers, but this setting represents nowadays a drawback and is not realistic enough in applications when considering how common it has become to have increasingly bigger data sets that grow with

†Boston University, Department of Mathematics and Statistics, 111 Cummington Mall, Boston, MA 02215, USA

E-mail addresses: bourguin@math.bu.edu, ycloud77@bu.edu.

2010 Mathematics Subject Classification. 60B20, 60F05, 60H07, 60G22.

Key words and phrases. Wishart matrices; high-dimensional regime; self-similar Gaussian processes; non-stationarity; Malliavin calculus; empirical spectral distributions; semicircular distribution; functional limit theorems.

S. Bourguin was supported in part by the Simons Foundation grant 635136.
time as one continuously collects new data. The high-dimensional regime, i.e., the case where both \( n \) and \( d \) grow to infinity, possibly at different paces, is much more difficult to apprehend and has triggered many different and complementary lines of work in the recent years (see e.g., [4, 5, 7, 8, 11, 15, 17, 20, 21, 24, 27]).

One particular setting that has received more attention than others is the one where the matrix \( Y \) has Gaussian entries, which is the most common case to arise in applications. More specifically, the case where the Gaussian entries are increments of Gaussian processes (such as a Brownian or fractional Brownian motion) is of particular interest when considering the modeling of systems that evolve in time. When all the entries of \( Y \) are considered to be independent and coming from a stationary Gaussian process, the fluctuations of the associated Wishart matrix have been studied in [7, 8]. The assumption of full independence of the entries has then been relaxed in [24] where the authors assume either row independence, with a possible correlation in each row separately, or overall correlation (depending on the setting), but keeping the assumption of stationarity of the underlying Gaussian process.

Our goal in this paper is to study the high-dimensional fluctuations of Wishart matrices based on Gaussian non-stationary entries with a self-similarity property, and hence relax the stationarity assumption made in [24] to allow for Gaussian entries coming from non-stationary Gaussian processes, such as bi-fractional or sub-fractional Brownian motions. We use the class of processes introduced in [16] and further studied in [10] characterized by the fact that the covariance function of the Gaussian processes that are members of this class satisfy hypotheses (H.1) and (H.2) stated below (the covariance function of the processes is a perturbation of the covariance function of a fractional Brownian motion, and includes among other important examples, bi-fractional Brownian motion, or sub-fractional Brownian motion).

We also address a question that has not been studied so far in this context to the best of our knowledge, namely that of functional convergence when the Wishart matrices are seen as matrix-valued processes. Deriving functional versions of limit theorems and convergence results is of utmost interest, especially when considering systems that naturally describe phenomena evolving in time, as illustrated by the fast-growing literature on this topic (see for instance [10, 22] for functional limit theorems related to the celebrated Breuer-Major theorem, [1, 3] for a quantitative approach based on Stein’s method in Banach spaces, among many other references). Our (non-functional) convergence results ensuring the convergence in the sense of finite-dimensional distributions, we prove that the sequences of Wishart matrices we consider are tight in \( C([a,b]; M_n(\mathbb{R})) \). We consider the indexing parameter of such matrix-valued processes to be part of the \( d \) dimension in the form of a dependency of this dimension on it by replacing \( d \) by \([dx]\), where \( x \) is the indexing parameter of the matrix process. An applied way of looking at this setting is to interpret \( x \) as time, and considering that the \( d \) dimension grows continuously with time, which is a very natural setting in many applications (such as financial time series or temperature readings for instance).

As an application of our results, we prove the convergence in expectation of the empirical spectral distributions of these Wishart matrices to the semicircular law (in the case where the Wishart matrices exhibit a central limit behavior), which is a central question in random matrix theory. Our methodology to derive this result complements the more classical methods such as the characteristic function approach, the method of moments or the Stieltjes transform.

Let us describe the class of processes we will be working with. Let \( X \) be a member of the
class of self-similar processes introduced in [16]. For reference, a stochastic process \( \{ X_s : s \geq 0 \} \) is called a self-similar process with self-similarity parameter \( H > 0 \) if for all \( c > 0 \),

\[
\{ X(cs) : s \geq 0 \} \overset{\text{dist}}{=} \{ c^H X(s) : s \geq 0 \},
\]

where \( \overset{\text{dist}}{=} \) denotes equality in distribution. In our case, \( \{ X_s : s \geq 0 \} \) is a centered, self-similar Gaussian process with self-similarity parameter \( \beta \in (0, 1) \). Define \( \phi : [1, \infty) \to \mathbb{R} \) by \( \phi(x) = \mathbb{E}[X_1 X_s] \), so that, for \( 0 < s \leq t \), we have

\[
\mathbb{E}[X_s X_t] = s^{2\beta} \mathbb{E}[X_1 X_\frac{t}{s}] = s^{2\beta} \phi \left( \frac{t}{s} \right).
\]

Hence, \( \phi \) characterizes the covariance function of \( X \). Moreover, the following two assumptions are assumed to hold for all members of this class of processes, and were both introduced in [16].

**\( (H.1) \)** There exists \( \alpha \in (0, 2\beta] \) such that \( \phi \) has the form

\[
\phi(x) = -\lambda (x - 1)^\alpha + \psi(x),
\]

where \( \lambda > 0 \) and \( \psi \) is twice-differentiable on an open set containing \( [1, \infty) \) and there exists a constant \( C > 0 \) such that, for any \( x \in (1, \infty) \),

(a) \( |\psi'(x)| \leq Cx^{\alpha-1} \)
(b) \( |\psi''(x)| \leq Cx^{-1}(x - 1)^{\alpha-1} \)
(c) \( \psi'(1) = \beta \psi(1) \) when \( \alpha \geq 1 \).

**\( (H.2) \)** There are constants \( C > 0 \) and \( 1 < \nu \leq 2 \) such that, for all \( x \geq 2 \),

(d) \( |\phi'(x)| = \begin{cases} C(x - 1)^{-\nu} & \alpha < 1, \\ C(x - 1)^{\alpha-2} & \alpha \geq 1. \end{cases} \)
(e) \( |\phi''(x)| = \begin{cases} C(x - 1)^{-\nu-1} & \alpha < 1, \\ C(x - 1)^{\alpha-3} & \alpha \geq 1. \end{cases} \)

The reader is referred to [16, Section 4] for worked out examples of Gaussian processes satisfying assumptions \( (H.1) \) and \( (H.2) \), among which, as pointed out earlier, the bi-fractional Brownian motion and the sub-fractional Brownian motion.

Now, for \( k \geq 0 \), define

\[
\Delta X_k = X_{k+1} - X_k \quad \text{and} \quad Y_k = \frac{\Delta X_k}{\|\Delta X_k\|_{L^2(\Omega)}}.
\]

We are now ready to introduce the Gaussian random matrices \( \mathcal{Y} \) our Wishart matrices will be built upon. Let \( \{ X^i : i \in \mathbb{N} \} \) be i.i.d. copies of \( X \) and write

\[
\Delta X^i_k = X^i_{k+1} - X^i_k \quad \text{and} \quad Y^i_k = \frac{\Delta X^i_k}{\|\Delta X^i_k\|_{L^2(\Omega)}}.
\]

For any \( x \in [a, b] \) where \( a < b \) are two positive constants, let \( \mathcal{Y} \) be a \( n \times \lfloor dx \rfloor \) matrix with entries given by \( Y^i_k, 1 \leq i \leq n, 1 \leq k \leq \lfloor dx \rfloor \). Whenever the parameter \( \alpha \) (appearing in \( (H.1) \)) of the process \( X \) is such that \( 0 < \alpha < \frac{3}{2} \), we define the Wishart matrix \( W_{n, \lfloor dx \rfloor} \) to be

\[
W_{n, \lfloor dx \rfloor} = \frac{\lfloor dx \rfloor}{\sqrt{d}} \left( \frac{1}{\lfloor dx \rfloor} \mathcal{Y} \mathcal{Y}^T - I \right).
\]

\( W_{n, \lfloor dx \rfloor} \) is a \( n \times n \) matrix with entries given by, for any \( 1 \leq i, j \leq n, \)

\[
W_{ij}(\lfloor dx \rfloor) = \frac{1}{\sqrt{d}} \sum_{k=1}^{\lfloor dx \rfloor} \left( Y^i_k Y^j_k - 1_{\{i=j\}} \right).
\]
Whenever \( \alpha = \frac{3}{2} \), we define the Wishart matrix \( \tilde{W}_{n,[dx]} \), which differs from \( W_{n,[dx]} \) by the normalization of its entries, by
\[
\tilde{W}_{n,[dx]} = \frac{|dx|}{\sqrt{d \ln \alpha}} \left( \frac{1}{|dx|} YY^T - I \right).
\]
Finally, whenever \( \alpha > \frac{3}{2} \), we define another version of the Wishart matrix (with yet another normalization of the entries) by
\[
\hat{W}_{n,[dx]} = \frac{|dx|}{d^{n-1}} \left( \frac{1}{|dx|} YY^T - I \right).
\]

**Remark 1.** The Wishart matrices introduced above, \( W_{n,[dx]} \), \( \tilde{W}_{n,[dx]} \) and \( \hat{W}_{n,[dx]} \) are essentially the same object, only differing by the normalization of their entries. The fact that several normalization are required depending on the value of the parameter \( \alpha \) corresponds to the different asymptotic regimes appearing depending on said values.

**Remark 2.** We let the parameter \( x \in [a,b] \) in order to study functional convergence of Wishart matrices as \( d \) grows to infinity. If functional convergence is not the topic of interest for applications, nothing prevents one from taking \( x = 1 \) to be fixed and recover a classical \( n \times d \) matrix. The assumption \( a > 0 \) allows us to sidestep the case \( x = 0 \) and ensure \( Y \) of size \( n \times |dx| \) is well-defined, as long as \( d \) is sufficiently large.

From now on, whenever \( x \) is considered to be fixed (Sections 3.1, 3.2 and 4), we will drop the \( x \) dependency in our notation and write, for example, \( W_{ij} \) in place of \( W_{ij}([dx]) \). Moreover, in what follows, \( C \) denotes a generic positive constant that may vary from line to line.

Our first main result establishes central convergence in the case where \( 0 < \alpha < \frac{3}{2} \). The notation \( d_W \) stands for the Wasserstein distance introduced and defined in Section 2.

**Theorem 1.** Let \( 0 < \alpha < \frac{3}{2} \). Then, the Wishart matrix \( W_{n,[dx]} \) is close to the Gaussian Orthogonal Ensemble matrix \( Z_n \) (in finite-dimensional distribution) defined to be a \( n \times n \) symmetric matrix with independent entries such that \( Z_{ii} \sim N(0,2\sigma^2) \) and \( Z_{ij} \sim N(0,\sigma^2) \) for \( i \neq j \), where \( \sigma^2 \) is defined in Lemma 5. Furthermore, the following quantitative bound holds
\[
d_W(W_{n,[dx]}, Z_n) \leq C \left( n^2r(\alpha, \nu) + nd^{2\alpha-3} + nd^{-1} \right),
\]
where
\[
r(\alpha, \nu) = \begin{cases} 
\frac{d^3(\alpha-3)}{2(\alpha-2\alpha)} & \text{if } \alpha < 1 \text{ and } \alpha + \nu < 2 \\
\frac{d^3}{2} & \text{if } \alpha < 1 \text{ and } \alpha + \nu \geq 2 \\
\frac{d^3}{4} & \text{if } 1 \leq \alpha < \frac{5}{4} \\
\frac{d^3}{2}(\ln d)^\frac{3}{2} & \text{if } \alpha = \frac{5}{4} \\
d^{2\alpha-3} & \text{if } \frac{5}{4} < \alpha < \frac{3}{2}
\end{cases}.
\]

**Remark 3.** Since we have \( -\frac{1}{2} < \frac{2\alpha-3}{2(\alpha-2\alpha)} < 0 \) for \( \alpha < 1 \), the above convergence rate in the case where \( \alpha < 1 \) and \( \alpha + \nu < 2 \) is weaker than the rate appearing in the stationary case treated in [24]. This comes from the fact that some estimates used in [24] are not valid in our increased generality, and other arguments are needed, giving rise to this different convergence rate. One can hence see our result as being complementary to those in [24] as their rate is better if the process \( X \) is stationary, but ours allows to cover the non-stationary case as well. It is not surprising that the fact that our result accommodates many more processes than the stationary ones comes at the price of a slightly less optimal rate. From a technical point of view, it comes from the fact that the estimates in Lemma 6 do not hold for all instances of \( \alpha \) and \( \beta \), and warrants new estimates.
Before going any further, we would like to illustrate the results of Theorem 1 on two interesting examples of processes that are not stationary (and hence not covered by, for instance, [24]), namely the bi-fractional Brownian motion and the sub-fractional Brownian motion.

**Example 1.** If $X$ is a bi-fractional Brownian motion, i.e., a centered Gaussian process with covariance function given by

$$E[X_t X_s] = 2^{-K} \left((t^{2H} + s^{2H})^K - |t - s|^{2HK}\right),$$

where $H \in (0, 1)$, $K \in [a, b]$, then it was derived in [16, Section 4.1] that $\alpha = 2\beta = 2HK$ and $\nu = (1 + 2H - 2HK) \land (2 - 2HK)$. Theorem 1 then yields

$$d_W (\mathcal{W}_{n,[dx]}, Z_n) \leq C \begin{cases} 
  n^2 d^{4HK - \frac{3}{2}} & \text{if } 2HK < 1 \text{ and } 2H < 1 \\
  n^2 d^{-\frac{1}{2}} & \text{if } 2HK < 1 \text{ and } 2H > 1 \\
  n^2 d^{-\frac{1}{2}} + nd^{4HK - 3} & \text{if } 1 \leq 2HK < \frac{5}{4} \\
  n^2 d^{-\frac{1}{2}} (\ln d)^{\frac{3}{2}} & \text{if } 2HK = \frac{5}{4} \\
  n^2 d^{4HK - 3} & \text{if } \frac{5}{4} < 2HK < \frac{3}{2}
\end{cases}.$$

If $X$ is a sub-fractional Brownian motion, i.e., a centered Gaussian process with covariance function given by

$$E[X_t X_s] = t^{2H} + s^{2H} - \frac{1}{2} \left((t + s)^{2H} + |t - s|^{2H}\right),$$

where $H \in (0, 1)$, then it was derived in [16, Section 4.2] that $\alpha = 2\beta = 2H$ and $\nu = 2 - 2H$. Theorem 1 then yields

$$d_W (\mathcal{W}_{n,[dx]}, Z_n) \leq C \begin{cases} 
  n^2 d^{-\frac{1}{2}} & \text{if } 0 \leq H < \frac{1}{2} \\
  n^2 d^{-\frac{1}{2}} + nd^{4H - 3} & \text{if } \frac{1}{2} \leq H < \frac{5}{8} \\
  n^2 d^{-\frac{1}{2}} (\ln d)^{\frac{3}{2}} & \text{if } H = \frac{5}{8} \\
  n^2 d^{4H - 3} & \text{if } \frac{5}{8} < H < \frac{3}{4}
\end{cases}$$

which are the same convergence rates obtained in [24] for fractional Brownian motion.

When $\alpha = \frac{3}{2}$, we still have central convergence, but under a different normalization of the entries of the Wishart matrix, hence giving rise to a different regime for central convergence when compared to the case $\alpha < \frac{3}{2}$.

**Theorem 2.** Let $\alpha = \frac{3}{2}$. Then, the Wishart matrix $\widetilde{\mathcal{W}}_{n,[dx]}$ is close to the Gaussian Orthogonal Ensemble matrix $\widetilde{Z}_n$ (in finite-dimensional distribution) defined to be a $n \times n$ symmetric matrix with independent entries such that $Z_{ij} \sim N(0, 2\rho^2)$ and $Z_{ij} \sim N(0, \rho^2)$ for $i \neq j$, where $\rho^2$ is defined in Lemma 8. Furthermore, the following quantitative bound holds

$$d_W \left(\widetilde{\mathcal{W}}_{n,[dx]}, \widetilde{Z}_n\right) \leq C n^{3/2} \frac{\ln n}{\ln d}.$$
The following theorem describes the non-central asymptotic regime where the above defined object appears as the limit.

**Theorem 3.** Let \( \frac{3}{2} < \alpha < 2 \). Then, the Wishart matrix \( \hat{W}_{n, [dx]} \) is close to the Rosenblatt-Wishart matrix \( R_n \) (in finite-dimensional distribution) defined in Definition 1. Furthermore, the following quantitative bound holds

\[
d_W(\hat{W}_{n, [dx]}, R_n) \leq C n d^\frac{3-2\alpha}{2}.
\]

**Example 2.** If \( X \) is a bi-fractional Brownian motion as defined in Example 1 with \( \frac{3}{4} < H K < 1 \), then \( \hat{W}_{n, [dx]} \) converges to a Rosenblatt-Wishart matrix at the rate \( nd^\frac{3-2\alpha}{2} \). The same limit distribution and convergence rate apply when \( X \) is a sub-fractional Brownian motion with \( \frac{3}{4} < H < 1 \).

Comparing our result to [24], Nourdin and Zheng obtain the same convergence rate when \( X \) is a fractional Brownian motion with \( \frac{3}{4} < H < 1 \). Here we note that \( \alpha = 2\beta = 2H \) for a fractional Brownian motion.

As functional limit theorems are taking an increasingly growing importance in the literature, it is a natural question to ask whether the convergence results we have obtained so far (regarding the finite dimensional distributions of the Wishart matrices under consideration) can be made functional under potentially additional assumptions. We have chosen to explore the case where the second dimension of the matrix \( Y \) grows continuously. One could think of the index \( x \) appearing in the second dimension as time for instance and consider the case where the second dimension of \( Y \) grows with time (because it is continuously being fed new data over time for example). Our next result strengthens and complements Theorems 1, 2 and 3 by settling the question of functional convergence in \( C([a, b]; M_n(\mathbb{R})) \).

**Theorem 4.** The convergences stated in Theorems 1, 2 and 3 hold in \( C([a, b]; M_n(\mathbb{R})) \).

We now present an application of Theorems 1 and 2 to proving that the empirical spectral distributions of the Wishart matrices \( \frac{1}{\sqrt{n}} W_{n, [dx]} \) and \( \frac{1}{\sqrt{n}} \hat{W}_{n, [dx]} \) converge in expectation to the semicircular distribution. Recall that for a \( n \times n \) symmetric matrix \( M_n \), the (normalized) empirical spectral distribution is defined as

\[
\mu_{\sqrt{n} M_n} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(M_n)} / \sqrt{n},
\]

where \( \lambda_1(M_n) \leq \cdots \leq \lambda_n(M_n) \) are the (real) eigenvalues of \( M_n \), counting multiplicity. Recall also that the semicircular distribution \( \nu_t \) with variance \( t > 0 \) is a probability distribution defined by

\[
\nu_t(dx) = \frac{1}{2\pi t} \sqrt{4t - x^2} \, dx.
\]

By Wigner’s semicircle law, we know that the empirical spectral distribution of the GOE matrices \( Z_n \) defined in Theorem 1 converges in expectation to \( \nu_{\sigma^2} \) (where \( \sigma^2 \) is the variance of the entries of the GOE matrices \( Z_n \) from Theorem 1). The following result is an application of Theorem 1 and highlights the spectral behavior of the class of Wishart matrices studied in this paper.
Theorem 5. The empirical spectral distribution of \( \frac{1}{\sqrt{n}} W_n,_{dx} \) converges in expectation to the semicircular distribution \( \nu_{\sigma^2} \). In other words, as \( n, d \to \infty \), it holds that

\[
\mathbb{E} \left[ \mu \frac{1}{\sqrt{n}} W_n,_{dx} \right] \longrightarrow \nu_{\sigma^2}.
\]

Remark 4. We emphasize that the above theorem is just one possible application of our main results, as it focuses on one particular statistic of the the Wishart matrices, namely the empirical spectral distribution. The link between the spectral statistics of Wishart and Wigner matrices, the latter being known as the Gaussian Orthogonal Ensemble if the entries are Gaussian, has been studied extensively. Tracy and Widom obtained the limiting distribution of the largest eigenvalue of Wigner matrices in \([28, 29]\), and it is now known as the Tracy-Widom law. Johnstone \([19]\) and El-Karoui \([14]\) obtained the same limit distribution for largest eigenvalues of real and complex Wishart matrices under the regime \( \frac{d}{n} \to c \in [0, \infty) \). More recent work on the transition from Wishart to Wigner matrices in the high dimensional setting and/or the corresponding transition of spectral statistics such as condition number, extremal eigenvalues and others includes \([7, 8, 12, 17, 27]\). In this context, Theorem 5 once again demonstrates how similar the spectrum of Wishart and Wigner matrices behave, even if the independence condition between the entries of \( Y \) is relaxed.

Proof of Theorem 5. We need to prove that for any fixed \( k \geq 1 \), the \( k \)-th moment

\[
\frac{1}{n} \mathbb{E} \left[ \text{Tr} \left( \left( \frac{1}{\sqrt{n}} W_n,_{dx} \right)^k \right) \right]
\]

converges to the \( k \)-th moment of the semicircular distribution \( \nu_{\sigma^2} \). By Theorem 1, we have, as \( d \to \infty \),

\[
\frac{1}{\sqrt{n}} W_n,_{dx} \longrightarrow \frac{1}{\sqrt{n}} Z_n,
\]

where the convergence holds in distribution. By the continuous mapping theorem, it follows that

\[
\frac{1}{n} \text{Tr} \left( \left( \frac{1}{\sqrt{n}} W_n,_{dx} \right)^k \right) \longrightarrow \frac{1}{n} \text{Tr} \left( \left( \frac{1}{\sqrt{n}} Z_n \right)^k \right),
\]

as \( d \to \infty \), where the convergence holds in distribution, due to the fact that the map \( M_n \mapsto \frac{1}{n} \text{Tr} (M_n^k) \) is continuous as a multivariate polynomial. Now, the fact that the sequence

\[
\left\{ \frac{1}{n} \text{Tr} \left( \left( \frac{1}{\sqrt{n}} W_n,_{dx} \right)^k \right) : d \geq 1 \right\}
\]

is uniformly integrable (by the hypercontractivity of the Wiener chaos and the fact that the entries of \( W_n,_{dx} \) have bounded variances – see Lemma 5) together with the convergence in distribution (2) yields

\[
\frac{1}{n} \mathbb{E} \left[ \text{Tr} \left( \left( \frac{1}{\sqrt{n}} W_n,_{dx} \right)^k \right) \right] \longrightarrow \frac{1}{n} \mathbb{E} \left[ \text{Tr} \left( \left( \frac{1}{\sqrt{n}} Z_n \right)^k \right) \right].
\]

Letting \( n \to \infty \) and invoking Wigner’s semicircle law now yields the desired fact that the \( k \)-th moment

\[
\frac{1}{n} \mathbb{E} \left[ \text{Tr} \left( \left( \frac{1}{\sqrt{n}} W_n,_{dx} \right)^k \right) \right]
\]

converges (as first \( d \), then \( n \) go to infinity) to the \( k \)-th moment of the semicircular distribution \( \nu_{\sigma^2} \), which concludes the proof. \( \square \)
Remark 5. Theorem 5 is stated for the matrices $W_{n,[dx]}$, but the same result holds for the matrices $\tilde{W}_{n,[dx]}$ appearing in Theorem 2 with no modifications of the proof, other than the limiting semicircular distribution being $\nu_{\rho^2}$ in this case (and invoking Theorem 2 instead of Theorem 1).

This paper is organized as follows. Section 2 provides the needed elements of Malliavin calculus, as well as some results related to Stein’s method. Section 3.1 is dedicated to the preparation of the proof and the proof of Theorem 1, Section 3.2 contains the proof of Theorem 2 while Section 4 addresses the proof of Theorem 3. The proof of the functional version of our results (Theorem 4) is given in Section 5.1. Section 6 gathers technical and auxiliary results needed for the proofs of the main results.

2. Preliminaries

2.1. Overview of Malliavin calculus. Let $\mathcal{H}$ be a real separable Hilbert space and $\{Z(h): h \in \mathcal{H}\}$ an isonormal Gaussian process indexed by it, that is, a centered Gaussian family of random variables such that $\mathbb{E}[Z(h)Z(g)] = \langle h, g \rangle_{\mathcal{H}}$. Denote by $I_n$ the multiple Wiener (or Wiener-Itô) stochastic integral of order $n \geq 0$ with respect to $Z$ (see [25, Section 1.1.2]). The mapping $I_n$ is an isometry between the Hilbert space $\mathcal{H} \otimes^n$ (symmetric tensor product) equipped with the scaled norm $\frac{1}{n!} \| \cdot \|_{\mathcal{H} \otimes^n}$ and the Wiener chaos of order $n$, which is defined as the closed linear span of the random variables

$$\{H_n(Z(h)): h \in \mathcal{H}, \|h\|_{\mathcal{H}} = 1\},$$

where $H_n$ is the $n$-th Hermite polynomial given by $H_0 = 1$ and for $n \geq 1$,

$$H_n(x) = \frac{(-1)^n}{n!} \exp\left(\frac{x^2}{2}\right) \frac{d^n}{dx^n} \left(\exp\left(-\frac{x^2}{2}\right)\right), \quad x \in \mathbb{R}.$$

Multiple Wiener integrals enjoy the following isometry property: for any integers $m,n \geq 1$,

$$\mathbb{E}[I_n(f)I_m(g)] = \mathbb{I}_{\{n=m\}} n! \langle \tilde{f}, \tilde{g} \rangle_{\mathcal{H} \otimes^n},$$

where $\tilde{f}$ denotes the symmetrization of $f$ and we recall that $I_n(f) = I_n(\tilde{f})$.

Recall the multiplication formula satisfied by multiple Wiener integrals: for any $n,m \geq 1$, and any $f \in \mathcal{H} \otimes^n$ and $g \in \mathcal{H} \otimes^m$, it holds that

$$I_n(f)I_m(g) = \sum_{r=0}^{\min(n,m)} \binom{n}{r} \binom{m}{r} I_{n+m-2r}(f \otimes_r g),$$

where the $r$-th contraction of $f$ and $g$ is defined by, for $0 \leq r \leq m \wedge n$,

$$f \otimes_r g = \sum_{i_1,...,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathcal{H} \otimes^r} \otimes \langle g, e_{i_1} \otimes \cdots \otimes e_{i_r} \rangle_{\mathcal{H} \otimes^r},$$

with $\{e_i: i \geq 1\}$ denoting a complete orthonormal system in $\mathcal{H}$.

Recall that any square integrable random variable $F$ which is measurable with respect to the $\sigma$-algebra generated by $Z$ can be expanded into an orthogonal sum of multiple Wiener integrals:

$$F = \sum_{n=0}^{\infty} I_n(f_n),$$

where $f_n \in \mathcal{H} \otimes^n$ are (uniquely determined) symmetric functions and $I_0(f_0) = \mathbb{E}(F)$. 
Let $L$ denote the Ornstein-Uhlenbeck operator, whose action on a random variable $F$ with chaos decomposition (4) and such that $\sum_{n=1}^{\infty} n^2 n! \|f_n\|_{\mathbb{S}^{\otimes n}}^2 < \infty$ is given by

$$LF = - \sum_{n=1}^{\infty} n I_n(f_n).$$

A pseudoinverse $L^{-1}$ can be introduced via spectral calculus as follows:

$$L^{-1} F = - \sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n).$$

It follows that

$$LL^{-1} F = F - E[F].$$

For $p > 1$ and $k \in \mathbb{R}$, the Sobolev-Watanabe spaces $\mathbb{D}^{k,p}$ are defined as the closure of the set of polynomial random variables with respect to the norm

$$\|F\|_{k,p} = \|(I - L)^{\frac{k}{2}} F\|_{L^p(\Omega)},$$

where $I$ denotes the identity operator. We denote by $D$ the Malliavin derivative that acts on smooth random variables of the form $F = g(Z(h_1), \ldots, Z(h_n))$, where $g$ is a smooth function with compact support and $h_i \in \mathcal{S}$, $1 \leq i \leq n$. Its action on such a random variable $F$ is given by

$$DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(Z(h_1), \ldots, Z(h_n))h_i.$$ 

The operator $D$ is closable and continuous from $\mathbb{D}^{k,p}$ into $\mathbb{D}^{k-1,p}(\mathcal{S})$.

Denote by $\delta$ the adjoint of $D$, which is known as the divergence operator. An element $u \in L^2(\Omega, \mathcal{S})$ belongs to dom($\delta$) only if there exists a constant $C_u$ depending only on $u$ such that

$$\left|E\left[\langle DF, u \rangle_{\mathcal{S}}\right]\right| \leq C_u \|F\|_{L^2(\Omega)}$$

for any $F \in \mathbb{D}^{1,2}$. In this case, we have the following integration by parts formula (or duality relation)

$$E[F\delta(u)] = E[\langle DF, u \rangle_{\mathcal{S}}].$$

**Remark 6.** A random variable $F$ is an element of dom $L = \mathbb{D}^{2,2}$ if and only if $F \in$ dom($\delta D$) (that is $F \in \mathbb{D}^{1,2}$ and $DF \in$ dom($\delta$)), and in this case,

$$LF = -\delta DF.$$

An important result from Malliavin calculus we will make use of in the sequel are Meyer’s inequalities: for any $1 \leq p \leq k$ and $u \in \mathbb{D}^{k,q}(\mathcal{S}^{\otimes p})$, there exists a constant $C > 0$ such that

$$\|\delta^p(u)\|_{k-p,q} \leq C \|u\|_{\mathbb{D}^{k,q}(\mathcal{S}^{\otimes p})}.\tag{5}$$

For a more complete treatment of Meyer’s inequalities, we refer to [23, Theorem 2.5.5] or [25, Section 1.5], and related results therein.
2.2. Distances between random matrices. We will use the Wasserstein distance between two random matrices taking values in $\mathcal{M}_n(\mathbb{R})$, which denotes the space of $n \times n$ real matrices. Given two $\mathcal{M}_n(\mathbb{R})$-valued random matrices $\mathcal{X}$ and $\mathcal{Y}$, the Wasserstein distance between them is given by

$$d_W(\mathcal{X}, \mathcal{Y}) = \sup_{\|g\|_{\text{Lip}} \leq 1} \left| \mathbb{E}[g(\mathcal{X})] - \mathbb{E}[g(\mathcal{Y})] \right|,$$

where the Lipschitz norm $\|\cdot\|_{\text{Lip}}$ of $g: \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ is defined by

$$\|g\|_{\text{Lip}} = \sup_{A \neq B \in \mathcal{M}_n(\mathbb{R})} \frac{|g(A) - g(B)|}{\|A - B\|_{\text{HS}}},$$

with $\|\cdot\|_{\text{HS}}$ denoting the Hilbert-Schmidt norm on $\mathcal{M}_n(\mathbb{R})$.

We will also make use of the Wasserstein distance between random vectors, defined analogously as in the matrix case. Namely, if $X, Y$ are two $n$-dimensional random vectors, then the Wasserstein distance between them is defined to be

$$d_W(X, Y) = \sup_{\|g\|_{\text{Lip}} \leq 1} |\mathbb{E}(g(X)) - \mathbb{E}(g(Y))|,$$

where the Lipschitz norm $\|\cdot\|_{\text{Lip}}$ of $g: \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\|g\|_{\text{Lip}} = \sup_{x \neq y \in \mathbb{R}^n} \frac{|g(x) - g(y)|}{\|x - y\|_{\mathbb{R}^n}},$$

with $\|\cdot\|_{\mathbb{R}^n}$ denoting the Euclidean norm on $\mathbb{R}^n$.

If $\mathcal{X} = (X_{ij})_{1 \leq i, j \leq n}$ is an $n \times n$ symmetric random matrix, we associate to it its “half-matrix” defined to be the $(n(n + 1)/2)$-dimensional random vector

$$\mathcal{X}^{\text{half}} = (X_{11}, X_{12}, \ldots, X_{1n}, X_{22}, X_{23}, \ldots, X_{2n}, \ldots, X_{nn}).$$

It turns out that, in the case of two symmetric matrices, the Wasserstein distance between said matrices can be bounded from above by a constant multiple of the Wasserstein distance between their associated half-matrices. More specifically, we have the following Lemma (see [24, Lemma 2.2]).

**Lemma 1.** Let $\mathcal{X}, \mathcal{Y}$ be two symmetric random matrices with values in $\mathcal{M}_n(\mathbb{R})$. Then

$$d_W(\mathcal{X}, \mathcal{Y}) \leq \sqrt{2}d_W(\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}}),$$

where $\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}}$ are the associated half-matrices defined in (6).

This shows that assessing the Wasserstein distance between symmetric random matrices can be shifted to the problem of estimating the Wasserstein distance between associated random vectors (see Lemma 1). In our context, a helpful result in this direction is [23, Theorem 6.1.1], which we restate here for convenience.

**Proposition 1** (Theorem 6.1.1 in [23]). Fix $m \geq 2$, and let $F = (F_1, \ldots, F_m)$ be a centered $m$-dimensional random vector with $F_i \in \mathbb{D}^{1,4}$ for every $i = 1, \ldots, m$. Let $C \in \mathcal{M}_m(\mathbb{R})$ be a symmetric and positive definite matrix, and let $Z \sim N_m(0, C)$. Then,

$$d_W(F, Z) \leq \|C^{-1}\|_{\text{op}} \|C\|_{\text{op}}^{1/2} \sqrt{\sum_{i,j=1}^{m} \mathbb{E} \left[ (C_{ij} - (DF_i, -DL^{-1}F_j))_2^2 \right]},$$

where $\|\cdot\|_{\text{op}}$ denotes the operator norm on $\mathcal{M}_m(\mathbb{R})$. 
3. Proofs of main central convergence results

This section is dedicated to the proofs of the main central convergence results, namely Theorems 1 and 2. Throughout this section and for the rest of the paper, \( f(x) = o(g(x)) \) is taken to mean

\[
\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0.
\]

3.1. Proof of Theorem 1. We start by embedding the covariance structure of \( (Y_{ik})_{i,k \in \mathbb{N}} \) in a Hilbert space \( \mathcal{H} \) such that for a collection of elements \( (e_{ik})_{i,k \in \mathbb{N}} \) in \( \mathcal{H} \), \( Y_{ik} = Z(e_{ik}) \) and \( \langle e_{ik}, e_{jl} \rangle_{\mathcal{H}} = \mathbb{E}[Y_{ik}Y_{jl}] \). Since the rows of \( Y_{n,\{dx\}} \) are independent and each entry of \( Y_{n,\{dx\}} \) is normalized, we have \( \langle e_{ik}, e_{jl} \rangle_{\mathcal{H}} = 0 \) for \( i \neq j \) and \( \|e_{ik}\|_{\mathcal{H}} = 1 \). For notational convenience, we will denote by \( \delta_{kj} = \langle e_{ik}, e_{ij} \rangle_{\mathcal{H}} \) in the sequel. With this structure handy, the entries of \( W_{n,\{dx\}} \) can be represented as

\[
W_{ij} = I_2(f_{ij})
\]

for any \( 1 \leq i, j \leq n \), where the kernel \( f_{ij} \) is defined by

\[
f_{ij} = \frac{1}{2\sqrt{d}} \sum_{k=1}^{[dx]} (e_{ik} \otimes e_{jk} + e_{jk} \otimes e_{ik}).
\]

Let \( \mathcal{G}_n \) denote a \( n \times n \) Gaussian matrix, having the same covariance structure as \( W_{n,\{dx\}} \). Using the notation introduced in (6), we write \( W_{\text{half}} \) to denote the half-matrix associated to \( W_{n,\{dx\}} \), that is

\[
W_{\text{half}} = (W_{11}, \ldots, W_{1n}, W_{2n}, \ldots, W_{2n}, \ldots, W_{nn}),
\]

and we define \( G_{\text{half}} \) in a similar way. As pointed out in Lemma 1, we have \( d_W(W_{n,\{dx\}}, \mathcal{G}_n) \leq \sqrt{2d_W(W_{\text{half}}, G_{\text{half}})} \) since the matrices \( W_{n,\{dx\}} \) and \( \mathcal{G}_n \) are symmetric.

**Remark 7.** Slightly abusing notation, we will continue to write \( W_{n,\{dx\}} \) and \( \mathcal{G}_n \) in place of \( W_{\text{half}} \) and \( G_{\text{half}} \), respectively.

Our goal being to estimate \( d_W(W_{n,\{dx\}}, \mathcal{Z}_n) \), we apply the triangle inequality to get

\[
d_W(W_{n,\{dx\}}, \mathcal{Z}_n) \leq d_W(W_{n,\{dx\}}, \mathcal{G}_n) + d_W(\mathcal{G}_n, \mathcal{Z}_n)
\]

and split the proof into two steps, the first one aiming at estimating \( d_W(W_{n,\{dx\}}, \mathcal{G}_n) \), and the second one dealing with the estimation of \( d_W(\mathcal{G}_n, \mathcal{Z}_n) \).

**Step 1: Estimation of** \( d_W(W_{n,\{dx\}}, \mathcal{G}_n) \).

According to Lemma 4, we can write

\[
\|f_{ij} \otimes f_{ik}\|_{\mathcal{H}^\otimes 2}^2 \leq \frac{1}{d^2} \sum_{k,l,m,p=1}^{[dx]} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp}.
\]

Based on Lemma 2, it holds that

\[
\|C^{-1}\|_{\text{op}} \|C\|_{\text{op}}^{1/2} = \sqrt{\frac{2d}{\sum_{k,l=1}^{[dx]} \delta_{kl}}}.
\]

where \( C \) denotes the covariance matrix of \( W_{n,\{dx\}}^{\text{half}} \). Applying Proposition 1 together with Lemmas 3 and 4, and observing that the cardinality of the set

\[
\{(i,j,k,l) \in \{1, \ldots, n\}^4 : i,j,k,l \text{ are not mutually distinct}\}
\]
is bounded by $6n^3$, yields

$$
d_W(W_n, |dx|, \mathcal{G}_n) \leq \sqrt{2} \left\| C^{-1} \right\|_{op} \left\| C \right\|_{op} \left( \sum_{i,j,k,l=1}^{n} \mathbb{E} \left[ \left( \mathbb{E}[W_{ij}W_{kl}] - \frac{1}{2} \langle DW_{ij}, DW_{kl} \rangle_{\mathcal{G}} \right)^2 \right] \right)^{\frac{1}{2}} \leq \sqrt{2} \left( \frac{2d}{\sum_{k,l=1}^{d} \delta_{kl}} \right)^{\frac{1}{2}} \left( \sum_{i,j,k,l=1}^{n} 8 \left\| f_{ij} \hat{\otimes} f_{lk} \right\|_{\mathcal{G}^2}^2 \right)^{\frac{1}{2}} \leq C \left( \frac{d}{\sum_{k,l=1}^{d} \delta_{kl}} \right)^{\frac{1}{2}} \left( \frac{n^3}{d^2} \sum_{k,l,m,p=1}^{d} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp} \right)^{\frac{1}{2}}.
$$

(10)

We now need to estimate the right hand side of (10). We divide this estimation into two cases, the first of which deals with the case where the process $X$ is such that $\alpha < 1$ and $\alpha + \nu < 2$, and the second one covering all other possibilities.

**Case 1** ($\alpha < 1$ and $\alpha + \nu < 2$). Lemma 5 implies that the first factor on the right hand side of (10) is bounded, so that it is sufficient to estimate the second factor. For $\frac{1}{2} < \alpha < \frac{3}{2}$, let $\theta$ be a constant in $(0, 1)$ and let $M_d = \lfloor (\alpha x)^{2} \rfloor$. Denote by $D_d$ the set of multi-indexes $\{1, \ldots, [dx]\}^4$ and decompose $D_d$ into $D_{1,M_d} \cup D_{2,M_d}$ according to

$$D_{1,M_d} = \{(k, l, m, p) \in D_d : |k - l| \leq M_d, |k - m| \leq M_d, |m - p| \leq M_d\}
$$

and

$$D_{2,M_d} = D_{3,M_d} \cup D_{4,M_d} \cup D_{5,M_d},
$$

where

$$\begin{cases}
D_{3,M_d} = \{(k, l, m, p) \in D_d : |k - l| > M_d\}, \\
D_{4,M_d} = \{(k, l, m, p) \in D_d : |k - m| > M_d\}, \\
D_{5,M_d} = \{(k, l, m, p) \in D_d : |m - p| > M_d\}.
\end{cases}
$$

Note that the cardinality of $D_{1,M_d}$ is bounded by $8dM_d^2$ since there are fewer than $d$ choices for $k$ and for each $k$, one has $2M_d$ choices for each $l, m, p$. Hence, we can write

$$\frac{n^3}{d^2} \sum_{k,l,m,p=1}^{d} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp} \leq \frac{n^3}{d^2} \sum_{(k,l,m,p) \in D_{1,M_d}} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp} + \frac{3n^3}{d^2} \sum_{(k,l,m,p) \in D_{3,M_d}} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp}$$

$$\leq 8n^3 M_d^2 \frac{d}{d^2} + \frac{3n^3}{d^2} \left( \sum_{(k,l,m,p) \in D_{3,M_d}} |\delta_{kl} \delta_{mp}|^2 \right)^{\frac{1}{2}} \left( \sum_{(k,l,m,p) \in D_{3,M_d}} |\delta_{km} |^2 \right)^{\frac{1}{2}}$$

$$\leq 8n^3 M_d^2 d^{-1} + 3n^3 \left( d^{-1} \sum_{1 \leq k,l \leq |dx| |k-l| > M_d} |\delta_{kl}|^2 \right)^{\frac{1}{2}} \left( d^{-1} \sum_{m,p=1}^{d} |\delta_{mp}|^2 \right)^{\frac{1}{2}}.$$
As Lemma 5 implies that \(d^{-1} \sum_{m,p=1}^{[dx]} |\delta_{mp}|^2 < \infty\), we get

\[
\frac{n^3}{d^2} \sum_{k,l,m,p=1}^{[dx]} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp} \leq C n^3 M_d^3 d^{-1} + 3 C n^3 \left( d^{-1} \sum_{1 \leq k,l \leq [dx]} \frac{|\delta_{kl}|^2}{|k-l|^M_d} \right)^{\frac{1}{2}}
\]

\[
\leq C n^3 d^{3\theta-1} + 3 C n^3 \left( d^{-1} \sum_{1 \leq k,l \leq [dx]} \frac{|\delta_{kl}|^2}{|k-l|^M_d} \right)^{\frac{1}{2}}.
\]

Now we use Remark 9 which states that the dominant part of \(d^{-1} \sum_{k,l=1}^{[dx]} \delta_{kl}\) is

\[
\frac{1}{d} \sum_{m=1}^{[dx]} \left( \frac{k}{k+m} \right)^{2\beta-\alpha} (|m+1|^{\alpha} + |m-1|^{\alpha} - 2|m|^{\alpha})^2 + \frac{x}{D},
\]

and the fact that

\[
|m+1|^{\alpha} + |m-1|^{\alpha} - 2|m|^{\alpha} = \frac{1}{2} \alpha(\alpha-1) |m|^{\alpha-2} + o(|m|^{\alpha-2})
\]

to get

\[
n^3 \left( d^{-1} \sum_{1 \leq k,l \leq [dx]} \frac{|\delta_{kl}|^2}{|k-l|^M_d} \right)^{\frac{1}{2}} \leq C n^3 \left( \sum_{m=|k-l|=M_d}^{[dx]-1} (|m+1|^{\alpha} + |m-1|^{\alpha} - 2|m|^{\alpha})^2 \right)^{\frac{1}{2}}
\]

\[
\leq C n^3 \left( \sum_{m=M_d}^{d-1} \alpha^2 (\alpha-1)^2 m^{2(\alpha-2)} \right)^{\frac{1}{2}}.
\]

This finally gives us

\[
n^3 \left( d^{-1} \sum_{1 \leq k,l \leq [dx]} \frac{|\delta_{kl}|^2}{|k-l|^M_d} \right)^{\frac{1}{2}} \leq C n^3 (M_d^{2\alpha-3})^{\frac{1}{2}} \leq C n^3 d^{\frac{1}{2}\theta(2\alpha-3)}.
\]

This gives us the convergence rate

\[
d_W (\mathcal{W}_{n,|dx|}, \mathcal{G}_n) \leq C \left( n^3 d^{-2} \sum_{k,l,m,p=1}^{[dx]} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp} \right)^{\frac{1}{2}} \leq C n^3 \left( d^{\frac{1}{2}(3\theta-1)} + d^{\frac{1}{2}\theta(2\alpha-3)} \right)
\]

\[
= C n^3 d^{\frac{1}{2}(3\theta-1) + \frac{1}{2}\theta(2\alpha-3)},
\]

which holds for any \(\theta \in (0, \frac{1}{3})\). Now, observe that for \(\theta \in (0, \frac{1}{3})\) and \(\alpha \in (0, \frac{3}{2})\), the function

\[
f(\theta) = \frac{1}{2}(3\theta - 1) + \frac{1}{4}\theta(2\alpha - 3) = \begin{cases} (1/4)\theta(2\alpha - 3) & \text{if } \theta \leq \frac{2}{9-2\alpha} \\ (1/2)(3\theta - 1) & \text{if } \theta > \frac{2}{9-2\alpha} \end{cases}
\]

attains its minimum \(\frac{2\alpha-3}{2(9-2\alpha)}\) at \(\theta = \frac{2}{9-2\alpha}\). This allows us to conclude that

\[
d_W (\mathcal{W}_{n,|dx|}, \mathcal{G}_n) \leq C n^3 d^{\frac{2\alpha-3}{2(9-2\alpha)}}.
\]
Remark 8. The above estimate works for all Wishart matrices such that $0 < \alpha < \frac{3}{2}$ and not just $\alpha < 1$ and $\alpha + \nu < 2$, but it turns out that we can obtain a better estimate for the other cases as the rest of the proof will show.

Case 2 ($1 \leq \alpha < \frac{3}{2}$ or $\alpha < 1$ and $\alpha + \nu \geq 2$). In this second case, we consider a process $X$ that satisfies any other assumptions appearing in Theorem 1 besides $\alpha < 1$ and $\alpha + \nu < 2$, and we obtain the same central convergence rate as the one appearing in [24] (recall that, as mentioned in the introduction, the results in [24] are only valid under the assumption of stationarity of $X$, which we do not impose here).

The fact that we get the rate $r(\alpha, \nu)$ in the case where $1 \leq \alpha < \frac{3}{2}$ or $\alpha < 1$ and $\alpha + \nu \geq 2$ follows directly from Lemma 6, the bound (10) and the fact that


d^{-1} \left( \sum_{m=-d+1}^{d-1} |a_\alpha(m)|^\frac{2}{3} \right)^3 \leq C \begin{cases} 
  d^{-1} & \text{if } 0 < \alpha < \frac{5}{4} \text{ and } 1 \geq \alpha < 2 \\
  d^{-1} \ln(d)^{\frac{3}{2}} & \text{if } \alpha = \frac{5}{4} \\
  d^{4\alpha - 6} & \text{if } \frac{5}{4} < \alpha < 2 
\end{cases}

Step 2: Estimation of $d_W(\mathcal{G}_n, \mathcal{Z}_n)$.

This step in concerned with bounding the Wasserstein distance between $\mathcal{G}_n$ and $\mathcal{Z}_n$. Applying Proposition 1 yields

\begin{align*}
  d_W(\mathcal{G}_n, \mathcal{Z}_n) & \leq \sqrt{2} \|C^{-1}_Z\|_{op} \|C_Z^\frac{1}{2}\| \left( \sum_{1 \leq i,j,l,k \leq n} \mathbb{E} \left( \left( \mathbb{E}[Z_{ij}Z_{lk}] - \langle DG_{ij}, DG_{lk}\rangle_0 \right)^2 \right) \right)^{\frac{1}{2}},
\end{align*}

where $C_Z$ denotes the covariance matrix of $\mathcal{Z}_n$. Recall that $Z_{ii} \sim N(0, 2\sigma^2)$, that $Z_{ij} \sim N(0, \sigma^2)$ for $i \neq j$ and that all the entries of $\mathcal{Z}_n$ are independent. Hence, Lemma 5 provides us with the exact values of $\mathbb{E}[Z_{ij}Z_{lk}]$. Lemma 5 also implies that $\sigma^2 \leq \|C_Z\|_{op} \leq 2\sigma^2$. Meanwhile, $G_{ij}$ and $G_{lk}$ are in the first Wiener chaos associated to $X$, so that $\langle DG_{ij}, DG_{lk}\rangle_0 = \mathbb{E}[G_{ij}G_{lk}]$. Thus, Lemma 2, Remark 9 and the fact that $\mathcal{G}_n$ and $\mathcal{W}_{n,|dx|}$ are identically distributed yield

\begin{align*}
  \mathbb{E}[G_{il}G_{il}] &= 2d^{-1} \sum_{k,j=1}^{[dx]} \delta^2_{kj} = 2d^{-1} \sum_{m=1}^{[dx]-1} \sum_{k=1}^{[dx]-m-1} \left( \frac{k}{k + m} \right) a^2_\alpha(m) + \frac{x^2}{4} + R_d, \\
  \mathbb{E}[G_{il}G_{il}] &= d^{-1} \sum_{k,j=1}^{[dx]} \delta^2_{kj} = 2d^{-1} \sum_{m=1}^{[dx]-1} \sum_{k=1}^{[dx]-m-1} \left( \frac{k}{k + m} \right) a^2_\alpha(m) + \frac{x^2}{4} + R_d,
\end{align*}

and for $i \neq l$,

\begin{align*}
  \mathbb{E}[G_{il}G_{il}] &= d^{-1} \sum_{k,j=1}^{[dx]} \delta^2_{kj} = 2d^{-1} \sum_{m=1}^{[dx]-1} \sum_{k=1}^{[dx]-m-1} \left( \frac{k}{k + m} \right) a^2_\alpha(m) + \frac{x}{4} + R_d,
\end{align*}

where $R_d = O(d^{2\alpha-3} + d^{-1})$.

Recalling that $a^2_\alpha(m) = \frac{1}{4}(|m + 1|^\alpha + |m - 1|^\alpha - 2|m|^\alpha)^2$, we are able to get

\begin{align*}
  d_W(\mathcal{G}_n, \mathcal{Z}_n) & \leq \sqrt{2} \frac{\sqrt{2\sigma^2}}{\sigma^2} \left( \frac{n(n + 1)}{2} \left( 2\sigma^2 - 2d^{-1} \sum_{k,j=1}^{[dx]} \delta^2_{kj} \right)^2 \right)^{\frac{1}{4}} \leq C \frac{\sqrt{n(n + 1)}}{\sigma} (A_1 + A_2 - R_d),
\end{align*}

where $A_1$ and $A_2$ are the constants from Lemma 5 and $R_d$ is the remainder term from the approximation of $\mathbb{E}[G_{il}G_{il}]$.
where
\[ A_1 = \frac{x}{2} \sum_{m \in \mathbb{Z}} a_3^2(m) - \frac{1}{d} \sum_{m=1}^{|dx| - 1} (|dx| - m - 1)a_3^2(m) - \frac{x}{2}a_3^2(0) \]
and
\[ A_2 = \frac{1}{d} \sum_{m=1}^{|dx| - 1} (|dx| - m - 1)a_3^2(m) - \frac{1}{d} \sum_{m=1}^{|dx| - 1} \sum_{k=1}^{|dx| - m - 1} \left( \frac{k}{k + m} \right)^{2\beta - \alpha} a_3^2(m). \]

As \( d \) gets sufficiently large, the term \( A_1 \) can be bounded by
\[ A_1 = \sum_{m=1}^{|dx| - 1} \frac{m + 1}{d} a_3^2(m) + x \sum_{m=|dx|}^\infty a_3^2(m) < Cd^{2\alpha - 3} \]
for \( \alpha < \frac{3}{2} \) and \( \alpha \neq 1 \). Meanwhile, to deal with the term \( A_2 \), we observe that as \( x \to \infty \), \( \ln |x| \leq |x|^\zeta \) for any positive value of \( \zeta \). We also know that \( a_\alpha(m) = \frac{1}{2}\alpha(\alpha - 1)m^{\alpha - 2} + o(m^{\alpha - 2}) \).
Thus, a Taylor expansion of \((1 - x)^p\) for \( p > 0 \) and \( 0 \leq x < 1 \) gives us
\[ A_2 = \frac{C}{d} \sum_{m=1}^{|dx| - 1} \int_1^{|dx| - m - 1} \frac{1}{y + m - 1} dy \]
\[ \leq \frac{C}{d} \sum_{m=1}^{|dx| - 1} \frac{m^{2\alpha - 3}}{m \ln |dx|} \]
\[ \leq \frac{C}{d} \sum_{m=1}^{|dx| - 1} m^{2\alpha - 3} (\frac{|dx|}{m})^\zeta \]
\[ \leq C d^{\zeta - 1} \sum_{m=1}^{|dx| - 1} m^{2\alpha - \zeta - 3} \leq C d^{2\alpha - 3}. \]

Earlier we have mentioned that \( Rd = o(d^{2\alpha - 3} + d^{-1}) \). This allows us to conclude that, for \( \alpha < \frac{3}{2} \) and \( \alpha \neq 1 \), we have
\[ d_W(G_n, Z_n) \leq C \sqrt{n(n + 1)} (d^{2\alpha - 3} + d^{-1}) \]
for any \( \zeta \in (0, 1) \). As for \( \alpha = 1 \), \( a_1(m) = 0 \) if \( m \neq 0 \) and \( a_1(0) = 1 \), so the estimate (11) becomes
\[ d_W(G_n, Z_n) \leq C \left( x - \frac{|dx|}{d} - \frac{1}{d} \right) \leq C \frac{d}{d}. \]

Finally the estimate in Theorem 1 follows immediately from (9). To conclude that \( W_{n, |dx|} \) is close to \( Z_n \) in the sense of finite-dimensional distributions, we refer to [23, Theorem 6.2.3] which states that for a sequence of vectors of multiple Wiener integrals, component-wise convergence to a Gaussian limit implies joint convergence.
3.2. Proof of Theorem 2. In the case $\alpha = \frac{3}{2}$, recall that we have to use a different normalization for the elements of the Wishart matrix (we hence adjust notation accordingly), namely for $i \leq j$,

$$
\widetilde{W}_{ij}(dx) = \frac{1}{\sqrt{d \ln d}} \sum_{k=1}^{\lfloor dx \rfloor} (Y^i_k Y^j_k - \mathbf{1}_{i=j}).
$$

Let $\widehat{G}_n$ be $n \times n$ Gaussian matrices with the same covariance structure as $\widetilde{W}_{n,\lfloor dx \rfloor}$, which as before denotes a half-matrix vector. In the same spirit as in the proof of Theorem 1, we will first estimate the Wasserstein distance from $\widehat{W}_{n,\lfloor dx \rfloor}$ to $\widehat{G}_n$ in Step 1 and the distance from $\widehat{G}_n$ to the G.O.E matrix $\widehat{Z}_n$ in Step 2. Theorem 2 then follows from the triangle inequality

$$
d_W(\widehat{W}_{n,\lfloor dx \rfloor}, \widehat{Z}_n) \leq d_W(\widehat{W}_{n,\lfloor dx \rfloor}, \widehat{G}_n) + d_W(\widehat{G}_n, \widehat{Z}_n).
$$

**Step 1: Estimation of $d_W(\widehat{W}_{n,\lfloor dx \rfloor}, \widehat{G}_n)$.**

We proceed in the same way as in the beginning of the proof of Theorem 1. Denote by $\widehat{C}$ the covariance matrix of $\widehat{W}_{n,\lfloor dx \rfloor}$ and $\widehat{G}_n$. Use Proposition 1 and observe that Lemmas 3 and 4 still hold for $\alpha = \frac{3}{2}$, modulo a change of normalizing factor. More precisely, we can write

$$
d_W(\widehat{W}_{n,\lfloor dx \rfloor}, \widehat{G}_n) \leq \sqrt{2} \|\widehat{C}^{-1}\|_{op} \|\widehat{C}\|_{op} \left( \sum_{i,j,k,l} \mathbb{E} \left[ \frac{1}{2} \left( D\widehat{W}_{ij}, D\widehat{W}_{kl} \right) \right] \right)^{1/2},
$$

at which point we can use Lemmas 7 and 8 to get

$$
d_W(\widehat{W}_{n,\lfloor dx \rfloor}, \widehat{G}_n) \leq C \left( \frac{n^3}{(d \ln d)^2} \sum_{k,l,m,p=1}^{\lfloor dx \rfloor} \delta_{kl}\delta_{mp}\delta_{kn}\delta_{lp} \right)^{1/2},
$$

where

$$
a_\frac{3}{2}(m) = \frac{1}{2} \left( |m + 1|^{\frac{3}{2}} + |m - 1|^{\frac{3}{2}} - 2 |m|^{\frac{3}{2}} \right) = \frac{3}{8} |m|^{-\frac{1}{2}} + o(|m|^{-\frac{1}{2}}).
$$

Combining with the fact that

$$
\sum_{m=1}^{\lfloor dx \rfloor} m^{-2/3} \leq \int_1^{\lfloor dx \rfloor} (y - 1)^{-2/3} dy \leq d^{1/3},
$$

we obtain

$$
d_W(\widehat{W}_{n,\lfloor dx \rfloor}, \widehat{G}_n) \leq C \left( \frac{n^3}{d(\ln d)^2} \sum_{m=-\lfloor dx \rfloor + 1}^{\lfloor dx \rfloor} |m|^{-2/3} + 1 \right)^{1/2} \leq C n^{3/2} \ln d.
$$

**Step 2: Estimation of $d_W(\widehat{G}_n, \widehat{Z}_n)$.** Denote by $\widehat{C}_\frac{2}{3}$ the covariance matrix of $\widehat{Z}_n$. Proposition
and all of the entries of $E$ done in Step 2 in the proof of Theorem 1). Now keep in mind that $a_1$ implies that $X$ first Wiener chaos associated to $\tilde{Z}$.

To estimate the above expression, we notice that since $\tilde{Z}_i \sim N(0, 2\rho^2)$, $\tilde{Z}_{ij} \sim N(0, \rho^2)$ for $i \neq j$ and all of the entries of $\tilde{Z}_n$ are independent, Lemma 8 provides us with the exact value of $\mathbb{E}[Z_{ij}Z_{lk}]$. Lemma 8 also implies that $\rho^2 \leq \|C_Z\|_{\text{op}} \leq 2\rho^2$. Meanwhile, $\tilde{G}_{ij}$ and $\tilde{G}_{ik}$ are in the first Wiener chaos associated to $X$, so that $\langle D\tilde{G}_{ij}, D\tilde{G}_{ik}\rangle_{\mathcal{B}} = \mathbb{E}[\tilde{G}_{ij}\tilde{G}_{ik}]$. Combined with the fact that $\tilde{G}_n$ and $\tilde{W}_{n,\lfloor dx \rfloor}$ are identically distributed, again an application of Lemma 8 yields

$$
\mathbb{E}[\tilde{G}_{ij}\tilde{G}_{ij}] = 2 \frac{1}{d \ln d} \sum_{k,j=1}^{\lfloor dx \rfloor} \delta_{kj}^2 = \frac{1}{d \ln d} \sum_{m=1}^{\lfloor dx \rfloor} \sum_{k=1}^{\lfloor dx \rfloor - m - 1} \left( \frac{k}{k + m} \right)^{2\beta - \frac{3}{2}} a_4(m)^2 + \tilde{R}_d
$$

and for $i \neq l$,

$$
\mathbb{E}[\tilde{G}_{il}\tilde{G}_{il}] = \frac{1}{d \ln d} \sum_{k,j=1}^{\lfloor dx \rfloor} \delta_{kj}^2 = \frac{1}{2 d \ln d} \sum_{m=1}^{\lfloor dx \rfloor - m - 1} \sum_{k=1}^{\lfloor dx \rfloor - m - 1} \left( \frac{k}{k + m} \right)^{2\beta - \frac{3}{2}} a_4(m)^2 + \tilde{R}_d,
$$

with $\tilde{R}_d = o(\frac{1}{d \ln d})$, which can be deduced from the proof of Lemma 8 (similarly to what was done in Step 2 in the proof of Theorem 1). Now keep in mind that $a_4(m)$ is defined in Lemma 7 and $a_4(m)^2 = \frac{9}{81} m^{-1} + o(m^{-1})$. This yields

$$
d_W(\tilde{G}_n, \tilde{Z}_n) \leq \sqrt{2} \sqrt{\frac{2\rho^2}{\rho^2}} \left( \frac{n(n + 1)}{2} \left( 2\rho^2 - 2d^{-1} \sum_{k,j=1}^{\lfloor dx \rfloor} \delta_{kj}^2 \right) \right)^{\frac{1}{2}} 
$$

$$
\leq C \frac{\sqrt{n(n + 1)}}{\rho} (A_3 + A_4 - \tilde{R}_d),
$$

where

$$
A_3 = x - \frac{1}{d \ln d} \sum_{m=1}^{\lfloor dx \rfloor - m - 1} (\lfloor dx \rfloor - m - 1) m^{-1}
$$

and

$$
A_4 = \frac{1}{d \ln d} \sum_{m=1}^{\lfloor dx \rfloor - m - 1} (\lfloor dx \rfloor - m - 1) m^{-1} - \frac{1}{d \ln d} \sum_{m=1}^{\lfloor dx \rfloor - m - 1} \sum_{k=1}^{\lfloor dx \rfloor - m - 1} \left( \frac{k}{k + m} \right)^{2\beta - \alpha} m^{-1}.
$$
For the term $A_3$, asymptotically, we can write

$$A_3 \leq C \left( x - \frac{1}{d \ln d} \int_1^{[dx]} \frac{|dx| - y - 1}{y} dy \right)$$

$$= C \left( x - \frac{1}{d \ln d} \int_1^{[dx]} \left( \frac{|dx|}{y} - 1 - \frac{1}{y} \right) dy \right)$$

$$= C \left( x - \frac{|dx| \ln (|dx| - 1)}{d \ln d} + \frac{1}{d \ln d} + \frac{\ln (|dx| - 1)}{d \ln d} \right)$$

$$\leq \frac{C}{\ln d}.$$

For the term $A_4$, observe that as $x \to \infty$, $\ln |x| \leq |x|^\zeta$ for any positive value of $\zeta$. Thus, a Taylor expansion of $(1 - x)^p$ for $p > 0$ and $0 \leq x < 1$ yields

$$A_4 = \frac{C}{d \ln d} \sum_{m=1}^{[dx]} \sum_{k=1}^{[dx]-1} \left( 1 - \left( 1 - \frac{m}{k + m} \right)^{2\beta - \alpha} \right) m^{-1}$$

$$\leq \frac{C}{d \ln d} \sum_{m=1}^{[dx]} \sum_{k=1}^{[dx]-1} \frac{m}{k + m} m^{-1}$$

$$\leq \frac{C}{d \ln d} \sum_{m=1}^{[dx]} \sum_{k=1}^{[dx]-1} \frac{1}{y + m - 1} dy$$

$$= \frac{C}{d \ln d} \sum_{m=1}^{[dx]} \ln \frac{|dx|}{m}$$

$$\leq \frac{C}{d \ln d} \sum_{m=1}^{[dx]} \left( \frac{|dx|}{m} \right)^\zeta \leq \frac{Cd^\zeta}{d \ln d} \sum_{m=1}^{[dx]} \left( 1 \right)^\zeta \leq \frac{C}{\ln d}.$$

Since $\tilde{R}_d = o\left( \frac{1}{\ln d} \right)$, the estimate (13) then becomes

$$d_W \left( \tilde{G}_n, \tilde{Z}_n \right) \leq C \sqrt{\frac{n(n + 1)}{\ln d}},$$

and the estimate in Theorem 2 follows from (12). Like for the proof of Theorem 1, the conclusion that $\tilde{W}_{n,\{dx\}}$ is close to the G.O.E matrix $\tilde{Z}_n$ in finite-dimensional distributions follows once again from [23, Theorem 6.2.3].

4. PROOF OF THE MAIN NON-CENTRAL CONVERGENCE RESULT

As pointed out in Theorem 3, the case where $\frac{3}{2} \leq \alpha < 2$ unveils a interesting phenomenon of non-central convergence, giving rise to a limiting object known as the Rosenblatt-Wishart matrix, introduced in [24]. This section is dedicated to the proof of this non-central convergence result.
4.1. Proof of Theorem 3. The self-similarity property of $X$ implies that the entries

$$\hat{Y}_k^i = \frac{X_k^{i+1} - X_k^i}{\|X_k^{i+1} - X_k^i\|_{L^2(\Omega)}}$$

of the $n \times d$ matrix $\hat{Y}$ are equal in distribution to the entries $Y_k^i$ of $\mathcal{Y}$. Given that the statement we aim to prove is distributional, we can work with (keeping the same denomination by a slight abuse of notation) the matrix

$$\tilde{W}_{n,[dx]} = \frac{|dx|}{d^{\alpha-1}} \left( \frac{1}{|dx|} \mathcal{Y} \mathcal{Y}^T - I \right)$$

with entries

$$\tilde{W}_{ij}([dx]) = d^{1-\alpha} \sum_{k=1}^{\lfloor dx \rfloor} \left( \hat{Y}_k^i \hat{Y}_k^j - 1_{\{i=j\}} \right)$$

in place of the original one given by (1). The existence of the limit (in the $L^2(\Omega)$ sense) of $\tilde{W}_{n,[dx]}$ as $d$ goes to infinity, called the Rosenblatt-Wishart matrix $R_n$, is ensured by Lemma 11. To estimate the Wasserstein distance between our Wishart matrix and the Rosenblatt-Wishart matrix, we need the following result from [6] applied to the half-matrices associated with $\tilde{W}_{n,[dx]}$ and $R_n$, which in our context, reads

$$d_W(\tilde{W}_{n,[dx]}, R_n) \leq \sqrt{2 \sum_{1 \leq i \leq j \leq n} \mathbb{E} \left[ (\tilde{W}_{ij}([dx]) - R_{ij})^2 \right]}.$$  \hspace{1cm} (14)

We hence need to evaluate, for all $1 \leq i \leq l \leq n$,

$$\mathbb{E} \left[ (\tilde{W}_{il}([dx]) - R_{il})^2 \right] = \mathbb{E} \left[ \tilde{W}_{il}([dx])^2 \right] - 2 \mathbb{E} \left[ \tilde{W}_{il}([dx]) R_{il} \right] + \mathbb{E} \left[ R_{il}^2 \right].$$  \hspace{1cm} (15)

[16, Lemma 3.2] states that $\mathbb{E} \left[ \left( X_k^{i+1} - X_k^i \right)^2 \right] = 2\lambda(\frac{k^2}{d})^{2\beta - \alpha}(\frac{1}{d})^\alpha (1 + \eta_{k,d})$ where $\eta_{k,d} = o(k^{\alpha-2})$. This fact combined with Lemma 9 allows us to estimate the first term on the right hand side of (15) as

$$\mathbb{E} \left[ \tilde{W}_{il}([dx])^2 \right] = \frac{1}{4\lambda^2} \sum_{1 \leq k, j \leq [dx]} d^{2(1-\alpha)} \left( \frac{k^2}{d} \right)^{2\beta - \alpha} \left( \frac{1}{d} \right)^\alpha (1 + \eta_{k,d})^{-1} \left( \int_{\frac{i}{d}}^{\frac{i+1}{d}} \int_{\frac{j}{d}}^{\frac{j+1}{d}} \partial_{s,t} \mathbb{E}[X_s X_t] ds dt \right)^2$$

$$= \frac{1}{4\lambda^2} \sum_{1 \leq k, j \leq [dx]} d^{2-2(\alpha-2\beta)} \left( \frac{k^2}{d} \right)^{\alpha-2\beta} + o \left( (kj)^{\alpha-2\beta} k^{\alpha-2} \right)$$

$$+ o \left( (kj)^{\alpha-2\beta} j^{\alpha-2} \right) \left( \int_{\frac{i}{d}}^{\frac{i+1}{d}} \int_{\frac{k}{d}}^{\frac{k+1}{d}} \partial_{s,t} \mathbb{E}[X_s X_t] ds dt \right)^2.$$
for which we have used \((1 + \eta_{k,d})^{-1} = 1 + o(k^{\alpha-2})\). Next, we use the substitution \(s = \frac{u + k}{d}, \quad t = \frac{v + j}{d}\) and apply the mean value theorem to get

\[
E\left[\hat{W}_{il}(\lfloor dx \rfloor)^2\right] = \frac{d^{2-2(\alpha-2\beta)}}{4\lambda^2} \sum_{1 \leq k, j \leq \lfloor dx \rfloor} \left((k \alpha-2\beta) + o\left((k \alpha-2\beta)k^{\alpha-2}\right)\right)
\]

\[
+ o\left((k \alpha-2\beta)j^{\alpha-2}\right) \partial_{u,v} E\left[X_{u+\frac{k}{d}} X_{v+\frac{j}{d}}\right]^{2} \bigg|_{u = u_0, v = v_0},
\]

with \(u_0, v_0 \in [a, b]\). For the second term on the right hand side of (15), an application of Lemmas 9 and 10 yields

\[
E\left[\hat{W}_{il}(\lfloor dx \rfloor) R_{il}\right] = \lim_{p \to \infty} E\left[\hat{W}_{il}(\lfloor dx \rfloor) \hat{W}_{il}(p)\right]
\]

\[
= \frac{1}{4\lambda^2} \sum_{1 \leq k, j \leq \lfloor dx \rfloor} d^{1-\alpha} \int_{0}^{1} \left(\left(\frac{k}{d}\right)^{2\beta-\alpha} \left(\frac{1}{d}\right)^{\alpha} \left(1 + \eta_{k,d}\right)\right)^{-1}
\]

\[
\times e^{\alpha-2\beta} \left(\int_{\frac{k}{d}}^{\frac{k+1}{d}} \partial_{s,t} E[X_{s} X_{t}] ds \right)^{2} dt
\]

\[
= \frac{1}{4\lambda^2} \sum_{1 \leq k, j \leq \lfloor dx \rfloor} d^{1-\alpha} \int_{\frac{k}{d}}^{\frac{k+1}{d}} \left(\left(\frac{k}{d}\right)^{2\beta-\alpha} \left(\frac{1}{d}\right)^{\alpha} \left(1 + \eta_{k,d}\right)\right)^{-1}
\]

\[
\times e^{\alpha-2\beta} \left(\int_{\frac{k}{d}}^{\frac{k+1}{d}} \partial_{s,t} E[X_{s} X_{t}] ds \right)^{2} dt
\]

\[
= \frac{d^{2-2(\alpha-2\beta)}}{4\lambda^2} \sum_{1 \leq k, j \leq \lfloor dx \rfloor} \int_{0}^{1} \left(k^{\alpha-2\beta} + k^{(\alpha-2\beta)+(\alpha-2)}\right)
\]

\[
(v + j)^{\alpha-2\beta} \left(\int_{0}^{1} \partial_{u,v} E[X_{u+\frac{k}{d}} X_{v+\frac{j}{d}}] du\right)^{2} dv
\]

\[
= \frac{d^{2-2(\alpha-2\beta)}}{4\lambda^2} \sum_{1 \leq k, j \leq \lfloor dx \rfloor} \left(k^{\alpha-2\beta} + k^{(\alpha-2\beta)+(\alpha-2)}\right)
\]

\[
(v_1 + j)^{\alpha-2\beta} \partial_{u,v} E[X_{u+\frac{k}{d}} X_{v_1+\frac{j}{d}}]^{2} \bigg|_{u = u_1, v = v_1},
\]
where $u_1, v_1 \in [a, b]$. To compute the last term of the right hand side of (15), we use Lemma 9 once more, which allows us to write

\[
\mathbb{E}[R_{ij}^2] = \lim_{d \to \infty} \mathbb{E}[\tilde{W}_{il}([dx])^2]
\]

\[
= \frac{1}{4\lambda^2} \int_0^1 \int_0^1 (st)^{\alpha - 2\beta} (\partial_{s,t}\mathbb{E}[X_sX_t])^2 dsdt
\]

\[
= \frac{1}{4\lambda^2} \sum_{1 \leq k,j \leq |dx|} \int_0^{\frac{k+1}{2}} \int_0^{\frac{j+1}{2}} (st)^{\alpha - 2\beta} (\partial_{s,t}\mathbb{E}[X_sX_t])^2 dsdt
\]

\[
= \frac{d^{2-2(\alpha - 2\beta)}}{4\lambda^2} \sum_{1 \leq k,j \leq |dx|} \int_0^1 \int_0^1 (u + k)^{\alpha - 2\beta} (v + j)^{\alpha - 2\beta} \left( \partial_{u,v}\mathbb{E}\left[X_{u+k}X_{v+j}\right] \right)^2 dudv
\]

\[
= \frac{d^{2-2(\alpha - 2\beta)}}{4\lambda^2} \sum_{1 \leq k,j \leq |dx|} (u_2 + k)^{\alpha - 2\beta} (v_2 + j)^{\alpha - 2\beta} \left( \partial_{u,v}\mathbb{E}\left[X_{u+k}X_{v+j}\right] \right)^2 |_{u = u_2, v = v_2},
\]

where $u_2, v_2 \in [a, b]$. Therefore, combining these estimates, we get

\[
\mathbb{E}\left[\left(\tilde{W}_{il}([dx]) - R_{il}\right)^2\right] \leq E_1 + E_2,
\]

where

\[
E_1 = Cd^{2-2(\alpha - 2\beta)} \sum_{1 \leq k,j \leq |dx|} \sup_{u_3, v_3 \in [a, b]} \left. \partial_{u,v}\mathbb{E}\left[X_{u+k}X_{v+j}\right]^2 \right|_{u = u_3, v = v_3} \left( (kj)^{\alpha - 2\beta} + (u_2 + k)^{\alpha - 2\beta} (v_2 + j)^{\alpha - 2\beta} - 2k^{\alpha - 2\beta} (v_1 + j)^{\alpha - 2\beta} \right)
\]

and

\[
E_2 = Cd^{2-2(\alpha - 2\beta)} \sum_{1 \leq k,j \leq |dx|} (kj)^{\alpha - 2\beta} \left( \left. \partial_{u,v}\mathbb{E}\left[X_{u+k}X_{v+j}\right]^2 \right|_{u = u_0, v = v_0} + \left. \partial_{u,v}\mathbb{E}\left[X_{u+k}X_{v+j}\right]^2 \right|_{u = u_2, v = v_2} - 2 \left. \partial_{u,v}\mathbb{E}\left[X_{u+k}X_{v+j}\right]^2 \right|_{u = u_1, v = v_1} \right).
\]

For $E_1$ we can use the estimate on $\partial_{s,t}\mathbb{E}[X_sX_t]$ in Lemma 9 and the symmetry of $j \leq k$ and $j \geq k$ to get

\[
E_1 \leq Cd^{2-2(\alpha - 2\beta)} \sum_{1 \leq j \leq k \leq |dx|} j^{2(\beta - \alpha)} k^{2(\alpha - 2)} d^{4\beta} \left( k^{\alpha - 2\beta} j^{\alpha - 2\beta - 1} \right)
\]

\[
\quad + Cd^{2-2(\alpha - 2\beta)} \sum_{1 \leq k \leq j \leq |dx|} j^{2(\beta - \alpha)} k^{2(\alpha - 2)} d^{4\beta} \left( k^{\alpha - 2\beta} j^{\alpha - 2\beta - 1} \right)
\]

\[
\leq Cd^{2-2\alpha} \sum_{k=1}^{|dx|} \frac{|dx|}{4} k^{2(\alpha - 2)} j^{-1} + Cd^{2-2\alpha} \sum_{j=1}^{|dx| - 1} \frac{|dx| - 1}{4} \sum_{k=j+1}^{2j} |k - j|^{2(\alpha - 2)} j^{-1}
\]

\[
\leq Cd^{2-2\alpha} \sum_{k=1}^{|dx|} \frac{k^{2(\alpha - 2)}}{4} \ln k + d^{2-2\alpha} \sum_{j=1}^{|dx| - 1} j^{2(\alpha - 2)}.
\]
As \( k \to \infty \), it holds that \( \ln k \leq k^\theta \) for any \( \theta \in (0, 1) \), so that
\[
E_1 \leq C d^{2 - 2\alpha} \sum_{k=1}^{[dx]} k^{2(\alpha - 2) + \theta} \leq C d^{\theta - 1},
\]
at which point we can take the infimum of this estimate over all \( \theta \in (0, 1) \) to get \( E_1 \leq \frac{C}{d} \). For \( E_2 \), in a similar fashion, we apply Lemma 9 to obtain \( E_2 \leq \frac{Cd^2}{2 - 2\alpha} \). Keeping in mind that \( \frac{3}{2} < \alpha < 2 \) is equivalent to \(-1 < 3 - 2\alpha < 0\), (14) implies that
\[
(16) \quad d_W\left(\hat{W}_{n,[dx]}, R_n\right) \leq C nd^\frac{3 - 2\alpha}{2}.
\]

Now to conclude that \( \hat{W}_{n,[dx]} \) and \( R_n \) are close with respect to finite-dimensional distributions, we need a slightly different version of (14), which can for example be found in [6]. Let \((x_1, \ldots, x_p)\) be a finite sequence in \([a,b]^p\), then we have
\[
d_W\left(\left(\hat{W}_{n,[dx_1]}, \ldots, \hat{W}_{n,[dx_p]}\right), (R_n, \ldots, R_n)\right) \leq \left( \sum_{1 \leq i \leq j \leq n} \sum_{1 \leq l \leq p} \mathbb{E} \left[ \left( \hat{W}_{ij}([dx_i]) - R_{ij} \right)^2 \right] \right)^{\frac{1}{2}},
\]
for which the same estimate as in (16) clearly holds.

5. Proof of the main functional convergence result

This section is dedicated to proving the main functional convergence result, Theorem 4, which provides a functional counterpart to Theorems 1, 2 and 3.

5.1. Proof of theorem 4. Our goal here is to show that our sequence of Wishart matrices converges in \( C([a,b]; M_n(\mathbb{R})) \) (without providing an estimate on the convergence rate). Recall from the introduction that \( a < b \) are positive constants. Joint convergence of a vector in \( C([a,b]; M_n(\mathbb{R})) \) is equivalent to marginal convergence of each component in \( C([a,b]; \mathbb{R}) \) per [13, Theorem 26.23], so that we only have to prove functional convergence of each entry of our matrix. Furthermore, it is a well-known fact that the condition
\[
||W_{ij}([dx]) - W_{ij}([dy])||_{L^p(\Omega)} \leq C_p |x - y|^\frac{1}{2}
\]
for some \( p > 2 \) and any \( x, y \in [a,b] \), combined with convergence of finite dimensional distributions is sufficient in order to guarantee tightness in \( C([a,b]; \mathbb{R}) \). Convergence in the sense of finite dimensional distributions has already been shown in Theorems 1, 2 and 3, so that we only need to verify the above condition ensuring tightness. In what follows, we reuse the notation and terminology introduced at the beginning of the proof of Theorem 1 in Subsection 3.1.

Case 1 (\( \alpha < 3/2 \)). Note that by definition of multiple Wiener integrals (see [23, Definition 2.7.1]),
\[
I_2(e_{ik} \otimes e_{jk}) = I_2(e_{ik} \otimes e_{jk}) = \frac{1}{2} \delta^2(e_{ik} \otimes e_{jk} + e_{jk} \otimes e_{ik}).
\]
Now, assume \( y \leq x \) (without loss of generality) for \( x, y \in [a, b] \). Apply the above transform together with Meyer’s inequality (5) to get

\[
\|W_{ij}([dx]) - W_{ij}([dy])\|_{L^p(\Omega)} = \left\| \delta^2 \left( \frac{1}{2 \sqrt{d}} \sum_{k=|dy|}^{dx} (e_{ik} \otimes e_{jk} + e_{ijk} \otimes e_{ik}) \right) \right\|_{L^p(\Omega)} \\
\leq C \sum_{m=0}^2 \left\| D^m \left( \frac{1}{2 \sqrt{d}} \sum_{k=|dy|}^{dx} (e_{ik} \otimes e_{jk} + e_{ijk} \otimes e_{ik}) \right) \right\|_{L^p(\Omega; S^{m+2})} \\
= C \left( d^{-1} \sum_{k=|dy|}^{dx} \delta_{kl}^2 \right)^{\frac{1}{2}}.
\]

By Remark 9 and the fact that \( a_\alpha(k - l) \leq C |k - l|^{\alpha - 2} \) as \( |k - l| \to \infty \), we have

\[
\|W_{ij}([dx]) - W_{ij}([dy])\|_{L^p(\Omega)} \leq C \left( d^{-1} \sum_{k,l=|dy|}^{dx} \left( \frac{k \land l}{k \lor l} \right)^{2\beta - \alpha} |k - l|^{2(\alpha - 2)} \right)^{\frac{1}{2}} \\
\leq C \left( d^{-1} \sum_{k,l=|dy|}^{dx} |k - l|^{2(\alpha - 2)} \right)^{\frac{1}{2}} \\
= C \left( d^{-1} \sum_{m=k-1}^{dx} (|dx| - |dy| - m) m^{2(\alpha - 2)} \right)^{\frac{1}{2}} \\
\leq C \left( \sum_{m \in \mathbb{Z}} |m|^{2(\alpha - 2)} \right)^{\frac{1}{2}} (x - y)^{\frac{1}{2}} \leq C (x - y)^{\frac{1}{2}},
\]

which is the desired estimate.

**Case 2 (\( \alpha = 3/2 \)).** The same procedure as in the previous case gives

\[
\left\| \hat{W}_{ij}([dx]) - \hat{W}_{ij}([dy]) \right\|_{L^p(\Omega)} \leq C \left( (\ln d)^{-1} \sum_{k,l=|dy|}^{dx} |k - l|^{-1} \right)^{\frac{1}{2}} \\
= C \left( (\ln d)^{-1} \sum_{m=1}^{y\lfloor - [dy] \rfloor} (|dx| - |dy| - m) m^{-1} \right)^{\frac{1}{2}} \\
\leq C \left( (\ln d)^{-1} \sum_{m=1}^{y\lfloor - [dy] \rfloor} |m|^{-1} \right)^{\frac{1}{2}} (x - y)^{\frac{1}{2}} \leq C (x - y)^{\frac{1}{2}},
\]

which gives the desired result.
Case 3 ($\alpha > 3/2$). For this final case, the above argument yields

$$
\left\| W_{ij}((dx)) - \hat{W}_{ij}((dy)) \right\|_{L^p(\Omega)} \leq C \left( d^{2-2\alpha} \sum_{k,l=|dy|} |k-l|^{2(\alpha-2)} \right)^{\frac{1}{2}}
$$

$$
= C \left( d^{2-2\alpha} \sum_{m=1}^{|dy|} (|dx| - |dy| - m)^{2(\alpha-2)} \right)^{\frac{1}{2}}
$$

$$
\leq C \left( \int_0^1 u^{2(\alpha-2)} du \right)^{\frac{1}{2}} (x-y)^{\frac{1}{2}} \leq C(x-y)^{\frac{1}{2}},
$$

which concludes the proof.

6. Technical Lemmas

This section gathers technical Lemmas used repeatedly in the proofs of our main results. For convenience, we group these auxiliary results by what proof they are related to. The notation used in all the results below is the one prevailing in Section 1.

6.1. Lemmas related to the proof of Theorem 1 in Subsection 3.1.

Lemma 2. The covariance structure of the half-matrix $W_{n,|dx|}^{\text{half}}$ is given by

$$
\begin{cases}
E[W_{il}W_{il}] = d^{-1} \sum_{k,j=1}^{|dx|} \delta_{kj}^2 & \text{for } i \neq l \\
E[W_{il}W_{il}] = 2d^{-1} \sum_{k,j=1}^{|dx|} \delta_{kj}^2 & \\
E[W_{il}W_{mn}] = 0 & \text{otherwise}
\end{cases}
$$

Thus, if we denote $C$ the covariance matrix of $W_{n,|dx|}^{\text{half}}$, then $C$ is diagonal with diagonal entries given by either $E[W_{il}W_{il}]$ or $E[W_{il}W_{il}]$.

Proof. For any $1 \leq i, l, m, n \leq n$, recalling the representation (7) of $W_{il}$, it holds that

$$
E[W_{il}W_{mn}] = E \left[ I_2 \left( \frac{1}{2\sqrt{d}} \sum_{k=1}^{|dx|} e_{ik} \otimes e_{ik} + e_{lk} \otimes e_{ik} \right) I_2 \left( \frac{1}{2\sqrt{d}} \sum_{j=1}^{|dx|} e_{mj} \otimes e_{nj} + e_{nj} \otimes e_{mj} \right) \right]
$$

$$
= 2! \left( \frac{1}{2\sqrt{d}} \sum_{k=1}^{|dx|} (e_{ik} \otimes e_{ik} + e_{lk} \otimes e_{ik}), \frac{1}{2\sqrt{d}} \sum_{j=1}^{|dx|} (e_{mj} \otimes e_{nj} + e_{nj} \otimes e_{mj}) \right)_{\mathcal{S}_2^{\otimes 2}}
$$

$$
= d^{-1} \sum_{k,j=1}^{|dx|} \langle e_{ik}, e_{mj} \rangle_{\mathcal{S}_2^{\otimes 2}} \langle e_{ik}, e_{nj} \rangle_{\mathcal{S}_2^{\otimes 2}} + d^{-1} \sum_{k,j=1}^{|dx|} \langle e_{ik}, e_{mj} \rangle_{\mathcal{S}_2^{\otimes 2}} \langle e_{lk}, e_{mj} \rangle_{\mathcal{S}_2^{\otimes 2}}.
$$

This shows that the only entries of the matrix $C$ that are non-zero are the ones for which $i = m$ and $l = n$ (note that we cannot encounter the case $i = n$ and $l = m$ as we are working with the half-matrix $W_{n,|dx|}^{\text{half}}$). This corresponds to entries of the form $E[W_{il}W_{il}]$. We hence only have
Lemma 3. For any $1 \leq i, j \leq n$, $W_{ij}$ belongs to the second Wiener chaos of the isonormal Gaussian process $X$ and has the representation (7) as a double Wiener integral, so that $W_{ij} = I_2(f_{ij})$, where $f_{ij} \in \mathcal{S}^{\otimes 2}$. Furthermore, for any $1 \leq i, j, m, n \leq n$, it holds that

$$E[W_{ij} W_{mn}] = E\left[\left(\frac{1}{2} \langle DW_{ij}, DW_{mn} \rangle^2 \right) \right] = 8 \| f_{ij} \otimes_1 f_{mn} \|_{\mathcal{S}^{\otimes 2}}^2.$$

Proof. Using the product formula (3) together the stochastic Fubini theorem, it is straightforward to check that $\langle DW_{ij}, DW_{mn} \rangle^2 = 4I_2(f_{ij} \otimes_1 f_{mn}) + 4 \langle f_{ij}, f_{mn} \rangle_{\mathcal{S}^{\otimes 2}}$, and hence deduce that $E[W_{ij} W_{mn}] = \frac{1}{4} E[\langle DW_{ij}, DW_{mn} \rangle^2]$, so that

$$E\left[\left(E[W_{ij} W_{mn}] - \frac{1}{2} \langle DW_{ij}, DW_{mn} \rangle^2 \right)^2 \right] = E\left[\frac{1}{2} E[\langle DW_{ij}, DW_{mn} \rangle^2] - \frac{1}{2} \langle DW_{ij}, DW_{mn} \rangle^2 \right]^2$$

$$= \frac{1}{4} \text{Var}(\langle DW_{ij}, DW_{mn} \rangle^2)$$

$$= 4\text{Var}(I_2(f_{ij} \otimes_1 f_{mn}) + \langle f_{ij}, f_{mn} \rangle_{\mathcal{S}^{\otimes 2}})$$

$$= 4E\left[I_2(f_{ij} \otimes_1 f_{mn})^2 \right]$$

$$= 8 \| f_{ij} \otimes_1 f_{mn} \|_{\mathcal{S}^{\otimes 2}}^2.$$

Lemma 4. For $f_{ij}$ and $f_{rs}$ as defined in (8), it holds that for any $1 \leq i, j, r, s \leq n$,

$$\| f_{ij} \otimes_1 f_{rs} \|_{\mathcal{S}^{\otimes 2}}^2 = \frac{1}{\tau d^2} \sum_{k,l,m,p=1}^{dx} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp},$$

where $\tau = 1$ if $i = j = r = s$, $\tau = 4$ if $i = j = r \neq s$ or $i = j = s \neq r$ or $i = r = s \neq j$ or $j = r = s \neq i$, $\tau = 8$ if $i = r \neq s$, $j = s \neq i$ or $i = s \neq j$, $j = r \neq s$ or $j = r \neq i$, $i = s \neq r$ and
and \( \tau = 16 \) if \( i = r, \ i \neq j, \ s \neq j, \ r \neq s \) or \( i = s, \ i \neq j, \ r \neq j \) or \( j = r, \ j \neq i, \ r \neq s \) or \( j = s, \ j \neq i, \ j \neq r, \ r \neq i \). In all other cases,
\[
\left\| f_{ij} \otimes_1 f_{rs} \right\|_{\mathcal{B}^2}^2 = 0.
\]

**Proof.** Recalling the definition of \( f_{ij} \) given in (8), we can write, for any \( 1 \leq i, j, r, s \leq n \),
\[
\left\| f_{ij} \otimes_1 f_{rs} \right\|_{\mathcal{B}^2}^2 = \left\| \frac{1}{2^n d} \sum_{k=1}^{[dx]} (e_{ik} \otimes e_{jk} + e_{jk} \otimes e_{ik}) \otimes_1 \frac{1}{2^n d} \sum_{t=1}^{[dx]} (e_{rl} \otimes e_{sl} + e_{sl} \otimes e_{rl}) \right\|_{\mathcal{B}^2}^2
\]
\[
= \frac{1}{16d^2} \sum_{k,l=1}^{[dx]} \left( e_{jk} \otimes e_{sl} \delta_{kl} \mathbb{I}_{\{i=r\}} + e_{jk} \otimes e_{rl} \delta_{kl} \mathbb{I}_{\{j=s\}} + e_{ik} \otimes e_{sl} \delta_{kl} \mathbb{I}_{\{i=r\}} + e_{ik} \otimes e_{rl} \delta_{kl} \mathbb{I}_{\{j=s\}} \right)
\]
\[
\left( e_{jk} \otimes e_{sl} \delta_{kl} \mathbb{I}_{\{i=r\}} + e_{jk} \otimes e_{rl} \delta_{kl} \mathbb{I}_{\{j=s\}} + e_{ik} \otimes e_{sl} \delta_{kl} \mathbb{I}_{\{i=r\}} + e_{ik} \otimes e_{rl} \delta_{kl} \mathbb{I}_{\{j=s\}} \right),
\]
from which the conclusion follows easily. \( \square \)

The following Lemma is borrowed from [16] and provides us with the asymptotic behaviour of the variance of the entries of \( \mathcal{W}_{n,[dx]} \).

**Lemma 5.** Denote by \( \sigma^2_{[dx]} \) the variance of the entries \( W_{ij}, i \neq j \) of \( \mathcal{W}_{n,[dx]} \). Let \( \alpha < \frac{3}{2} \) and define \( a_\alpha(m) = \frac{1}{2} (|m+1|^{\alpha} + |m-1|^{\alpha} - 2|m|^\alpha) \). Then,
\[
\sigma^2_{[dx]} = \frac{1}{d} \sum_{k,l=1}^{[dx]} |\delta_{kl}|^2
\]
and
\[
\sigma^2 = \lim_{d \to \infty} \sigma^2_{[dx]} = \frac{x}{2} \sum_{m \in \mathbb{Z}} a_\alpha(m)^2.
\]

**Proof.** Refer to the proof of [16, Lemma 5.1]. \( \square \)

**Remark 9.** An important observation coming from the proof of [16, Lemma 5.1] is that
\[
\sigma^2_{[dx]} = \left( \frac{1}{d} \sum_{m=1}^{[dx]-1} \sum_{k=1}^{m-1} \left( \frac{k}{k+m} \right)^{2\beta-\alpha} a_\alpha(m)^2 + \frac{x}{2} a_\alpha(0) + R_d, \right.
\]
where \( R_d \) is a remainder term with \( R_d = o(d^{2\alpha-3} + d^{-1}) \).

**Lemma 6.** Assume that \( \alpha < 1 \) and \( \alpha + \nu > 2 \), or that \( 1 \leq \alpha < \frac{3}{2} \). Then, it holds that
\[
\frac{1}{d^2} \sum_{k,l,m,p=1}^{[dx]} |\delta_{kl} | \delta_{km} | \delta_{lp} | ^2 \leq C \frac{1}{d^2} \sum_{k,l,m,p=1}^{[dx]} |a_\alpha(k-l) a_\alpha(m-p) a_\alpha(k-m) a_\alpha(l-p)|,
\]
where as in Lemma 5, \( a_\alpha(i) = \frac{1}{2}(|i + 1|^\alpha + |i - 1|^\alpha - 2|i|^\alpha) \). As a result, we have

\[
\frac{1}{d^2} \sum_{k,l,m,p=1}^{\lfloor dx \rfloor} \delta_{kl}\delta_{mp}\delta_{km}\delta_{lp} \leq \frac{C}{d} \left( \sum_{m=-\lfloor dx \rfloor+1}^{\lfloor dx \rfloor-1} |a_\alpha(m)|^{\frac{4}{d}} \right)^3.
\]

**Proof.** In order to obtain the first estimate, it is sufficient to show that \( |\delta_{kl}| \leq C |a_\alpha(k-l)| \) for any \( 1 \leq k \leq l \leq [dx] \). By symmetry, we also only need to examine the case where \( l \leq k \), which we separate into three separate cases. If \( \lfloor \frac{k}{2} \rfloor \leq l \leq k-2 \), which implies that \( k-l \leq 2l \), then \( [16, \text{Lemma 3.1, Lemma 3.2, Part (b)}] \) implies that there exist a constant \( a_\alpha \leq \frac{1}{2} \alpha \alpha-1 (k-l) \alpha-1 + \alpha 2 \alpha-2 \). Since \( a_\alpha(k-l) = \frac{1}{2} \alpha \alpha-1 (k-l) \alpha-2 + o((k-l) \alpha-2) \), it follows that for \( k-l \leq 2l \), \( l^{-1}(k-l) \alpha-1 \leq Ca_\alpha(k-l) \) and \( l^{\alpha-2} \leq Ca_\alpha(k-l) \). Thus,

\[
|\delta_{kl}| \leq C |a_\alpha(k-l)|.
\]

If \( 1 \leq l \leq \lfloor \frac{k}{2} \rfloor \), which implies that \( 2l \leq k-l+3 < C(k-l) \), then \( [16, \text{Lemma 3.1, Lemma 5.1}] \) yields, whenever \( \alpha + \nu > 2 \),

\[
|\delta_{kl}| \leq C(lk)^{\frac{\alpha}{2} - \beta} \left( l^{2\beta-\alpha} a_\alpha(k-l) + l^{2\beta-\alpha-1}(k-l) \alpha-1 + l^{2\beta-2} \right) \\
\leq C \left( (l^{\alpha+\nu-2}(k-l)^{-\nu}) \lor (k-l)^{\alpha-2} \right) \\
\leq C |a_\alpha(k-l)|.
\]

If \( k = l \), then \( \delta_{lk} = a_\alpha(l-k) = 1 \). Now, consider the case when \( l = k-1 \). Since \( a_\alpha(1) = 2^{\alpha-1} - 1 \), \( [16, \text{Lemma 3.1, Lemma 3.2, Part (a)}] \) implies that there exists a constant \( C > 0 \) such that for \( k \) large enough, one has

\[
|\delta_{(k-1)k}| \leq C \left( \frac{k}{k-1} \right)^{\beta - \frac{\alpha}{2}} \leq C |a_\alpha(1)|.
\]

Combining all three of these cases yields

\[
\frac{1}{d^2} \sum_{k,l,m,p=1}^{\lfloor dx \rfloor} |\delta_{kl}\delta_{mp}\delta_{km}\delta_{lp}| \leq \frac{1}{d^2} \sum_{k,l,m,p=1}^{\lfloor dx \rfloor} |\delta_{kl}\delta_{mp}\delta_{km}\delta_{lp}| \\
\leq \frac{C}{d^2} \sum_{k,l,m,p=1}^{\lfloor dx \rfloor} |a_\alpha(k-l)a_\alpha(m-p)a_\alpha(k-m)a_\alpha(l-p)|.
\]

The second estimate is due to a result in \([23, \text{Pages } 134–135]\), which states

\[
\frac{C}{d^2} \sum_{k,l,m,p=1}^{\lfloor dx \rfloor} a_\alpha(k-l)a_\alpha(m-p)a_\alpha(k-m)a_\alpha(l-p) \\
\leq \frac{C}{d^2} \sum_{k,l,m,p=1}^{\lfloor dx \rfloor} |a_\alpha(k-l)a_\alpha(m-p)a_\alpha(k-m)a_\alpha(l-p)| \leq \frac{C}{d^2} \left( \sum_{m=-\lfloor dx \rfloor+1}^{\lfloor dx \rfloor-1} |a_\alpha(m)|^{\frac{4}{d}} \right)^3.
\]

Finally, note that the fact that we impose \( \alpha < 1 \) and \( \alpha + \nu > 2 \), or that \( 1 \leq \alpha < \frac{3}{2} \) is due to our hypothesis that \( 1 < \nu \leq 2 \). \( \square \)
6.2. Lemmas related to the proof of Theorem 2 in Subsection 3.2.

Lemma 7. Assume that \( \alpha = \frac{3}{2} \). Let \( a_\alpha(m) = \frac{1}{2}(|m+1|^{\alpha} + |m-1|^{\alpha} - 2|m|^{\alpha}). \) Then,

\[
\frac{1}{(d \ln d)^2} \sum_{k,l,m,p=1}^{[dx]} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp} \leq \frac{C}{(d \ln d)^2} \sum_{k,l,m,p=1}^{[dx]} |a_\alpha(k-l)a_\alpha(m-p)a_\alpha(k-m)a_\alpha(l-p)|.
\]

As a result,

\[
\frac{1}{(d \ln d)^2} \sum_{k,l,m,p=1}^{[dx]} \delta_{kl} \delta_{mp} \delta_{km} \delta_{lp} \leq \frac{C}{d(d \ln d)^2} \left( \sum_{m=-[dx]+1}^{[dx]-1} \left| a_\alpha'(m) \right|^\frac{3}{2} \right)^3.
\]

Proof. The proof follows in the exact same way as the proof of Lemma 6. \( \square \)

Lemma 8. Denote by \( \rho^2_{[dx]} \) the variance of the non-diagonal entries of \( \tilde{W}_{\alpha, [dx]} \). Then,

\[
\rho^2_{[dx]} = \frac{1}{d \ln d} \sum_{k,l=1}^{[dx]} |\delta_{kl}|^2
\]

and

\[
\rho^2 = \lim_{d \to \infty} \rho^2_{[dx]} = \frac{9x}{32}.
\]

Proof. We adapt the ideas in [16, Proof of Lemma 5.1] which does not cover the case \( \alpha = \frac{3}{2} \).

Step 1. Define \( \xi_{j,d} = \left\| \Delta X_{\frac{j}{2}} \right\|_{L^2(\Omega)}. \) Choose \( \gamma \in (0, \frac{1}{2}) \) and let \( \tau = ([dx])^\gamma \). We will perform the decomposition

\[
\rho^2_{[dx]} = \frac{1}{d \ln d} \sum_{k,l=1}^{[dx]} |\delta_{kl}|^2 = A_{1,d} + A_{2,d} + A_{3,d} + A_{4,d},
\]

where

\[
A_{1,d} = \frac{1}{d \ln d} \sum_{j \in D_1, k \in D_1} \xi_{j,d,k,d}^{-2} \left( E \left[ \Delta X_{\frac{j}{2}} \Delta X_{\frac{k}{2}} \right] \right)^2,
\]

\[
A_{2,d} = \frac{1}{d \ln d} \sum_{j \in D_2, k \in D_2} \xi_{j,d,k,d}^{-2} \left( E \left[ \Delta X_{\frac{j}{2}} \Delta X_{\frac{k}{2}} \right] \right)^2,
\]

\[
A_{3,d} = \frac{1}{d \ln d} \sum_{j \in D_1, k \in D_2} \xi_{j,d,k,d}^{-2} \left( E \left[ \Delta X_{\frac{j}{2}} \Delta X_{\frac{k}{2}} \right] \right)^2,
\]

\[
A_{4,d} = \frac{1}{d \ln d} \sum_{j \in D_2, k \in D_1} \xi_{j,d,k,d}^{-2} \left( E \left[ \Delta X_{\frac{j}{2}} \Delta X_{\frac{k}{2}} \right] \right)^2,
\]

with

\[
D_1 = \{ l : 1 \leq l \leq \tau \land [dx] \}
\]

\[
D_2 = \{ l : \tau < l \leq [dx] \}.
\]

\( A_{1,d} \) is bounded by \( Cd^{2\gamma-1}(\ln d)^{-1} \) by the Cauchy-Schwarz inequality, so that it converges to zero as \( d \) goes to infinity.
Step 2. We further decompose and bound $A_{2,d}$ and $A_{4,d}$. $A_{3,d}$ can be bounded in the same way as $A_{4,d}$. We have

$$A_{2,d} = \frac{1}{d \ln d} \left( [dx] - \tau \right) + \frac{1}{d \ln d} \sum_{j,k \in \mathbb{D}_2: |j-k|=1} \xi_{j,k}^{-2} \left( \mathbb{E} \left[ \Delta X_{\frac{j}{d}} \Delta X_{\frac{k}{d}} \right] \right)^2$$

$$+ \frac{1}{d \ln d} \sum_{j,k \in \mathbb{D}_2: |j-k| \geq 2} \xi_{j,k}^{-2} \left( \mathbb{E} \left[ \Delta X_{\frac{j}{d}} \Delta X_{\frac{k}{d}} \right] \right)^2$$

$$= B_{1,d}^{(1)} + B_{2,d}^{(1)} + B_{3,d}^{(1)}$$

and

$$A_{4,d} = \frac{1}{d \ln d} s_{\tau,d}^{-2} \xi_{\tau,d}^{-2} \left( \mathbb{E} \left[ \Delta X_{\frac{1}{d}} \Delta X_{\frac{-1}{d}} \right] \right)^2$$

$$+ \frac{1}{d \ln d} \sum_{j \in \mathbb{D}_2, k \in \mathbb{D}_1: |j-k| \geq 2} \xi_{j,k}^{-2} \left( \mathbb{E} \left[ \Delta X_{\frac{j}{d}} \Delta X_{\frac{k}{d}} \right] \right)^2$$

$$= B_{1,d}^{(2)} + B_{2,d}^{(2)}.$$

$B_{1,d}^{(1)}$ clearly goes to 0 as $d \to \infty$. $B_{1,d}^{(2)}$ and $B_{2,d}^{(1)}$ also go to 0 by the Cauchy-Schwarz inequality. For $B_{2,d}^{(1)}$ in particular, we can write

$$\frac{1}{d \ln d} \sum_{j,k \in \mathbb{D}_2: |j-k|=1} \xi_{j,k}^{-2} \left( \mathbb{E} \left[ \Delta X_{\frac{j}{d}} \Delta X_{\frac{k}{d}} \right] \right)^2 \leq \frac{1}{d \ln d} \left( [dx] - \tau \right).$$

Step 3. In this step, we argue that all terms with $k < \left\lfloor \frac{j}{d} \right\rfloor$ have no contribution to $\rho^2$ as $d \to \infty$. The case $j \leq \left\lfloor \frac{k}{d} \right\rfloor$ can be treated similarly. [16, Lemma 5.1] gives the bound for $k < \left\lfloor \frac{j}{d} \right\rfloor$

$$\left( \mathbb{E} \left[ \Delta X_{\frac{j}{d}} \Delta X_{\frac{k}{d}} \right] \right)^2 \leq C d^{-4 \beta} k^{4 \beta - 3} (j - k)^{-1}.$$

Meanwhile, [16, Lemma 3.1] states that

$$\xi_{j,k}^2 = 2 \lambda j^{2 \beta - \frac{3}{d} d^{-2 \beta}} (1 + \eta_{j,d})$$

$$\xi_{j,k}^2 = 2 \lambda k^{2 \beta - \frac{3}{d} d^{-2 \beta}} (1 + \eta_{k,d}),$$

such that $\eta_{j,d} \leq C d^{-\gamma d}$ and $\eta_{k,d} \leq C d^{-\gamma d}$. Hence,

$$\frac{1}{d \ln d} \sum_{j=3}^{\left\lfloor \frac{k}{d} \right\rfloor} \sum_{k=1}^{\left\lfloor \frac{j}{d} \right\rfloor} \xi_{j,k}^{-2} \left( \mathbb{E} \left[ \Delta X_{\frac{j}{d}} \Delta X_{\frac{k}{d}} \right] \right)^2 \leq \frac{C}{d \ln d} \sum_{j=3}^{\left\lfloor \frac{k}{d} \right\rfloor} \sum_{k=1}^{\left\lfloor \frac{j}{d} \right\rfloor} (j - k)^{-1}$$

$$\leq \frac{C}{d \ln d} \sum_{j=3}^{\left\lfloor \frac{k}{d} \right\rfloor} \int_1^{\left\lfloor \frac{j}{d} \right\rfloor} (j - y)^{-1} dy$$

$$\leq \frac{C}{d \ln d} \sum_{j=3}^{\left\lfloor \frac{k}{d} \right\rfloor} \ln (1 - [1/3])$$

$$\leq \frac{C}{d \ln d}.$$

Step 4. In this step, we study those terms which belongs to $B_{3,d}^{(1)}$ and $B_{2,d}^{(2)}$ and were not considered in Step 3. In the case $k \leq j - 2$, we use the covariance representation from [16,
Lemma 3.1 and Lemma 3.2] in order to get
\[
C_d = \frac{1}{d \ln d} \sum_{j \in D_2, \lfloor \frac{j}{3} \rfloor \leq k \leq j-2} \xi_{j,d}^{-2} \xi_{k,d}^{-2} \left( E \left[ \Delta X_{j,d} \Delta X_{k,d} \right] \right)^2
\]
\[
= \frac{1}{d \ln d} \sum_{j \in D_2, \lfloor \frac{j}{3} \rfloor \leq k \leq j-2} \left( \left( \frac{k}{j} \right)^{\beta - 3/4} \alpha_{\frac{3}{2}} (j-k) + R_{j,k} \right)^2
\]
\[
= \frac{1}{d \ln d} \sum_{j \in D_2, \lfloor \frac{j}{3} \rfloor \leq k \leq j-2} \left( \left( \frac{k}{j} \right)^{2\beta - 3/2} \alpha_{\frac{3}{2}} (j-k) \right)
\]
\[
+ \left( \frac{2}{d \ln d} \sum_{j \in D_2, \lfloor \frac{j}{3} \rfloor \leq k \leq j-2} \left( \left( \frac{k}{j} \right)^{\beta - 3/4} \alpha_{\frac{3}{2}} (j-k) R_{j,k} + R_{j,k}^2 \right) \right)
\]
\[= D_d + O_d,
\]
where according to [16, Lemma 3.2] and the fact that \( \frac{1}{3} \leq k \leq j-2 \Rightarrow j-k \leq 2k, \)
\[R_{j,k} \leq C \left( \frac{k}{j} \right)^{\beta - 3/4} k^{-1}(j-k-1)^{\frac{2}{3}} + C \left( \frac{k}{j} \right)^{\beta - 3/4} k^{-\frac{2}{3}} \leq C k^{-\frac{2}{3}}.
\]
Now since \( a_o(m) = \frac{3}{8} m^{-1/2} + o(m^{-1/2}) \), \( O_d \) can be bounded via
\[
O_d \leq \frac{C}{d \ln d} \sum_{j=\lceil \frac{d \tau}{3} \rceil} \sum_{k=\lfloor \frac{j}{3} \rfloor} (j-k)^{-\frac{2}{3} k^{-1/2}}
\]
\[
\leq \frac{C}{d \ln d} \sum_{k=1} d\sum_{j=k+2} (j-k)^{-\frac{2}{3} k^{-1/2}} \leq \frac{C}{d},
\]
which vanishes as \( d \to \infty \).

**Step 5.** The last term \( D_d \) is the only one with a non-trivial contribution to \( \rho_{[dx]}^2 \) as \( d \to \infty \).

We will show that
\[
(17) \quad \lim_{d \to \infty} D_d = \lim_{d \to \infty} \frac{1}{d \ln d} \sum_{j=3}^{d} \sum_{k=1}^{j-2} \left( \frac{k}{j} \right)^{2\beta - 3} a_{\frac{3}{2}}^2 (j-k) = \frac{9x}{64}.
\]
Since \( a_o(m) = \frac{3}{8} m^{-\frac{1}{2}} + \delta_{\frac{3}{2}} (m) \), where \( \delta_{\frac{3}{2}} (m) = o(m^{-\frac{1}{2}}) \), it follows that
\[
D_d' = \lim_{d \to \infty} \frac{1}{d \ln d} \sum_{j=3}^{d} \sum_{k=1}^{j-2} a_{\frac{3}{2}}^2 (j-k) = \lim_{d \to \infty} \frac{1}{d \ln d} \sum_{m=1}^{d} \sum_{k=1}^{m-1} a_{\frac{3}{2}}^2 (m)
\]
\[
= \lim_{d \to \infty} \frac{1}{d \ln d} \sum_{m=1}^{d} \frac{[dx] - m - 1}{d} \left( \frac{9}{64} m^{-1} + \frac{3}{4} m^{-\frac{3}{2}} \delta_{\frac{3}{2}} (m) + \delta_{\frac{3}{2}}^2 (m) \right)
\]
\[
= \lim_{d \to \infty} \frac{9}{64 \ln d} \sum_{m=1}^{[dx]} \frac{[dx] - m - 1}{d} m^{-1} + E_d.
\]
The fact that \( \lim_{m \to \infty} \frac{\delta_{3/2}(m)}{m^{-1/2}} = 0 \) implies that for \( \epsilon < 1 \), there exists \( M_\epsilon \in \mathbb{N} \) such that \( \delta_{3/2}(m) \leq \epsilon m^{-1/2} \) for all \( m \geq M_\epsilon \). In addition, this means that \( \delta_{3/2}(m) \) is bounded by some constant \( C > 1 \). Hence,

\[
E_d = \lim_{d \to \infty} \frac{1}{\ln d} \sum_{m=1}^{\lfloor dx \rfloor} \frac{|dx| - m - 1}{d} \left( \frac{3}{4} m^{-\frac{3}{2}} \delta_{3/2}(m) + \delta_{3/2}^2(m) \right)
\]

\[
\leq \lim_{d \to \infty} \frac{1}{\ln d} \sum_{m=1}^{M_\epsilon - 1} \left( \frac{3}{4} m^{-\frac{3}{2}} \delta_{3/2}(m) + \delta_{3/2}^2(m) \right) + \lim_{d \to \infty} \frac{1}{\ln d} \sum_{m=M_\epsilon}^{\lfloor dx \rfloor - 1} \left( \frac{3}{4} m^{-\frac{3}{2}} \delta_{3/2}(m) + \delta_{3/2}^2(m) \right)
\]

\[
\leq \lim_{d \to \infty} \frac{1}{\ln d} \sum_{m=1}^{M_\epsilon - 1} \left( \frac{3}{4} m^{-\frac{3}{2}} \delta_{3/2}(m) + \delta_{3/2}^2(m) \right) + \lim_{d \to \infty} \frac{1}{\ln d} \sum_{m=M_\epsilon}^{\lfloor dx \rfloor - 1} \left( \frac{3}{4} \epsilon m^{-1} + \epsilon^2 m^{-1} \right)
\]

\[
\leq \lim_{d \to \infty} \frac{3CM_\epsilon}{\ln d} + \lim_{d \to \infty} \frac{2\epsilon}{\ln d} \sum_{m=M_\epsilon}^{\lfloor dx \rfloor - 1} m^{-1}.
\]

The first term in the above inequality is clearly 0. For the second term, observe that

\[
\lim_{d \to \infty} \frac{1}{\ln d} \sum_{m=1}^{\lfloor dx \rfloor - 1} m^{-1} = 1,
\]

so that \( \frac{1}{\ln d} \sum_{m=1}^{\lfloor dx \rfloor - 1} m^{-1} \) is uniformly bounded by some constant \( C_1 \) for all \( d \geq 1 \). We hence deduce that, for all \( d \geq 1 \), \( E_d \leq \epsilon C_1 \) which holds for all \( \epsilon < 1 \). This implies \( E_d = 0 \) and

\[
D_d^* = \lim_{d \to \infty} \frac{9}{64 \ln d} \sum_{m=1}^{\lfloor dx \rfloor} \frac{|dx| - m - 1}{m^{-1}}
\]

\[
= \frac{9x}{64} \lim_{d \to \infty} \frac{1}{\ln d} \sum_{m=1}^{\lfloor dx \rfloor} m^{-1} - \frac{9}{64} \lim_{d \to \infty} \frac{1}{d \ln d} \sum_{m=1}^{\lfloor dx \rfloor} (1 + m^{-1}) = \frac{9x}{64}.
\]

Now, if \( r_d = |D_d^* - D_d| \) and \( \lim_{d \to \infty} r_d = 0 \), then (17) holds. To this end we have

\[
r_d = \frac{1}{d \ln d} \sum_{j=3}^{\lfloor dx \rfloor - 2} \sum_{k=1}^{j-2} \left( 1 - \left( \frac{k}{j} \right)^{2 - \frac{3}{2}} \right) a_{3/2}^2(j-k)
\]

\[
\leq \frac{1}{d \ln d} \sum_{m=1}^{\lfloor dx \rfloor - 1} a_{3/2}^2(m) \sum_{k=1}^{\lfloor dx \rfloor - m - 1} \left( 1 - \left( \frac{m}{k + m} \right)^{2 - \frac{3}{2}} \right).
Note that as \( x \to \infty \), \( \ln |x| \leq |x|^\zeta \) for any positive value of \( \zeta \). Also, \( a_{2\zeta} (m) \leq \frac{4}{7} m^{-\frac{3}{7}} \) as \( m \to \infty \). Thus, a Taylor expansion of \((1 - x)^p\) for \( p > 0 \) and \( 0 \leq x < 1 \) yields

\[
rd \leq C \frac{d!}{d\ln d} \sum_{m=1}^{|dx|} \frac{|dx|-1}{m-1} \sum_{k=1}^{|dx|-m-1} \frac{m}{k+m} \int_a^b E[X_s X_t] dsdt.
\]

which implies \( \lim_{d \to \infty} r_d = 0 \). Finally, combining the cases \( k \leq j \) and \( j < k \) yields

\[
\lim_{d \to \infty} \rho^2 \left( \frac{C}{d} \ln d \right) \leq \frac{9x}{32}.
\]

6.3. Lemmas related to the proof of Theorem 3 in Subsection 4.1.

**Lemma 9.** The covariance structure of \( X \) can be written as a double integral as

\[
E[(X_a - X_{a-\epsilon})(X_b - X_{b-\delta})] = \int_b^a \int_{a-\epsilon}^b \partial_{s,t} E[X_s X_t] dsdt.
\]

Moreover, whenever \( \frac{a-t}{s-t} < \frac{1}{2} \), the following bound holds

\[
|\partial_{s,t} E[X_s X_t]| \leq C(s \wedge t)^{2\beta - \alpha} (s \vee t)^{\alpha - 2},
\]

and whenever \( \frac{1}{2} < \frac{a-t}{s-t} < 1 \), we have

\[
|\partial_{s,t} E[X_s X_t]| \leq C(s \wedge t)^{2\beta - \alpha} |s - t|^{\alpha - 2}.
\]

**Proof.** The first assertion can be deduced from writing

\[
E[(X_a - X_{a-\epsilon})(X_b - X_{b-\delta})] = E[X_a X_b] - E[X_{a-\epsilon} X_b] - E[X_a X_{b-\delta}] + E[X_{a-\epsilon} X_{b-\delta}]
\]

\[
= \int_b^a \int_{a-\epsilon}^b \partial_{s,t} E[X_s X_t] dsdt.
\]

The bounds on \( |\partial_{s,t} E[X_s X_t]| \) are consequences of hypotheses (H.1) and (H.2) as we will show next. Without loss of generality, let’s assume \( s \leq t \). In the case \( \frac{a-t}{s-t} < \frac{1}{2} \), (H.2) implies that

\[
\partial_{s,t} E[X_s X_t] = (2\beta - 1)s^{2\beta-2} \phi'(\frac{t}{s}) - s^{2\beta-3} \phi''(\frac{t}{s}) \leq Cs^{2\beta-2} \left( \frac{t}{s} \right)^{\alpha - 2} + Cs^{2\beta-3} \left( \frac{t}{s} \right)^{\alpha - 3} \leq Cs^{2\beta-\alpha} t^{\alpha - 2}.
\]
Meanwhile, whenever \( \frac{1}{2} < \frac{s}{t} \leq 1 \), (H.1) implies that

\[
\partial_{s,t} \mathbb{E}[X_s X_t] = (1 - 2\beta)\lambda \alpha x^{2\beta - 2} \left( \frac{t}{s} - 1 \right)^{\alpha - 1} + \lambda \alpha (\alpha - 1) s^{2\beta - 3} t \left( \frac{t}{s} - 1 \right)^{\alpha - 2} + (2\beta - 1) s^{2\beta - 2} \phi' \left( \frac{t}{s} \right) - x^{2\beta - 3} t \phi'' \left( \frac{t}{s} \right)
\]

\[
\leq C s^{2\beta - 2} \left( \frac{t}{s} - 1 \right)^{\alpha - 1} + C s^{2\beta - 3} t \left( \frac{t}{s} - 1 \right)^{\alpha - 2} + C s^{2\beta - 1} t^{\alpha - 1} + C s^{2\beta - 2} \left( \frac{t}{s} - 1 \right)^{\alpha - 1}
\]

\[
\leq C s^{2\beta - 3} t \left( \frac{t}{s} - 1 \right)^{\alpha - 2}
\]

\[
= C s^{\beta - 1} t (t - s)^{\alpha - 2}
\]

\[
\leq C (s \land t)^{\beta - \alpha} |s - t|^{\alpha - 2}.
\]

□

**Lemma 10.** Assuming the integrals appearing on the right hand sides of the equalities below are well defined, it holds that

\[
\lim_{d,p \to \infty} \mathbb{E} \left[ \hat{W}_{ij} \left( \lfloor dx \rfloor \right) \hat{W}_{ij} \left( \lfloor px \rfloor \right) \right] = \int_0^1 \int_0^1 g(u,v) f(u,v)^2 dudv
\]

as well as

\[
\lim_{p \to \infty} \mathbb{E} \left[ \hat{W}_{ij} \left( \lfloor dx \rfloor \right) \hat{W}_{ij} \left( \lfloor px \rfloor \right) \right] = d \sum_{k=0}^{d-1} \int_0^1 g \left( \frac{k}{d} v \right) \left( \int_{\frac{k}{d}}^{\frac{k+1}{d}} f(u,v) du \right)^2 dv.
\]

**Proof.** Both limits follow directly from the mean value theorem. □

**Lemma 11.** For any \( 1 \leq i,j \leq n \), the sequence \( \left\{ \hat{W}_{ij}(\lfloor dx \rfloor) \colon d \in \mathbb{N} \right\} \) is Cauchy in \( L^2(\Omega) \).

**Proof.** A sequence \( \{a_n \colon n \in \mathbb{N}\} \) in a Hilbert space \( K \) is Cauchy in \( K \) if and only if \( \langle a_n, a_m \rangle_K \to C \) as \( n, m \to \infty \), for some constant \( C \) as

\[
\|a_m - a_n\|_K^2 = \langle a_m, a_m \rangle_K + \langle a_n, a_n \rangle_K - 2 \langle a_m, a_n \rangle_K.
\]

Based on this observation, we only need to show that

\[
I = \lim_{d,p \to \infty} \mathbb{E} \left[ \hat{W}_{ij} \left( \lfloor dx \rfloor \right) \hat{W}_{ij} \left( \lfloor px \rfloor \right) \right] < \infty
\]
for any $1 \leq i, j \leq n$. Thus, we use the first part of Lemma 9 and [16, Lemma 3.1] to write

$$I = \lim_{d,p \to \infty} \sum_{1 \leq k \leq [dx]} \frac{\mathbb{E} \left[ (X_{i+1} - X_i) \left( X_{j+1} - X_j \right) \right]^2}{\mathbb{E} \left[ (X_{i+1} - X_i)^2 \right] \mathbb{E} \left[ (X_{j+1} - X_j)^2 \right]}$$

$$= \frac{1}{4\lambda^2} \lim_{d,p \to \infty} \sum_{1 \leq j \leq [dx]} \int_0^x \int_0^x (st)^{\alpha-2\beta} \left( \frac{s}{t} \right)^{\alpha-2\beta} \left( \int_0^{\frac{x}{t}} \int_0^{\frac{x}{s}} \partial_s \partial_t \mathbb{E}[X_s X_t] ds dt \right) ds dt$$

$$= \frac{1}{4\lambda^2} \int_0^x \int_0^x (st)^{\alpha-2\beta} \left( \frac{s}{t} \right)^{\alpha-2\beta} \mathbb{E}[X_s X_t] ds dt$$

$$= \frac{1}{2\lambda^2} \int_0^x \int_0^x (st)^{\alpha-2\beta} \left( \partial_s \mathbb{E}[X_s X_t] \right)^2 ds dt$$

$$= \frac{1}{2\lambda^2} \int_0^x \int_0^x \left( \partial_s \mathbb{E}[X_s X_t] \right)^2 ds dt + \frac{1}{2\lambda^2} \int_0^x \int_0^x (st)^{\alpha-2\beta} \left( \partial_s \mathbb{E}[X_s X_t] \right)^2 ds dt$$

$$= I_1 + I_2.$$ 

To handle $I_1$, we can use the second part of Lemma 9, which implies

$$I_1 \leq C \int_0^x \int_0^x \left( (s \wedge t)^{\alpha-2\beta} \left( s^{\alpha-\beta} \wedge t^{\alpha-\beta} \right)^2 \right) ds dt$$

$$= C \int_0^x \int_0^x \left( s^{2\beta-\alpha} t^{2\beta-\alpha} \right)^2 ds dt$$

$$= C \int_0^x \int_0^x s^{2\beta-\alpha} t^{2\beta-\alpha} ds dt$$

$$= C \int_0^x t^{2\alpha-3} ds dt,$$

which converges for $\alpha > \frac{3}{2}$. To deal with $I_2$, we appeal to Lemma 9 once more to get

$$I_2 = C \int_0^x \int_0^x \left( \partial_s \mathbb{E}[X_s X_t] \right)^2 ds dt$$

$$\leq C \int_0^x \int_0^x \left( s^{2\beta-\alpha} \left( t-s \right)^{\alpha-2} \right)^2 ds dt$$

$$\leq C \int_0^x \int_0^x \left( t-s \right)^{2(\alpha-2)} ds dt$$

$$= C \int_0^x t^{2\alpha-3} dt,$$

which is finite as well. \qed

**Acknowledgments.** The authors would like to thank Guangqu Zheng for helpful and enjoyable discussions. The authors are also grateful to two anonymous referees for their insightful comments and remarks.

**References**

[1] A. Barbour, *Stein’s Method for Diffusion Approximations*. Probability Theory and Related Fields, vol. 84, pp. 297-322, 1990.
[2] P. Billingsley, *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics, 1999.
[3] S. Bourguin, S. Campese, *Approximation of Hilbert-valued Gaussian measures on Dirichlet structures*. arXiv, 2019.
[4] S. Bourguin, C. Diez, C. Tudor, *Limiting behavior of large correlated Wishart matrices with chaotic entries*. Bernoulli, to appear, 2020+.
[5] M. Brennan, G. Bresler, D. Nagaraj, *Phase transitions for detecting latent geometry in random graphs*. Probability Theory and Related Fields, 2020.
[6] J. Breton, *Convergence in variation of the joint laws of multiple Wiener-Itô integrals*. Statistics & Probability Letters, vol. 76, pp. 1904-1913, 2006.
[7] S. Bubeck, J. Ding, R. Eldan, M. Rácz, *Testing for high-dimensional geometry in random graphs*. Random Structures and Algorithms, vol. 49, pp. 503-532, October 2016.
[8] S. Bubeck, S. Ganguly, *Entropic CLT and phase transition in high-dimensional Wishart matrices*. International Mathematics Research Notices, pp. 1-19, 2015.
[9] A. Bishop, P. Del Moral, A. Niclas, *An Introduction to Wishart Matrix Moments*. arXiv, Oct 2017.
[10] S. Campese, I. Nourdin, D. Nualart, *Continuous Breuer-Major theorem: tightness and non-stationarity*. Ann. Probab., vol. 48, no. 1, pp. 147-177, 2020.
[11] S. Chatterjee, *Fluctuations of eigenvalues and second order Poincaré inequalities*. Probability Theory and Related Fields, vol. 143 (1-2), pp. 1-40, 2009.
[12] D. Chételat, M. Wells, *The middle-scale asymptotics of Wishart matrices*. The Annals of Statistics, Volume 47, Number 5, 2019.
[13] J. Davidson, *Stochastic Limit Theory: An Introduction for Econometricians*. Oxford University Press. 1994.
[14] N. El-Karoui, *Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices*. The Annals of Probability, Volume 35, Number 2, 2007.
[15] X. Fang, Y. Koike, *New error bounds in multivariate normal approximations via exchangeable pairs with applications to Wishart matrices and fourth moment theorems*. arXiv, 2020.
[16] D. Harnett, D. Nualart, *Central limit theorem for functionals of a generalized self-similar Gaussian process*. Stochastic Processes and their Applications, vol. 128, Issue 2, pages 404-425, February 2018.
[17] T. Jiang, D. Li *Approximation of Rectangular Beta-Laguerre Ensembles and Large Deviations*. Journal of Theoretical Probability, vol. 28, pp. 804-847, 2015.
[18] I. Johnstone, *High Dimensional Statistical Inference and Random Matrices*. ICM2006.
[19] I. Johnstone, *On the distribution of the largest eigenvalue in principal components analysis*. The Annals of Statistics Volume 29, Number 2, 2001.
[20] V.A. Marcenko, L.A. Pastur, *Distribution of eigenvalues in certain sets of random matrices*. Mat. Sb., vol 72, 507-536, 1967.
[21] D. Mikulincer, *A CLT in Stein’s distance for generalized Wishart matrices and higher order tensors*. arXiv, 2020.
[22] I. Nourdin, D. Nualart, *The functional Breuer-Major theorem*. Probability Theory and Related Fields, 176, 2020.
[23] I. Nourdin, G. Peccati, *Normal approximations with Malliavin calculus: From Stein’s method to universality*. Cambridge University press, 2012.
[24] I. Nourdin, G. Zheng, *Asymptotic behavior of large Gaussian correlated Wishart matrices*. arXiv, April 2018.
[25] D. Nualart, *The Malliavin calculus and related topics*. Probability and Its Application. Springer, 2006.
[26] C.E. Rasmussen, C. Williams, *Gaussian processes for machine learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 2006.
[27] M.Z. Racz., J. Richey, *A smooth transition from Wishart to GOE*. J. Theoret. Probab., vol. 32, pp. 898-906, 2019.
[28] A. Tracy, H. Widom, *Level-spacing distributions and the Airy kernel*. Communications in Mathematical Physics, Volume 159, Number 1, 1994.
[29] A. Tracy, H. Widom, *On orthogonal and symplectic matrix ensembles*. Communications in Mathematical Physics, Volume 177, Number 3, 1996.
[30] J. Wishart, *The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population*. Biometrika, vol. 20A, pp. 32-52, 1928.