Electronic Journal of Statistics
Vol. 16 (2022) 3176–3211
ISSN: 1935-7524
https://doi.org/10.1214/22-EJS2021

Asymptotics of Yule's nonsense correlation for Ornstein-Uhlenbeck paths: A Wiener chaos approach*

Soukaina Douissi

National School of Applied Sciences, Marrakech, Morocco
e-mail: s.douissi@uca.ma

Khalifa Es-Sebaiy

Department of Mathematics, Faculty of Science, Kuwait University, Kuwait
e-mail: khalifa.essebaiy@ku.edu.kw

and

Frederi Viens

Department of Statistics and Probability, Michigan State University, USA
e-mail: viens@msu.edu

Abstract: In this paper, we study the distribution of the so-called "Yule's nonsense correlation statistic" on a time interval $[0, T]$ for a time horizon $T > 0$, when T is large, for a pair (X_1, X_2) of independent Ornstein-Uhlenbeck processes. This statistic is by definition equal to:

$$\rho(T) := \frac{Y_{12}(T)}{\sqrt{Y_{11}(T)}\sqrt{Y_{22}(T)}},$$

where the random variables $Y_{ij}(T)$, $i, j = 1, 2$ are defined as

$$Y_{ij}(T) := \int_0^T X_i(u)X_j(u)du - T\bar{X}_i\bar{X}_j, \qquad \bar{X}_i := \frac{1}{T}\int_0^T X_i(u)du.$$

We assume X_1 and X_2 have the same drift parameter $\theta > 0$. We also study the asymptotic law of a discrete-type version of $\rho(T)$, where $Y_{ij}(T)$ above are replaced by their Riemann-sum discretizations. In this case, conditions are provided for how the discretization (in-fill) step relates to the long horizon T. We establish identical normal asymptotics for standardized $\rho(T)$ and its discrete-data version. The asymptotic variance of $\rho(T)T^{1/2}$ is θ^{-1}. We also establish speeds of convergence in the Kolmogorov distance, which are of Berry-Esséen-type (constant*$T^{-1/2}$) except for a $\ln T$ factor. Our method is to use the properties of Wiener-chaos variables, since $\rho(T)$ and its discrete version are comprised of ratios involving three such variables in the 2nd Wiener chaos. This methodology accesses the Kolmogorov distance thanks to a relation which stems from the connection between the Malliavin calculus and Stein's method on Wiener space.

MSC2020 subject classifications: Primary 60F05; secondary 60G15.

---

arXiv: 2108.02857

*The first and third authors' research was partially supported by the US NSF award DMS 1811779.

Keywords and phrases: Yule's nonsense correlation, Ornstein-Uhlenbeck paths, Wiener chaos approach.

Received August 2021.

Contents

1 Introduction 3177
2 Preliminaries 3182
2.1 Elements of analysis on Wiener space 3182
3 Continuous observations 3184
3.1 Asymptotic distribution of $\frac{Y_{12}(T)}{\sqrt{T}}$: 3185
3.2 The denominator term 3191
4 Discrete observations 3196
4.1 Convergence in law of $\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}$ 3197
4.2 The denominator term 3202
5 Numerical results 3207
Acknowledgements 3210
References 3210

1. Introduction

In this paper, we study the normal asymptotics in law of the so-called "Yule's nonsense correlation statistic" on a time interval $[0, T]$ when the time horizon $T > 0$ tends to infinity, for two independent paths of the Ornstein-Uhlenbeck (OU) stochastic processes. This statistic is defined as:

$$\rho(T) := \frac{Y_{12}(T)}{\sqrt{Y_{11}(T)}\sqrt{Y_{22}(T)}}, \tag{1.1}$$

where the random variables $Y_{ij}(T)$, $i, j = 1, 2$ are given by

$$Y_{ij}(T) := \int_0^T X_i(u)X_j(u)du - T\bar{X}_i\,(T)\,\bar{X}_j\,(T)\,, \qquad \bar{X}_i\,(T) := \frac{1}{T}\int_0^T X_i(u)du, \tag{1.2}$$

and (X_1, X_2) is a pair of two independent OU processes with the same known drift parameter $\theta > 0$, namely X_i solves the linear SDE, for $i = 1, 2$

$$dX_i(t) = -\theta X_i(t)dt + dW^i(t), \quad t \geq 0 \tag{1.3}$$

with $X_i(0) = 0$, $i = 1, 2$, where the driving noises $(W^1(t))_{t\geq 0}$, $(W^2(t))_{t\geq 0}$ are two independent standard Brownian motions (Wiener processes). We also study the asymptotic law of a discrete-data version of $\rho(T)$, denote by $\tilde{\rho}(n)$ for n observations, where the Riemann integrals in (1.2) are replaced by Riemann-sum approximations.

It has been known since 1926 that a discrete version of the statistic ρ, which is the Pearson correlation coefficient, does not behave the same way when the data from X_1 and X_2 are i.i.d. and as when they are the discrete-time observations of a random walk. As is universally known for i.i.d. data, and also holds for shorter-memory models, $\tilde{\rho}(n)$ converges in probability to 0 under all but the most extreme circumstances (data coming from a distribution with no second moment), but G. Udny Yule showed in [12] that when the data come from a random walk, $\tilde{\rho}(n)$ does not concentrate, and has a law which seems to converge instead to a diffuse distribution on $(-1, 1)$. The exact variance and other statistical properties of this law remained unknown with mathematical precision, though a 1986 paper [5] by P.C.B. Phillips showed that the limiting law of $\tilde{\rho}(n)$ with simple symmetric random-walk data rescaled to the time interval $[0, 1]$ is the same as the law of $\rho(1)$ for two independent Wiener processes, which is indeed necessarily diffuse and fully supported on $(-1, 1)$. This advance prompted several talented prominent probabilists to look for ways of computing statistics of $\rho(1)$, if even only its variance, but this remained elusive until 2017, when Ph. Ernst and two collaborators (one posthumous) provided a closed-form expression for $Var[\rho(1)]$ in [1]. Since then, other advances on the moments of $\rho(1)$ have been made, particularly [2], and recent progress was recorded in [3] on how to compute the momens of $\tilde{\rho}(n)$ when the paths (X_1, X_2) are independent Gaussian simple-symmetric random walks. In all cases mentioned in this paragraph, the asymptotic behavior of $\tilde{\rho}(n)$ in law (scaled appropriately in time) is necessarily that of $\rho(1)$ for two independent Wiener processes.

This leaves open the question of what happens to $\tilde{\rho}(n)$ when the paths (X_1, X_2) deviate substantially from Wiener paths or random walks. Wiener (resp. random walk) paths have the property of exact (resp. approximate) self-similarily. We take up the question of using different kinds of paths, with the simplest possible example of a clear alternative to self-similar processes, namely the ubiquitous mean-reverting OU processes. The property of mean reversion is so distinct from self-similarily, that the behavior of $\rho(T)$ changes drastically from one to the other. Note that these two classes of processes are simply those satisfying (1.3) with $\theta = 0$ (Wiener process) or $\theta \neq 0$ (OU process). To illustrate the point of how distinct these processes are, let us extend the scope of this paper momentarily, to include all processes defined by (1.3) (with or without $\theta = 0$), by replacing W with a fractional Brownian motion (fBm) denoted by B^H, for some $H \in (0, 1)$. Like the Wiener process, which corresponds to $H = 1/2$, the self-similar property of B^H simply states that for any fixed real constant a, $B^H(a\cdot) = a^H B^H(\cdot)$ in law. By using this property with $a = T$ via the change of variable $u' = u/T$ in the Riemann integrals defining $\rho(T)$, we obtain immediately the equality in law

$$\begin{aligned} Y_{ij}(T) &= \int_0^1 T^H X_i(u') T^H X_j(u') T du' - T^{-1} T^{H+1} \bar{X}_i(1) T^{H+1} \bar{X}_j(1) \\ &= T^{2H+1} Y_{ij}(1) \end{aligned}$$

and therefore

$$\mathcal{L}(\rho(T)) = \mathcal{L}(\rho(1)).$$

In fact, we only used the property of self-similarity to get the above. In other words, for any pair of self-similar processes, the law of the nonsense correlation $\rho(T)$ is constant as the time horizon T increases. In stark contrast, in this paper, we show that, for a pair of OU processes, the law of $\rho(T)$ converges to the Dirac mass at 0. As mentioned, we show more: a central limit theorem for $\rho(T)\sqrt{T}$ (the mean of $\rho(T)$ is always 0), with asymptotic variance equal to θ^{-1}, and a speed of convergence of $\mathcal{L}\left(\rho(T)T^{1/2}\right)$ to $N\left(0,\theta^{-1}\right)$ in Kolmogorov metric at the rate $T^{-1/2}\ln T$.

The discrete-observation part of this paper simply replaces the Riemann integrals by the Riemann sums, for instance replacing the first integral in $Y_{ij}(T)$ by $\Delta_n\sum_{k=0}^{n-1}X_i(t_k)X_j(t_k)$ where $\Delta_n = T/n$ and $t_k = kT/n$. We denote the resulting empirical correlation by $\tilde{\rho}(n)$ rather than $\rho(T)$. It is convenient to note that $T = n\Delta_n$ can be thought of as depending on n, and we will systematically emphasize this by denoting $T = T_n$. It is assumed that the discretization step Δ_n converges to 0 while $T_n = T$ tends to ∞, which means that $n \gg T_n$ in our asymptotics. We provide a full range of speeds of convergence in central limit theorem depending on how fast Δ_n converges to 0. We find in fact that we must have $T_n\Delta_n = n\Delta_n^2 \to 0$ as $n\to\infty$, and we also note that $n\Delta_n = T_n = T \to \infty$, as well it should. Our convergence result, which immediately implies the central limit theorem $\lim_{n\to\infty}\mathcal{L}\left(\tilde{\rho}(n)T_n^{1/2}\right) = \mathcal{N}\left(0,\theta^{-1}\right)$, is

$$\begin{aligned} d_{Kol}\left(\sqrt{\theta}\sqrt{T_n}\tilde{\rho}(n),\mathcal{N}(0,1)\right) &\leqslant c(\theta)\times\ln(n\Delta_n)\max\left((n\Delta_n)^{-1/2},(n\Delta_n^2)^{\frac{1}{3}}\right)\\ &= c(\theta)\times\ln(T_n)\max\left(T_n^{-1/2},(T_n\Delta_n)^{\frac{1}{3}}\right). \end{aligned}$$

From this, we can immediately read off that a rather optimal rate of sampling of our discrete data is one for which the two terms in the max are of the same order, i.e. $T_n^{-1/2} \asymp (T_n\Delta_n)^{\frac{1}{3}}$, which is equivalent to requiring that Δ_n be of order $T_n^{-5/2}$, which in turn, since $T_n = n\Delta_n$, is equivalent to Δ_n of order $n^{-5/7}$. This is explained in more detail in the conclusion of the section on discrete data. In any case, in Kolmogorov distance, we see that the best rate of convergence of $\sqrt{T_n}\tilde{\rho}(n)$ to $N\left(0,\theta^{-1}\right)$ is of order $T_n^{-1/2}\ln T_n$, which is exactly the same rate as in the case of continuous data, and which occurs for a relatively frequence of observations of order $\Delta_n^{-1} = T_n^{5/2}$ over unit intervals. Lower frequency of observations lead to slower convergence rate in Kolmogorov distance in the scale of the time horizon T_n compared to continuous observation. Higher frequency of observation leads to the same rate $T_n^{-1/2}\ln T_n$ as with continuous observations, but this can be considered wasteful since the same rate was achieved at the optimal frequency of $\Delta_n^{-1} = T_n^{5/2}$ per unit time. It is worth stating again that these results in discrete time, pertaining to the convergence rate in the CLT for $\sqrt{T_n}\tilde{\rho}(n)$, are a second-order result compared to the CLT itself, i.e.

$$\lim_{n\to\infty}\sqrt{T_n}\tilde{\rho}(n) = \lim_{T\to\infty}\sqrt{T}\rho(T) = \mathcal{N}(0,\theta^{-1})$$

which holds in law identically in both the discrete and continuous data cases.

As mentioned, we use techniques from analysis on Wiener chaos to prove the above results. Some of these results are technical and novel, and we provide here a few points in the hopes of enlightening the methods. A key element comes from the connection discovered by I. Nourdin and G. Peccati (see [8]) between the Malliavin calculus and Stein's method. In that connection, the distance in law between a random variable X and the standard normal law can be measured to some extent by comparing the Hilbert-space norm of the Malliavin derivative DX to the value 1, which is the value one would find for the norm of the Malliavin derivative of a standard normal variable N under any reasonable coupling of X and N, i.e. under any reasonable representation of X on Wiener space. The question of how to represent X on Wiener space is typically trivial when dealing with functionals of stochastic processes based on Wiener processes, and this is certainly the case in our paper. The question of whether DX is an adeqate functional of X to make the comparison with $\mathcal{N}(0,1)$ is less trivial. The original work in [4] noted that it is sufficient for variables on Wiener chaos, and used an auxiliary random variable G_X which is slightly more involved than $\|DX\|^2$ to establish broader convergence in law beyond fixed chaos. That same random variable G_X was used in [9] to characterize laws on Wiener chaos at the level of densities, and was used specifically in [6, Theorem 2.4] to measure distances between laws in the Kolmogorov metric. We use that theorem herein, by applying it separately to all three Wiener chaos components Y_{11}, Y_{22}, Y_{12} which are used to calculate $\rho(T)$, noting as in [4] that $G_X = 2\|DX\|^2$.

The above elements are explained in the preliminary section on analysis on Wiener space below. They are used herein via standard computations of variances and differentiation and product rules for variables on Wiener chaos which are represented as double Wiener integrals with respect to the W_i's, leading to computing the asymptotic variance of the rescaled numerator $T^{-1/2}Y_{12}(T)$, namely $1/4\theta^3$, and the speed of convergence of the variances to this limit, including precise estimations of the constants in this rate of convergence as functions of θ. It turns out that the rescaled denominator $T^{-1}(Y_{11}(T)\,Y_{22}(T))^{1/2}$ does not have normal fluctuations, but rather converges to the constant $1/2\theta$. We establish this too. Adding to this that the numerator, as a second-chaos variable, has mean zero, this indicates that the entire rescaled fraction ρ should converge in law to $\mathcal{N}\left(0,\theta^{-1}\right)$. Finding a presumably sharp rate of this convergence in Kolmogorov metric is the main technical issue we tackle in this paper. Establishing this for the numerator alone is a key quantitative estimate. We use [6, Theorem 2.4] and our ability to compute the norm of the Malliavin derivative of the first double Wiener integral $\int_0^T X_i(u)X_j(u)du$ in the expression for $Y_{12}(T)$, and we find a rate of normal convergence of order $T^{-1/2}$. However, we must also handle the second term in $Y_{12}(T)$, which is the (rescaled) product $\bar{X}_1(T)\bar{X}_2(T)$ of two independent normal variables, which are both non-independent from the first part of $Y_{12}(T)$. For this, we appeal to a 1971 theorem of Michel and Pfanzagl [11] which allows us to decouple the dependence of a sum (resp. a ratio) of two variables when comparing them to a normal law in Kolmogorov distance. We specialize this theorem to the case when the second summand (resp. the denominator) is a product normal variable (resp. the root of a product normal), establishing an optimal use of it in this special case. See Proposition 3.4, Corollary 3.6, Proposition 3.7 and estimate (3.17), and estimate (3.25).

This optimal use of this decoupling technique comes at the very small cost of adding a factor of $\ln T$ to our rate of convergence. We believe this factor is optimal given our use of [11], and is determined by the weight of the tail of a product normal law, which is asymptotically the same as the tail of a chi-square variable with one degree of freedom, which in logarithmic scale, is the same as an exponential tail. The interested reader can check that any use of Holder's inequality or similar methods based on moments, cannot achieve this more efficient method, leading instead to a rate of convergence of $T^{-\alpha}$ for $\alpha < 1/2$. The reader will also observe our use of the fine structure of the second Wiener chaos as a separable Hilbert space, to deal with the tail distribution of ρ's denominator terms. This structure is documented for instance in [8, Section 2.7.4] where it is shown that every second-chaos variable can be represented as a series $\sum_k \lambda_k \chi_k^2$ where $\left(\chi_k^2\right)_k$ is a sequence of i.i.d. mean-zero chi-square variables with one degree of freedom, and $(\lambda_k)_k$ is in ℓ^2. In our case, the reader will observe that the terms in the denominator of ρ also contain non-zero expectations, that their λ_k's are positive and in ℓ^1, and that the expectations equal $\sum_k \lambda_k$. This fact is essential to us being able to control the denominator.

The techniques used to establish results in the case of discrete observations are similar to those in the continuous case. Additional ingredients include the rate of convergence of the Riemann-sum version of the first integral in $Y_{12}(T_n)$ to its limit. This rate turns out to be $n\Delta_n^2$ where, as mentioned, n is the number of observations in $(0, T_n]$, and the regular mesh is Δ_n. The use of the aforementioned Michel-Pfanzagl theorem from [11] to deal with the product-normal term in the numerator has to be optimized against this dicretization error; this is where the term $\left(n\Delta_n^2\right)^{1/3}$ comes from, whereas the term $(n\Delta_n)^{-1/2}$ is none other than the same convergence rate $T^{-1/2}$ for the numerator in Kolmogorov distance as in the continuous case. See Lemma 4.1 and Proposition 4.3. The use of the sum version of the Michel-Pfanzagl theorem from our Corollary 3.6 leads again to a leading log correction factor $\ln T_n = \ln(n\Delta_n)$. The denominator terms also require a careful analysis, though no additional ideas are needed beyond what was already established in the continuous case.

With this roadmap summary complete, the structure of this paper should appear as straightforward. We begin with a section of preliminaries presenting the tools needed from analysis on Wiener space, followed by a section covering the convergence in the continuous case, and then a section dealing with the case of discrete data. The final section provides some numerics to illustrate the convergence rates in practice, wherein we find that in discrete time, the time-scaled $\tilde{\rho}(n)$ does indeed behave in distribution largely like a normal with variance θ^{-1}, even without using the optimal observation frequency.

2. Preliminaries

2.1. Elements of analysis on Wiener space

With $(\Omega, \mathcal{F}, \mathbf{P})$ denoting the Wiener space of a standard Wiener process W, for a deterministic function $h \in L^2(\mathbf{R}_+) =: \mathcal{H}$, the Wiener integral $\int_{\mathbf{R}_+} h(s)\, dW(s)$ is also denoted by $W(h)$. The inner product $\int_{\mathbf{R}_+} f(s)\, g(s)\, ds$ will be denoted by $\langle f, g\rangle_{\mathcal{H}}$.

- **The Wiener chaos expansion**. For every $q \geq 1$, $\mathcal{H}_q$ denotes the qth Wiener chaos of W, defined as the closed linear subspace of $L^2(\Omega)$ generated by the random variables $\{H_q(W(h)), h \in \mathcal{H}, \|h\|_{\mathcal{H}} = 1\}$ where H_q is the qth Hermite polynomial. Wiener chaos of different orders are orthogonal in $L^2(\Omega)$. The so-called Wiener chaos expansion is the fact that any $X \in L^2(\Omega)$ can be written as

$$X = \mathbf{E}[X] + \sum_{q=1}^{\infty} X_q \tag{2.1}$$

for some $X_q \in \mathcal{H}_q$ for every $q \geq 1$. This is summarized in the direct-orthogonal-sum notation $L^2(\Omega) = \oplus_{q=0}^{\infty} \mathcal{H}_q$. Here $\mathcal{H}_0$ denotes the constants.

- **Relation with Hermite polynomials. Multiple Wiener integrals**. The mapping $I_q(h^{\otimes q}) := q! H_q(W(h))$ is a linear isometry between the symmetric tensor product $\mathcal{H}^{\odot q}$ (equipped with the modified norm $\|.\|_{\mathcal{H}^{\odot q}} = \sqrt{q!}\|.\|_{\mathcal{H}^{\otimes q}}$) and $\mathcal{H}_q$. Hence, for X and its Wiener chaos expansion (2.1) above, each term X_q can be interpreted as a multiple Wiener integral $I_q(f_q)$ for some $f_q \in \mathcal{H}^{\odot q}$.
- **Isometry Property-Product formula**. For any integers $1 \leqslant q \leqslant p$ and $f \in \mathcal{H}^{\odot p}$ and $g \in \mathcal{H}^{\odot q}$, we have

$$\mathbf{E}[I_p(f) I_q(g)] = \begin{cases} p! \langle f, g\rangle_{\mathcal{H}^{\otimes p}} & \text{if } p = q \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

For any integers $p, q \geq 1$ and symmetric integrands $f \in \mathcal{H}^{\odot p}$ and $g \in \mathcal{H}^{\odot q}$,

$$I_p(f) I_q(g) = \sum_{r=0}^{p \wedge q} r! C_p^r C_q^r I_{p+q-2r}(f \tilde{\otimes}_r g); \tag{2.3}$$

where $f \otimes_r g$ is the contraction of order r of f and g which is an element of $\mathcal{H}^{\otimes(p+q-2r)}$ defined by

$$(f \otimes_r g)(s_1, \ldots, s_{p-r}, t_1, \ldots, t_{q-r}) \\ = \int_{\mathbf{R}_+^{p+q-2r}} f(s_1, \ldots, s_{p-r}, u_1, \ldots, u_r) g(t_1, \ldots, t_{q-r}, u_1, \ldots, u_r)\, du_1 \cdots du_r.$$

while $(f\tilde{\otimes}_r g)$ denotes its symmetrization. More generally the symmetrization $\tilde{f}$ of a function f is defined by $\tilde{f}(x_1,...,x_p) = \frac{1}{p!}\sum_\sigma f(x_{\sigma(1)},...,x_{\sigma(p)})$ where the sum runs over all permutations σ of $\{1,...,p\}$. The special case for $p = q = 1$ in (2.3) is particularly handy, and can be written in its symmetrized form:

$$I_1(f)I_1(g) = 2^{-1}I_2\left(f\otimes g + g\otimes f\right) + \langle f,g\rangle_{\mathcal{H}}. \tag{2.4}$$

where $f\otimes g$ means the tensor product of f and g.

- **Hypercontractivity in Wiener chaos**. For $h \in \mathcal{H}^{\otimes q}$, the multiple Wiener integrals $I_q(h)$, which exhaust the set $\mathcal{H}_q$, satisfy a hypercontractivity property (equivalence in $\mathcal{H}_q$ of all L^p norms for all $p \geq 2$), which implies that for any $F \in \oplus_{l=1}^q \mathcal{H}_l$ (i.e. in a fixed sum of Wiener chaoses), we have

$$\left(\mathbf{E}\big[|F|^p\big]\right)^{1/p} \leqslant c_{p,q}\left(\mathbf{E}\big[|F|^2\big]\right)^{1/2} \quad \text{for any } p \geq 2. \tag{2.5}$$

  The constants $c_{p,q}$ above are known with some precision when $F \in \mathcal{H}_q$: by Corollary 2.8.14 in [8], $c_{p,q} = (p-1)^{q/2}$.
- **Malliavin derivative**. For any function $\Phi \in C^1(\mathbf{R})$ with bounded derivative, and any $h \in \mathcal{H}$, the Malliavin derivative D of the random variable $X := \Phi(W(h))$ is defined to be consistent with the following chain rule:

$$DX : X \mapsto D_r X := \Phi'(W(h))\,h(r) \in L^2(\Omega\times\mathbf{R}_+).$$

  A similar chain rule holds for multivariate Φ. One then extends D to the so-called Gross-Sobolev subset $\mathbf{D}^{1,2} \subsetneqq L^2(\Omega)$ by closing D inside $L^2(\Omega)$ under the norm defined by $\|X\|_{1,2}^2 := \mathbf{E}\left[X^2\right] + \int_{\mathbf{R}_+}\mathbf{E}|D_r X|^2 dr$. All Wiener chaos random variable are in the domain $\mathbf{D}^{1,2}$ of D. In fact this domain can be expressed explicitly for any X as in (2.1): $X \in \mathbf{D}^{1,2}$ if and only if $\sum_q qq!\|f_q\|^2_{\mathcal{H}^{\otimes q}} < \infty$.
- **Generator L of the Ornstein-Uhlenbeck semigroup**. The linear operator L is defined as being diagonal under the Wiener chaos expansion of $L^2(\Omega)$: $\mathcal{H}_q$ is the eigenspace of L with eigenvalue $-q$, i.e. for any $X \in \mathcal{H}_q$, $LX = -qX$. We have $Ker(L) = \mathcal{H}_0$, the constants. The operator $-L^{-1}$ is the negative pseudo-inverse of L, so that for any $X \in \mathcal{H}_q$, $-L^{-1}X = q^{-1}X$.
- **Kolmogorov distance.** Recall that, if X, Y are two real-valued random variables, then the Kolmogorov distance between the law of X and the law of Y is given by

$$d_{Kol}(X,Y) = \sup_{z\in\mathbf{R}}\left|\mathbf{P}\left[X \leqslant z\right] - \mathbf{P}\left[Y \leqslant z\right]\right|$$

  If $X \in \mathbb{D}^{1,2}$, with $\mathbf{E}[X] = 0$ and $Y = \mathcal{N}(0,1)$, then (Theorem 2.4 in [6]), then

$$d_{Kol}(X,Y) \leqslant \sqrt{\mathbf{E}[(1-\langle DX, -DL^{-1}X\rangle_{\mathcal{H}})^2]}$$

  If moreover, $X = I_q(f)$ for some $q \geq 2$, $f \in \mathcal{H}^{\odot q}$, then $\langle DX, -DL^{-1}X\rangle_{\mathcal{H}} = q^{-1}\|DX\|^2_{\mathcal{H}}$, and thus in this case

$$d_{Kol}(X,Y) \leqslant \sqrt{\mathbf{E}[(1-q^{-1}\|DX\|^2_{\mathcal{H}})^2]} \tag{2.6}$$

Lemma 2.1. *Let $\gamma > 0$. Let $(Z_n)_{n\in\mathbb{N}}$ be a sequence of random variables. If for every $p \geq 1$ there exists a constant $c_p > 0$ such that for all $n \in \mathbb{N}$,*

$$\|Z_n\|_{L^p(\Omega)} \leqslant c_p \cdot n^{-\gamma},$$

then for all $\varepsilon > 0$ there exists a random variable η_ε which is almost surely finite such that

$$|Z_n| \leqslant \eta_\varepsilon \cdot n^{-\gamma+\varepsilon} \quad \textit{almost surely}$$

for all $n \in \mathbb{N}$. Moreover, $E|\eta_\varepsilon|^p < \infty$ for all $p \geq 1$.

3. Continuous observations

In this section, we compute the asymptotic variance of $\rho(T)$ and its normal fluctuations for large T, by working with each of the three terms which appear in its definition. For the sake of convienence and compactness of notation, we construct a two-sided Brownian motion $(W(t))_{t\in\mathbb{R}}$ from the two independent Brownian motions $(W^1(t))_{t\geq 0}$ and $(W^2(t))_{t\geq 0}$ as follows:

$$W(t) := W^1(t)\mathbf{1}_{\{t\geq 0\}} + W^2(-t)\mathbf{1}_{\{t<0\}}, \quad t \in \mathbb{R}.$$

The following lemma will be convenient in the sequel.

Lemma 3.1. *Let f, $g \in L^2(\mathbb{R}_+)$, then*

$$I_1^{W^1}(f)I_1^{W^2}(g) = I_2^W(\bar{f} \otimes \bar{\bar{g}})$$

where $\bar{f}$, $\bar{\bar{g}}$ in $L^2(\mathbb{R})$ are defined by

$$\bar{f}(x) = f(x)\mathbf{1}_{\{x\geq 0\}}, \quad \bar{\bar{g}}(x) = -g(-x)\mathbf{1}_{\{x<0\}}.$$

Proof. Using the product formula of multiple integrals, we have

$$\begin{aligned}
I_2^W(\bar{f}\otimes\bar{\bar{g}}) &= I_1^W(\bar{f})I_1^W(\bar{\bar{g}}) - E\left[I_1^W(\bar{f})I_1^W(\bar{\bar{g}})\right] \\
&= \left(\int_{\mathbb{R}} \bar{f}(x)dW_x\right)\left(\int_{\mathbb{R}} \bar{\bar{g}}(x)dW_x\right) - E\left\langle \bar{f}, \bar{\bar{g}}\right\rangle_{L^2(\mathbb{R})} \\
&= \left(\int_0^\infty f(x)dW_x^1\right)\left(\int_0^\infty g(x)dW_x^2\right) \\
&= I_1^{W^1}(f)I_1^{W^2}(g),
\end{aligned}$$

which completes the proof. □

3.1. Asymptotic distribution of $\frac{Y_{12}(T)}{\sqrt{T}}$:

The numerator of $\rho(T)$ can be written as follows

$$\frac{Y_{12}(T)}{\sqrt{T}} = F_T - \sqrt{T}\bar{X}_1(T)\bar{X}_2(T) \tag{3.1}$$

where $F_T := \frac{1}{\sqrt{T}}\int_0^T X_1(t)X_2(t)dt$. Using the notation I_1^W for the Wiener integral with respect to W, since $X_i(t) = \int_0^t e^{-\theta(t-u)}dW^i(u) = I_1^{W^i}(f_t)$, $i = 1,2$ where $f_t(.) := e^{-\theta(t-.)}\mathbf{1}_{[0,t]}(.)$ we can write using Lemma 3.1

$$\begin{aligned} F_T &= \frac{1}{\sqrt{T}}\int_0^T I_1^{W^1}(f_t)I_1^{W^2}(f_t)dt \\ &= I_2^W(h_T), \end{aligned} \tag{3.2}$$

with $h_T \in L^2([-T,T]^2)$ is given

$$\begin{aligned} h_T : [-T,T]^2 &\to \mathbb{R} \\ (x,y) &\mapsto \frac{1}{\sqrt{T}}\int_0^T \bar{f}_t(x)\bar{\bar{f}}_t(y)dt \end{aligned} \tag{3.3}$$

On the other hand, we have

$$\begin{aligned} h_T(x,y) &= \frac{1}{\sqrt{T}}\int_0^T -e^{-2\theta t}e^{\theta x}e^{-\theta y}\mathbf{1}_{[0,t]}(x)\mathbf{1}_{[-t,0]}(y)dt \\ &= \frac{1}{\sqrt{T}}\int_0^T -e^{-2\theta t}e^{\theta x}e^{-\theta y}\mathbf{1}_{[x\vee -y,T]}(t)\mathbf{1}_{[0,T]}(x)\mathbf{1}_{[-T,0]}(y)dt \\ &= \frac{1}{2\theta}\frac{1}{\sqrt{T}}e^{\theta x}e^{-\theta y}\left[e^{-2\theta T} - e^{-2\theta(x\vee -y)}\right]\mathbf{1}_{[0,T]}(x)\mathbf{1}_{[-T,0]}(y). \end{aligned}$$

Note that the kernel h_T is not symmetric, in the sequel we will denote $\tilde{h}_T$ its systematization defined by $\tilde{h}_T(x,y) := \frac{1}{2}(h_T(x,y)+h_T(y,x))$. We are now ready to compute the asymptotic variance of the main term in the numerator of ρ.

Lemma 3.2. *With F_T defined in (3.2), then*

$$\left|\mathbf{E}[F_T^2] - \frac{1}{4\theta^3}\right| \leqslant \frac{C(\theta)}{T},$$

where $C(\theta) := \frac{9}{16\theta^4}$. In particular, $\lim_{T\to\infty}\mathbf{E}[F_T^2] = \frac{1}{4\theta^3}$.

Proof. We have

$$\begin{aligned} \mathbf{E}[F_T^2] &= \mathbf{E}[I_2^W(h_T)^2] \\ &= 2 \times \|\tilde{h}_T\|^2_{L^2([-T,T]^2)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{T}\frac{1}{4\theta^2}\int_{-T}^{0}\int_{0}^{T} e^{2\theta x}e^{-2\theta y}\left[e^{-2\theta T}-e^{-2\theta(x\vee -y)}\right]^2 dxdy \\
&= \frac{1}{T}\frac{1}{4\theta^2}\int_{0}^{T}\int_{0}^{T} e^{2\theta x}e^{2\theta z}\left[e^{-2\theta T}-e^{-2\theta(x\vee z)}\right]^2 dxdz \\
&= \frac{1}{T}\frac{1}{2\theta^2}\int_{0}^{T}\int_{0}^{z} e^{2\theta x}e^{2\theta z}\left[e^{-2\theta T}-e^{-2\theta z}\right]^2 dxdz \\
&= \frac{1}{T}\frac{1}{4\theta^3}\left[\int_{0}^{T}(e^{-2\theta(T-y)}-1)^2dy-\int_{0}^{T} e^{-2\theta y}(e^{-2\theta(T-y)}-1)^2dy\right] \\
&=: A_1(T)+A_2\,(T)
\end{aligned}
$$

where

$$
\begin{aligned}
|A_1(T)-\frac{1}{4\theta^3}| &:= \left|\frac{1}{T}\frac{1}{4\theta^3}\int_{0}^{T}(e^{-2\theta(T-y)}-1)^2dy-\frac{1}{4\theta^3}\right| \\
&= \left|\frac{1}{T}\frac{1}{4\theta^3}[\frac{1}{4\theta}(1-e^{-4\theta T})+\frac{1}{\theta}(e^{-2\theta T}-1)+T]-\frac{1}{4\theta^3}\right| \\
&\leqslant \frac{5}{16\theta^4}\times\frac{1}{T},
\end{aligned}
$$

and

$$
\begin{aligned}
|A_2(T)| &:= \left|\frac{1}{T}\frac{1}{4\theta^3}\int_{0}^{T} e^{-2\theta y}(e^{-2\theta(T-y)}-1)^2dy\right| \\
&\leqslant \frac{1}{T}\frac{1}{8\theta^4}(1-e^{-4\theta T})+\frac{1}{2\theta^3}e^{-2\theta T} \\
&\leqslant \frac{1}{4\theta^4}\frac{1}{T}.
\end{aligned}
$$

□

Proposition 3.3. *Let $F_T^\theta := 2\theta^{3/2}F_T$ and $N\sim\mathcal{N}(0,1)$, then we have*

$$
d_{Kol}(F_T^\theta, N)\leqslant\frac{c(\theta)}{\sqrt{T}}.
$$

where $c(\theta) := \sqrt{(\frac{9}{4\theta})^2+\frac{3^3}{4\theta}}$. Consequently $F_T \xrightarrow{\mathcal{L}} \mathcal{N}\left(0,\frac{1}{4\theta^3}\right)$ as $T\to+\infty$.

Proof. We will use the estimate (2.6) recalled in the preliminaries in order to prove this proposition. We have $D_tF_T^\theta = 4\theta^{3/2}I_1^W(\tilde{h}_T(.,t))$, $t\in[-T,T]$, hence

$$
\begin{aligned}
\frac{1}{2}\|DF_T^\theta\|^2_{L^2([-T,T])} &= \frac{1}{2}\int_{-T}^{T}(D_tF_T^\theta)^2dt \\
&= 8\theta^3\int_{-T}^{T} I_1^W(\tilde{h}_T(.,t))^2dt \\
&= 8\theta^3\left[\int_{-T}^{T} I_2^W(\tilde{h}_T(.,t)\otimes\tilde{h}_T(.,t))dt+\int_{-T}^{T}\|\tilde{h}_T(.,t)\|^2_{L^2([-T,T])}\right]
\end{aligned}
$$

$$= 8\theta^3 \left[I_2^W(\tilde{h}_T \otimes_1 \tilde{h}_T) + \|\tilde{h}_T\|^2_{L^2([-T,T]^2)} \right],$$

where we used the product formula (2.4) and the fact that the kernel $\tilde{h}_T \otimes_1 \tilde{h}_T$ is symmetric. Thus

$$\mathbf{E}\left[(1 - \frac{1}{2}\|DF_T^\theta\|^2_{L^2([-T,T])})^2 \right] = (\mathbf{E}[(F_T^\theta)^2] - 1)^2 + 2^7\theta^6 \times \|\tilde{h}_T \otimes_1 \tilde{h}_T\|^2_{L^2([-T,T]^2)} \tag{3.4}$$

We have,

$$\begin{aligned}
&(\tilde{h}_T \otimes_1 \tilde{h}_T)(x,y) \\
&= \int_{[-T,T]} \tilde{h}_T(x,z)\tilde{h}_T(y,z)dz \\
&= \frac{1}{4}\int_{[-T,0]} h_T(x,z)h_T(y,z)\mathbf{1}_{[0,T]}(x)\mathbf{1}_{[0,T]}(y)dz \\
&\quad + \frac{1}{4}\int_{[0,T]} h_T(z,x)h_T(z,y)\mathbf{1}_{[-T,0]}(x)\mathbf{1}_{[-T,0]}(y)dz
\end{aligned}$$

Hence

$$\begin{aligned}
&\|\tilde{h}_T \otimes_1 \tilde{h}_T\|^2_{L^2([-T,T]^2)} \\
&\leqslant \frac{1}{8}\int_{[0,T]^2} \left(\int_{[-T,0]} h_T(x,z)h_T(y,z) \right)^2 dxdydz \\
&\quad + \frac{1}{8}\int_{[-T,0]^2} \left(\int_{[0,T]} h_T(z,x)h_T(z,y) \right)^2 dxdydz
\end{aligned} \tag{3.5}$$

On the other hand by Fubini's theorem

$$\begin{aligned}
&\int_{[0,T]^2} \left(\int_{[-T,0]} h_T(x,z)h_T(y,z)dz \right)^2 dxdy \\
&= \frac{1}{T^2}\int_{[0,T]^2} \left[\int_{[-T,0]}\int_{[0,T]^2} \bar{f}_r(x)\bar{\bar{f}}_r(z)\bar{f}_s(y)\bar{\bar{f}}_s(z)drdsdz \right]^2 dxdy \\
&= \frac{1}{T^2}\int_{[0,T]^2} \left(\int_{[0,T]^2} \bar{f}_r(x)\bar{f}_s(y)\langle \bar{\bar{f}}_r, \bar{\bar{f}}_s\rangle_{L^2([-T,0])}drds \right)^2 dxdy \\
&= \frac{1}{T^2}\int_{[0,T]^2}\int_{[0,T]^4} \\
&\quad \bar{f}_r(x)\bar{f}_s(y)\bar{f}_v(x)\bar{f}_u(y)\langle \bar{\bar{f}}_r, \bar{\bar{f}}_s\rangle_{L^2([-T,0])}\langle \bar{\bar{f}}_v, \bar{\bar{f}}_u\rangle_{L^2([-T,0])}drdsdudvdxdy
\end{aligned}$$

Using the fact that the other term of (3.5) can be treated similarly and that $\langle \bar{\bar{f}}_r, \bar{\bar{f}}_s\rangle_{L^2([-T,0])} = \langle \bar{f}_r, \bar{f}_s\rangle_{L^2([0,T])} = \mathbf{E}[X^i(r)X^i(s)]$, $i = 1, 2$, we get

$$\|\tilde{h}_T \otimes_1 \tilde{h}_T\|^2_{L^2([-T,T]^2)}$$

$$\leqslant \frac{1}{4}\frac{1}{T^2}\int_{[0,T]^4} \mathbf{E}[X^i(r)X^i(v)]\mathbf{E}[X^i(s)X^i(u)]\mathbf{E}[X^i(r)X^i(s)]\mathbf{E}[X^i(v)X^i(u)]drdsdudv$$

On the other hand, since for $i = 1,2$, $\mathbf{E}[X^i(r)X^i(s)] = \frac{e^{-\theta(r+s)}}{2\theta}[e^{2\theta(r\wedge s)} - 1] \leqslant \frac{1}{2\theta}e^{-\theta|r-s|} = \mathbf{E}[Z_i(r)Z_i(s)] := Q(r-s)$, where $Z_i(r) := \int_{-\infty}^{r} e^{-\theta(r-t)}dW^i(t)$, $i = 1,2$ we get

$$\begin{aligned}
&\|\tilde{h}_T \otimes_1 \tilde{h}_T\|^2_{L^2([-T,T]^2)} && (3.6)\\
&\leqslant \frac{1}{4}\frac{1}{T^2}\int_{[0,T]^4} Q(u-v)Q(v-r)Q(r-s)Q(s-u)dudvdrds && (3.7)\\
&= \frac{1}{4}\frac{1}{T^2}\int_{[0,T]^2} dudr \int_{\mathbb{R}^2} dvdsQ_T(u-v)Q_T(s-r)Q_T(u-s)Q_T(v-r) && (3.8)\\
&= \frac{1}{4}\frac{1}{T^2}\int_{[0,T]^2} dudr \int_{\mathbb{R}^2} dvdsQ_T(y)Q_T(u-r-x)Q_T(x)Q_T(u-r-y), && (3.9)
\end{aligned}$$

where $Q_T(x) := |Q(x)|\mathbf{1}_{\{|x|\leqslant T\}}$ and we used the change of variables $y = u - v$, $x = u - s$. Therefore, applying Young's inequality, we can conclude

$$\begin{aligned}
\|\tilde{h}_T \otimes_1 \tilde{h}_T\|^2_{L^2([-T,T]^2)} &\leqslant \frac{1}{4}\frac{1}{T^2}\int_{[0,T]^2} dudr(Q_T * Q_T)(u-r)^2\\
&\leqslant \frac{1}{4}\frac{1}{T}\int_{\mathbb{R}}(Q_T * Q_T)(z)^2dz\\
&= \frac{1}{4}\frac{1}{T}\|Q_T * Q_T\|^2_{L^2(\mathbb{R})}\\
&\leqslant \frac{1}{4}\frac{1}{T}\|Q_T\|^4_{L^{4/3}(\mathbb{R})} = \frac{1}{4}\frac{1}{T}\left(\int_{[-T,T]}|Q(t)|^{4/3}dt\right)^{3/4}\\
&= \frac{1}{4}\frac{1}{T}\left(\frac{3^3}{2^7\theta^7}(1-e^{-4\theta T/3})^3\right). \qquad (3.10)
\end{aligned}$$

The desired result follows using (2.6) and the estimates (3.4), (3.10) and Lemma 3.2. □

We will need the following Proposition due to Michel and Pfanzagl (1971) [11] in the sequel which gives upper bounds for Kolmogorov's distance between respectively the sum and the ratio of two random variables and a standard Gaussian random variable.

Proposition 3.4. *Let X, Y and Z be three random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that $\mathbf{P}(Z > 0) = 1$. Then, for all $\varepsilon > 0$, we have*

1. $d_{Kol}(X+Y, N) \leqslant d_{Kol}(X,N) + \mathbf{P}(|Y| > \varepsilon) + \varepsilon$.
2. $d_{Kol}(\frac{X}{Z}, N) \leqslant d_{Kol}(X,N) + \mathbf{P}(|Z-1| > \varepsilon) + \varepsilon$.

where $N \sim \mathcal{N}(0,1)$.

Proposition 3.5. *Let Y be a r.v. such that $Y = N \times N'$ where $N \sim \mathcal{N}(0, \sigma_1^2)$ and $N' \sim \mathcal{N}(0, \sigma_2^2)$ two independent Gaussian r.v defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Then, there exists a constant $\beta > \frac{2\sqrt{3}}{3}\sigma_1\sigma_2$ such that*

$$\mathbf{E}\left[e^{\frac{Y}{\beta}}\right] < 2.$$

Moreover, there exists a constant $C > \frac{\sqrt{3}}{3}\pi$ such that $\beta < C \times \mathbf{E}[|Y|]$.

Proof. By the independence of N and N', it's easy to check that for any $\beta > \sigma_1\sigma_2$, we have

$$\begin{aligned}\mathbf{E}\left[e^{\frac{Y}{\beta}}\right] &= \frac{1}{2\pi\sigma_1\sigma_2}\int_{\mathbb{R}^2} e^{\frac{xy}{\beta}} e^{-\frac{x^2}{2\sigma_1^2}} e^{-\frac{y^2}{2\sigma_2^2}} dxdy \\ &= \frac{1}{\sqrt{(1 - \frac{\sigma_1^2\sigma_2^2}{\beta^2})}}.\end{aligned}$$

Thus the constraint $\mathbf{E}\left[e^{\frac{Y}{\beta}}\right] < 2$ implies that β should be such that $\beta > \frac{2\sqrt{3}}{3}\sigma_1\sigma_2$. On the other hand since $\mathbf{E}[|Y|] = \mathbf{E}[|N|] \times \mathbf{E}[|N'|] = \frac{2\sigma_1\sigma_2}{\pi}$, thus there exists a constant $C > \frac{\sqrt{3}}{3}\pi$ such that $\beta < C \times \mathbf{E}[|Y|]$. □

Corollary 3.6. *Let X, Y be two r.v. defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that $Y = N \times N'$ where $N \sim \mathcal{N}(0, \sigma_1^2)$ and $N' \sim \mathcal{N}(0, \sigma_2^2)$. Then, there exists a constant $\frac{2\sqrt{3}}{3}\sigma_1\sigma_2 < \beta \leqslant ((C \times \mathbf{E}[|Y|]) \wedge 4)$ such that*

$$d_{Kol}(X + Y, \mathcal{N}(0,1)) \leqslant d_{Kol}(X, \mathcal{N}(0,1)) + \beta\left(1 + \ln\left(\frac{4}{\beta}\right)\right).$$

Proof. Let X, Y be two random variables, then from Michel and Pfanzagl (1971), for all $\varepsilon > 0$,

$$d_{Kol}(X + Y, \mathcal{N}(0,1)) \leqslant d_{Kol}(X, \mathcal{N}(0,1)) + \mathbf{P}(|Y| > \varepsilon) + \varepsilon.$$

Since $Y = N \times N'$ where $N \sim \mathcal{N}(0, \sigma_1^2)$ and $N' \sim \mathcal{N}(0, \sigma_2^2)$ then, by Proposition 3.5 and Markov's inequality, we have

$$\mathbf{P}(|Y| > \varepsilon) = 2 \times \mathbf{P}(Y > \varepsilon) \leqslant 2 \times \mathbf{E}\left[e^{\frac{Y}{\beta}}\right] e^{-\frac{\varepsilon}{\beta}} < 4e^{-\frac{\varepsilon}{\beta}}.$$

Thus we can write

$$d_{Kol}(X + Y, \mathcal{N}(0,1)) \leqslant d_{Kol}(X, \mathcal{N}(0,1)) + \inf_{\varepsilon > 0} g_\beta(\varepsilon).$$

where $g_\beta(\varepsilon) := 4e^{-\frac{\varepsilon}{\beta}} + \varepsilon$. Since g_β is convex on $\mathbb{R}_+$, $\arg\inf_{\varepsilon>0} g_\beta(\varepsilon) = \varepsilon^*(\beta) = \beta\ln(\frac{4}{\beta})$, $\beta < ((C \times \mathbf{E}[|Y|]) \wedge 4)$, with C the constant from Proposition 3.5. The desired result follows. □

To prove the convergence in law of $\frac{Y_{12}(T)}{\sqrt{T}}$, recall that $T^{-1/2}Y_{12}(T) = F_T - \sqrt{T}\bar{X}_1(T)\bar{X}_2(T)$. We can write $\bar{X}_i(T) := I_1^{W_i}(g_T)$, $i = 1, 2$ where $g_T := T^{-1}\int_0^T f_t dt$ and we have for $i = 1, 2$

$$\begin{aligned}
\mathbf{E}[\bar{X}_i^2(T)] &= \|g_T\|^2_{L^2([0,T])} \\
&= \frac{1}{T^2}\int_0^T (\int_0^T f_t(u)dt)^2 du \\
&= \frac{1}{T^2}\int_0^T e^{2\theta u}(\int_u^T e^{-\theta t}dt)^2 du \\
&= \frac{1}{T^2}\frac{1}{\theta^2}\int_0^T (1 - e^{-\theta(T-u)})^2 du \leqslant \frac{1}{\theta^2}\frac{1}{T}.
\end{aligned} \tag{3.11}$$

Hence by the independence of X_1 and X_2 and denoting $Y(T) := \sqrt{T}\bar{X}_1(T)\bar{X}_2(T)$, we get

$$\mathbf{E}[Y(T)^2] \leqslant \frac{1}{T}\frac{1}{\theta^4}. \tag{3.12}$$

Then in virtue of Proposition 3.5 and Corollary 3.6, there exists a constant β with $\frac{4\sqrt{3}}{3}\sqrt{T}\theta^{3/2}\|g_T\|^2_{L^2([0,T])} < \beta < \frac{4C}{\pi\sqrt{T\theta}} \wedge 4$ such that

$$d_{Kol}\left(\frac{Y_{12}(T)}{\sqrt{T}}, \mathcal{N}(0, \frac{1}{4\theta^3})\right) \leqslant d_{Kol}\left(F_T, \mathcal{N}(0, \frac{1}{4\theta^3})\right) + \beta\left(1 + \ln\left(\frac{4}{\beta}\right)\right).$$

On the other hand, since the function $x \mapsto x(1 + \ln(\frac{4}{x}))$ is increasing on $(0, 4)$, we have for T large enough

$$d_{Kol}\left(\frac{Y_{12}(T)}{\sqrt{T}}, \mathcal{N}(0, \frac{1}{4\theta^3})\right) \leqslant d_{Kol}\left(F_T, \mathcal{N}(0, \frac{1}{4\theta^3})\right) + \frac{c(\theta)}{2}\frac{\ln(T)}{\sqrt{T}}.$$

The following proposition follows.

Proposition 3.7. *There exists a constant $C(\theta)$ depending only on θ, such that*

$$d_{Kol}\left(\frac{Y_{12}(T)}{\sqrt{T}}, \mathcal{N}\left(0, \frac{1}{4\theta^3}\right)\right) \leqslant C(\theta) \times \frac{\ln(T)}{\sqrt{T}}.$$

In particular, $\frac{Y_{12}(T)}{\sqrt{T}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \frac{1}{4\theta^3})$ *as* $T \to +\infty$.

Having just completed the study of the convergence in law of the numerator in $\rho(T)$, in order to study the convergence in law of $\sqrt{T}\rho(T)$, we will use Proposition 3.4 assertion 2 and the fact that

$$\sqrt{\theta}\sqrt{T}\rho(T) = \frac{2\theta^{3/2}\frac{Y_{12}(T)}{\sqrt{T}}}{2\theta\sqrt{\frac{Y_{11}(T)}{T} \times \frac{Y_{22}(T)}{T}}} \tag{3.13}$$

to show in the next subsection that the denominator concentrates to the value 1, and that the behavior of $\sqrt{T}\rho(T)$ is thus given by that of the numerator above.

3.2. The denominator term

Let us denote the denominator term

$$D(T) := D := 2\theta\sqrt{\frac{Y_{11}(T)\times Y_{22}(T)}{T\times T}}, \tag{3.14}$$

According to Proposition 3.4 assertion 2. we need to estimate $\mathbf{P}(|D-1|>\varepsilon)$ for instance for $0<\varepsilon<1$. Using the fact that $D\geq 0$ a.s. then $|D-1|\leqslant|D^2-1|$ a.s. Now using the shorthand notation $\bar{Y}_{ii}(T):=\frac{Y_{ii}(T)}{T/2\theta}$, $i=1,2$, thus, we have a.s.

$$\begin{aligned}|D-1|\leqslant|D^2-1|&\leqslant|\bar{Y}_{11}(T)\bar{Y}_{22}(T)-1|\\ &\leqslant|\bar{Y}_{11}(T)-1|\times|\bar{Y}_{22}(T)-1|+|\bar{Y}_{11}(T)-1|+|\bar{Y}_{22}(T)-1|\end{aligned} \tag{3.15}$$

Thus, using the fact that $\bar{Y}_{11}(T)$ and $\bar{Y}_{22}(T)$ are equal in law, we get for any $\varepsilon<1$,

$$\begin{aligned}\mathbf{P}\left(|D-1|>\varepsilon\right)&\leqslant\mathbf{P}\left(|\bar{Y}_{11}(T)-1|>\frac{\varepsilon}{3}\right)+\mathbf{P}\left(|\bar{Y}_{22}(T)-1|>\frac{\varepsilon}{3}\right)\\ &+2\mathbf{P}\left(|\bar{Y}_{11}(T)-1|^2>\frac{\varepsilon}{3}\right)\leqslant 4\times\mathbf{P}\left(|\bar{Y}_{11}(T)-1|>\frac{\varepsilon}{3}\right).\end{aligned} \tag{3.16}$$

We lighten the notation by writing ε instead of $\varepsilon/3$, therefore by Proposition 3.4 assertion 2. applied to $\rho(T)$ in (3.13), we get

$$d_{Kol}\left(\sqrt{\theta}\sqrt{T}\rho(T),N\right)\leqslant d_{Kol}\left(2\theta^{3/2}\frac{Y_{12}(T)}{\sqrt{T}},N\right)+4\times\mathbf{P}\left(|\bar{Y}_{11}(T)-1|>\varepsilon\right)+3\varepsilon. \tag{3.17}$$

where $N\sim\mathcal{N}(0,1)$.

The next step is to control the term $\mathbf{P}\left(|\bar{Y}_{11}(T)-1|>\varepsilon\right)$. We have:

$$\begin{aligned}\bar{Y}_{11}(T)&=\frac{2\theta}{T}\int_0^T\left(X_1^2(u)-\mathbf{E}[X_1^2(u)]\right)du+\frac{2\theta}{T}\int_0^T\mathbf{E}[X_1^2(u)]du-2\theta\bar{X}_1^2(T)\\ &=\frac{2\theta}{T}\int_0^T\left((I_1^{W_1}(f_u))^2-\|f_u\|^2_{L^2([0,T])}\right)du+\frac{1}{T}\int_0^T\mathbf{E}[X_1^2(u)]du-\bar{X}_1^2(T)\\ &=2\theta I_2^{W_1}(k_T)+\frac{2\theta}{T}\int_0^T\mathbf{E}[X_1^2(u)]du-2\theta\bar{X}_1^2(T)\\ &:=A_\theta(T)+\mu_\theta(T)-2\theta\bar{X}_1^2(T).\end{aligned}$$

where

$$\begin{aligned}k_T(x,y)&:=\frac{1}{T}\int_0^T f_u^{\otimes 2}(x,y)du\\ &=\frac{1}{T}\int_0^T e^{-\theta(u-x)}e^{-\theta(u-y)}\mathbf{1}_{[0,u]}(x)\mathbf{1}_{[0,u]}(y)du\end{aligned}$$

$$= \frac{1}{T}\frac{1}{2\theta}e^{\theta x}e^{\theta y}\left(e^{-2\theta(x\vee y)} - e^{-2\theta T}\right)\mathbf{1}_{[0,T]}(x)\mathbf{1}_{[0,T]}(y)$$

and

$$\begin{aligned}
\mu_\theta(T) &= \frac{2\theta}{T}\int_0^T \mathbf{E}[X_1^2(u)]du = \frac{2\theta}{T}\int_0^T \|f_u\|^2_{L^2([0,T]}du \\
&= \frac{2\theta}{T}\int_0^T\int_0^T f_u^2(t)dtdu \\
&= \frac{2\theta}{T}\int_0^T\int_0^u e^{-2\theta(u-t)}dtdu \\
&= \frac{1}{T}\int_0^T (1-e^{-2\theta u})du \\
&= 1 - \frac{1}{2\theta T}\left(1-e^{-2\theta T}\right)
\end{aligned}$$

Then we immediately get the mean concentration around 1:

$$|\mu_\theta(T) - 1| \leqslant \frac{1}{2\theta T} = O(\frac{1}{T}). \tag{3.18}$$

For the term $A_\theta(T)$, in a similar way to the calculus in the proof of Proposition 3.2 and since k_T is symmetric, we get

$$\begin{aligned}
\mathbf{E}\left[I_2^{W_1}(k_T)^2\right] &= 2\|k_T\|^2_{L^2([0,T]^2)} \\
&= \frac{1}{T^2}\frac{1}{2\theta^2}\int_{[0,T]^2} e^{2\theta x}e^{2\theta y}\left(e^{2\theta(x\vee y)} - e^{-2\theta T}\right)^2 dxdy \\
&= \frac{1}{T^2}\frac{1}{2\theta^3}\left(\frac{1}{4\theta}(1-e^{-4\theta T}) + \frac{1}{\theta}(e^{-2\theta T}-1) + T(1+2e^{-2\theta T})\right. \\
&\quad \left. -\frac{1}{2\theta}(1-e^{-4\theta T})\right) \leqslant \frac{1}{2\theta^3}\left(3+\frac{7}{4\theta}\right)\frac{1}{T}.
\end{aligned}$$

Therefore, we have $\mathrm{Var}(A_\theta(T)) = O(\frac{1}{T})$, since

$$\begin{aligned}
\mathrm{Var}(A_\theta(T)) &= \mathrm{Var}(2\theta I_2^{W_1}(k_T)^2) \\
&= 4\theta^2\mathbf{E}\left[I_2^{W_1}(k_T)^2\right] \leqslant \frac{2}{\theta}\left(3+\frac{7}{4\theta}\right)\frac{1}{T}.
\end{aligned}$$

Finally, by equation (3.11), $2\theta\mathbf{E}[\bar{X}_1^2(T)] = O(\frac{1}{T})$. On the other hand, we can write

$$\bar{Y}_{11}(T) = \tilde{Y}_{11}(T) + y_{11}(T)$$

where

$$\begin{cases}
\tilde{Y}_{11}(T) := A_\theta(T) - 2\theta\left(\bar{X}_1^2(T) - \mathbf{E}\left[\bar{X}_1^2(T)\right]\right), \\
y_{11}(T) := \mu_\theta(T) - 2\theta\mathbf{E}\left[\bar{X}_1^2(T)\right].
\end{cases}$$

By the product formula (2.3), the r.v. $\tilde{Y}_{11}(T)$ belongs to the second Wiener chaos while $y_{11}(T)$ is deterministic. Moreover, we have

$$\begin{aligned}\mathrm{Var}(\tilde{Y}_{11}(T)) &= \mathrm{Var}\left(A_\theta(T) - 2\theta \bar{X}_1^2(T)\right) \\ &\leqslant 2\mathrm{Var}\left(A_\theta(T)\right) + 8\theta^2 \mathrm{Var}\left(\bar{X}_1^2(T)\right) \\ &\leqslant \frac{4}{T\theta}\left(3 + \frac{7}{4\theta}\right) + \frac{8^2}{\theta^2 T^2} \leqslant \frac{cst(\theta)}{T} := \left[\frac{4}{\theta}\left(3 + \frac{7}{4\theta}\right) + \frac{8^2}{\theta^2}\right]\frac{1}{T},\end{aligned} \tag{3.19}$$

where we used the hypercontractivity property (2.5) on Wiener chaos for $\bar{X}_1^2(T)$, since $\mathrm{Var}(\bar{X}_1^2(T)) = \mathbf{E}[I_1^{W_1}(g_T)^4] - \left(\mathbf{E}[I_1^{W_1}(g_T)^2]\right)^2$ and $\mathbf{E}[I_1^{W_1}(g_T)^2] = \mathbf{E}[\bar{X}_1^2(T)] \leqslant \frac{1}{T\theta^2}$. The last estimate plus the estimate (3.18) on $\mu_\theta(T)$ established earlier imply that

$$|1 - y_{11}(T)| \leqslant \frac{5}{2}\frac{1}{\theta T}. \tag{3.20}$$

With all those estimates in place, we return to our main target to control $\mathbf{P}\left(|\bar{Y}_{11}(T) - 1| > \varepsilon\right)$ for some $0 < \varepsilon < 1$, Or

$$\begin{aligned}\mathbf{P}\left(|\bar{Y}_{11}(T) - 1| > \varepsilon\right) &= \mathbf{P}\left(|\tilde{Y}_{11}(T) + y_{11}(T) - 1| > \varepsilon\right) \\ &= \mathbf{P}\left(\left\{\tilde{Y}_{11}(T) > \varepsilon + 1 - y_{11}(T)\right\}\right. \\ &\quad \left.\cup \left\{-\tilde{Y}_{11}(T) > \varepsilon + y_{11}(T) - 1\right\}\right)\end{aligned}$$

Of course second chaos r.v. are not symmetric, but since we do not know the sign of $1 - y_{11}(T)$ and second Wiener chaos can be skewed in either direction, there is no loss of efficiency to treat $\tilde{Y}_{11}(T)$ and $-\tilde{Y}_{11}(T)$ in the same fashion. Recall the from Proposition 2.7.13 of [8] that any second chaos r.v. F has the following representation

$$F = \sum_{n=1}^{+\infty} \lambda_n \left(Z_n^2 - 1\right)$$

where $\{\lambda_n, n \geq 1\}$ is a sequence of reals for which $|\lambda_n|$ is decreasing and $(Z_n)_{n\geq 1}$ are independent standard Gaussian random variables and

$$\mathrm{Var}(F) = 2\sum_{n=1}^{+\infty} |\lambda_n|^2 < +\infty.$$

Moreover, by the product formula (2.3), if F is a quadratic functional of a Gaussian process, then $\sum_{n=1}^{+\infty} \lambda_n$ is the expectation of that functional. Therefore with $F = \tilde{Y}_{11}(T)$, there exists $\{\lambda_n(T), n \geq 1\}$ and $(Z_n)_{n\geq 1}$ iid $\mathcal{N}(0,1)$, such that

$$\tilde{Y}_{11}(T) = \sum_{n=1}^{+\infty} \lambda_n(T) \left(Z_n^2 - 1\right).$$

One immediately checks that the expression to be added to $\tilde{Y}_{11}(T)$ to make it a quadratic functional is $\mu_\theta(T) - 2\theta\mathbf{E}[\bar{X}_1^2(T)]$ which is equal to $\sum_{n=1}^{+\infty}\lambda_n(T)$. Therefore, the sequence $\{\lambda_n(T), n \geq 1\}$ satisfies

$$\begin{cases} 2\sum\limits_{n=1}^{+\infty}\lambda_n^2(T) = \mathrm{Var}(\tilde{Y}_{11}(T)) \leqslant \frac{c(\theta)}{T}, \\ \\ |\sum\limits_{n=1}^{+\infty}\lambda_n(T)| = 1 + O(\frac{1}{T}). \end{cases}$$

From the representation we just established, we will set a general global tail for any r.v. in the second Wiener chaos, which is convenient for our purposes. Let $Y = \sum_{n=1}^{+\infty}\lambda_n(Z_n^2 - 1)$ let $\sigma := \sqrt{\mathrm{Var}(Y)} = \sqrt{2\sum_{n=1}^{+\infty}|\lambda_n|^2}$ and $v = \sum_{n=1}^{+\infty}\lambda_n$. Assume that $|v| < +\infty$. Let $\beta > 0$ which is a constant to be chosen later. Then by Markov's inequality, we have for all y

$$\begin{aligned} \mathbf{P}(Y > y) &\leqslant e^{-\frac{y}{\beta}} \times \mathbf{E}\left[e^{\frac{Y}{\beta}}\right] = e^{-\frac{y}{\beta}} \times \mathbf{E}\left[e^{\frac{1}{\beta}\sum\limits_{n=1}^{+\infty}\lambda_n(Z_n^2-1)}\right] \\ &= e^{-\frac{y}{\beta}}\prod_{n=1}^{+\infty}\mathbf{E}\left[e^{\lambda_n\frac{Z_n^2}{\beta}}\right] \times e^{-\frac{v}{\beta}} \\ &= e^{-\frac{y+v}{\beta}} \times \left(\prod_{n=1}^{+\infty}\frac{1}{\sqrt{1-\frac{2\lambda_n}{\beta}}}\right), \end{aligned} \tag{3.21}$$

This formula requires that $2\lambda_n/\beta < 1$ for all $n \geq 1$, but since the sign of λ_n is unknown and $|\lambda_n|$ decreases, it is sufficient to require that $\beta > 2|\lambda_1|$. Since $2\sum_{n=1}^{+\infty}|\lambda_n|^2 = \mathrm{Var}(Y)$, we can say that $\sqrt{2}|\lambda_n| < \sqrt{\mathrm{Var}(Y)} = \sigma$ for any n. Therefore, to be completely safe we choose to require that $\beta \geq 2\sqrt{2}\sigma$. This implies that for every n, $2\lambda_n/\beta \leqslant \sqrt{2}\sigma/\beta \leqslant 1/2$. We must also check that the product in (3.21) converges. In fact, notice that since $\forall u \in [0, 1/2], -\ln(1-u) \leqslant u + \frac{u^2}{2} + u^3$. Thus, since we just checked that $2\lambda_n/\beta \leqslant 1/2$, for all n, we have

$$\ln\left(\prod_{n=1}^{+\infty}\frac{1}{\sqrt{1-\frac{2\lambda_n}{\beta}}}\right) \leqslant \frac{1}{2}\sum_{n=1}^{+\infty}\left(\frac{2\lambda_n}{\beta} + \frac{2\lambda_n^2}{\beta^2} + \frac{8\lambda_n^3}{\beta^3}\right).$$

Thus, we get

$$\mathbf{P}(Y > y) \leqslant e^{-\frac{y}{\beta}} \times e^{\frac{\mathrm{Var}(Y)}{2\beta^2}} \times e^{\frac{k_3(Y)}{2\beta^3}}. \tag{3.22}$$

where we used the fact that (see Proposition 2.7.13 of [8])

$$\sum_{n=1}^{+\infty}\lambda_n^3 = \frac{1}{8}k_3(Y),$$

where $k_3(Y)$ denotes the third cumulant of Y. Since Y is centered then this cumulant is equal to the third moment: $k_3(Y) = \mathbf{E}[Y^3]$. Applying inequality (3.22) to $Y = \tilde{Y}_{11}(T)$ where

$$\sigma \leqslant \sqrt{\frac{cst(\theta)}{T}} \quad \text{and} \quad |v| = 1 + O(\frac{1}{T}),$$

we can pick any $\beta \geq 2\sqrt{2}\sigma$, thus we can chose $\beta = 2\sqrt{\frac{2cst(\theta)}{T}}$. This implies that $\frac{\text{Var}(\tilde{Y}_{11}(T))}{2\beta^2} = \frac{\sigma^2}{2\beta^2} \leqslant \frac{1}{16}$. On the other hand, since $\tilde{Y}_{11}(T)$ is in the second Wiener chaos, then by the hypercontractivity property in Section 2, $k_3(\tilde{Y}_{11}(T)) \leqslant 9\text{Var}(\tilde{Y}_{11}(T))^{3/2}$, and consequently $\frac{1}{2}\frac{k_3(\tilde{Y}_{11}(T))}{\beta^3} \leqslant \frac{9}{2} \times \left(2\sum_{n=1}^{+\infty}\lambda_n^2/\beta^2\right)^{3/2}$. Finally, since $\beta^2 > 8\sum_{n=1}^{+\infty}\lambda_n^2$, thus we get the following:

$$\mathbf{P}(\tilde{Y}_{11}(T) > y) \leqslant K \times \exp\left(-\frac{y\sqrt{T}}{2\sqrt{2cst(\theta)}}\right) \tag{3.23}$$

where

$$K := \exp\left(\frac{1}{16} + \frac{9}{2} \times \frac{1}{8^{3/2}}\right) = \exp\left(\frac{9\sqrt{2}+4}{64}\right).$$

We now replace y by $\varepsilon + 1 - y_{11}(T)$. Note that we must also evaluate $\mathbf{P}(-\tilde{Y}_{11}(T) > \varepsilon + 1 - y_{11}(T))$ but since the signs of λ_n and $1 - y_{11}(T)$ are not known, this will yield exactly to the same estimate as $\mathbf{P}(\tilde{Y}_{11}(T) > \varepsilon + 1 - y_{11}(T))$. Thus from the estimate (3.23), we get

$$\mathbf{P}\left(|\bar{Y}_{11}(T) - 1| > \varepsilon\right) \leqslant 2K \exp\left(-\frac{y\sqrt{T}}{2\sqrt{2cst(\theta)}}\right)$$

Let us denote $c := 2\sqrt{2cst(\theta)}$, we must choose ε. Let

$$\varepsilon = d \times \frac{\ln(T)}{\sqrt{T}}, \tag{3.24}$$

where d is some constant to be chosen as well. Thus since we proved in (3.20) that $|1 - y_{11}| \leqslant \frac{5}{2}\frac{1}{\theta T}$, we get

$$y := \varepsilon + 1 - y_{11}(T) \geq \varepsilon - |1 - y_{11}(T)| \geq d \times \frac{\ln(T)}{\sqrt{T}} - \frac{5}{2\theta T}.$$

For T large enough, for instance for $T > T^* = e \vee \left(\frac{5}{d\theta}\right)^2$, we have

$$d \times \frac{\ln(T)}{\sqrt{T}} > \frac{5}{\theta T}$$

so that

$$y > d \times \frac{\ln(T)}{2\sqrt{T}}$$

and hence

$$\exp\left(-\frac{y\sqrt{T}}{2\sqrt{2cst(\theta)}}\right) \leqslant \exp\left(-\frac{d\ln T}{2c}\right) = T^{-d/(2c)}.$$

Thus it's sufficient to choose $d = c$, obtaining

$$\mathbf{P}\left(|\bar{Y}_{11}(T) - 1| > \varepsilon\right) + 3\varepsilon \leqslant \frac{K}{\sqrt{T}} + 3c\frac{\ln(T)}{\sqrt{T}} \sim 3c\frac{\ln(T)}{\sqrt{T}} \quad \text{for } T \quad \text{large} \tag{3.25}$$

Summarizing, with $K := e^{(9\sqrt{2}+4)/64}$ and

$$c := 2\sqrt{2cst(\theta)} = 2\sqrt{8(3 + 7/(4\theta))/\theta + 128/\theta^2},$$

for $T > T^*(\theta) := e \vee (\frac{5}{2\theta\sqrt{2cst(\theta)}})^2$, then according to inequality (3.15), with the choice for ε in (3.24) with $d = c$, we have the following estimate for the denominator term $D(T)$ defined in (3.14), as follows

$$\mathbf{P}(|D-1| > \varepsilon) + \varepsilon \leqslant \frac{4K}{\sqrt{T}} + \frac{12c\ln(T)}{\sqrt{T}}, \quad \text{for } T > T^*(\theta) := \max\left(e, \frac{25}{8\theta^2 cst(\theta)}\right).$$

Thus, from inequalities (3.17) and (3.25), we get the following theorem for the convergence in law of Yule's statistic $\rho(T)$ as $T \to +\infty$.

Theorem 3.8. *There exists a constant $c(\theta)$ depending only on θ such that for T large enough, we have*

$$d_{Kol}\left(\sqrt{\theta}\sqrt{T}\rho(T), \mathcal{N}(0,1)\right) \leqslant c(\theta)\frac{\ln(T)}{\sqrt{T}}.$$

In particular,

$$\sqrt{\theta}\sqrt{T}\rho(T) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \quad \textit{as} \quad T \to +\infty.$$

Remark 3.9. From expression $cst(\theta) = 4(3 + 7/(4\theta))/\theta + 64/\theta^2$ at the end of the calculation preceding Theorem 3.8, and similar estimates elsewhere above, a detailed analysis of how these constants depend on θ show that for $\theta > 1$,

$$0 < c(\theta) < \frac{c_u}{\sqrt{\theta}},$$

where c_u is a universal constant. This analysis is omitted for conciseness. In particular, $c(\theta) \to 0$ as $\theta \to +\infty$.

4. Discrete observations

We assume now that the pair (X_1, X_2) of Ornstein-Uhlenbeck processes is observed at n equally spaced discrete time instants $t_k := k \times \Delta_n$, where Δ_n is the observation mesh and $T_n := n\Delta_n$ is the length of the "observation window". We assume $\Delta_n \to 0$ and $T_n \to +\infty$, as $n \to +\infty$. The aim of this section is to prove a CLT for the following statistic $\tilde{\rho}(n)$ which can be considered as a discrete version of Yule's nonsense correlation statistic $\rho(T)$, defined by

$$\tilde{\rho}(n) := \frac{\tilde{Y}_{12}(n)}{\sqrt{\tilde{Y}_{11}(n) \times \tilde{Y}_{22}(n)}} \tag{4.1}$$

where $\tilde{Y}_{ij}(n)$, $i, j = 1, 2$ are the Riemann-type discretization of $Y_{ij}(T)$ defined as follows

$$\tilde{Y}_{ij}(n) := \Delta_n \sum_{k=0}^{n-1} X_i(t_k) X_j(t_k) - T_n \tilde{X}_i(n) \tilde{X}_j(n), \quad i, j = 1, 2, \tag{4.2}$$

with $\tilde{X}_i(n)$ denoting the empirical mean-process of X_i, $i = 1, 2$, namely

$$\tilde{X}_i(n) := \frac{1}{n} \sum_{k=0}^{n-1} X_i(t_k), \quad i = 1, 2$$

As in the continuous case, we make use of the following expression of $\tilde{\rho}(n)$ along with Proposition 3.4 in order to prove its convergence in law to a Gaussian distribution:

$$\sqrt{\theta}\sqrt{T_n}\tilde{\rho}(n) = \frac{2\theta^{3/2} \frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}}{2\theta \sqrt{\frac{\tilde{Y}_{11}(n)}{T_n} \times \frac{\tilde{Y}_{22}(n)}{T_n}}}. \tag{4.3}$$

4.1. Convergence in law of $\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}$

From the expression of $\tilde{Y}_{12}(n)$ given in (4.2), we can write

$$\frac{\tilde{Y}_{12}(n)}{\sqrt{T}_n} = A(n) - B(n), \tag{4.4}$$

where

$$A(n) := \frac{\sqrt{T}_n}{n} \sum_{k=0}^{n-1} X_1(t_k) X_2(t_k) \quad \text{and} \quad B(n) := \sqrt{T_n} \tilde{X}_1(n) \tilde{X}_2(n).$$

We also defined the following random sequence

$$\delta(n) := A(n) - F_{T_n},$$

where F_{T_n} is defined in (3.2). The following lemma holds.

Lemma 4.1. *Assume that $\Delta_n \to 0$ as $n \to +\infty$ and that $T_n = n\Delta_n \to +\infty$, then, there exists a constant $C_\theta := 4 \times \max\left(\frac{8}{9\theta}, \frac{\sqrt{2}}{3}\frac{1}{\theta^{1/2}}, \frac{1}{4}\right)$, such that*

$$\mathbf{E}[\delta^2(n)] \leqslant C_\theta \times n\Delta_n^2,$$

In particular if $n\Delta_n^2 \to 0$ as $n \to +\infty$, $\mathbf{E}[\delta^2(n)] \to 0$, as $n \to +\infty$.

Proof. We have

$$\begin{aligned}\delta(n) &= A(n) - F_{T_n}\\ &= \sqrt{T_n}\left(\frac{1}{n}\sum_{k=0}^{n-1} X_1(t_k)X_2(t_k) - \frac{1}{T_n}\int_0^{T_n} X_1(u)X_2(u)du\right)\\ &= \frac{1}{\sqrt{T_n}}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} (X_1(t_k)X_2(t_k) - X_1(u)X_2(u))du\end{aligned}$$

By Cauchy Schwartz inequality and the fact that $\sup_{t\geq 0}\mathbf{E}[X_i^2(t)] = \frac{1}{2\theta}$, $i = 1, 2$, we get

$$\begin{aligned}&\mathbf{E}[\delta^2(n)]\\ &= \frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}} \mathbf{E}\left[(X_1(t_{k_1})X_2(t_{k_1})\right.\\ &\quad \left. -X_1(u)X_2(u))(X_1(t_{k_2})X_2(t_{k_2}) - X_1(v)X_2(v))\right] dudv\\ &\leqslant \frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}} \|X_1(t_{k_1})X_2(t_{k_1})\\ &\quad - X_1(u)X_2(u)\|_{L^2}\|X_1(t_{k_2})X_2(t_{k_2}) - X_1(v)X_2(v)\|_{L^2}dudv\\ &\leqslant \frac{2}{\theta}\frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}} \|X_i(t_{k_1})\\ &\quad - X_i(u)\|_{L^2}\|X_i(t_{k_2}) - X_i(v)\|_{L^2}dudv, \quad i = 1, 2,\end{aligned}$$

where we used the fact that X_1 and X_2 are two Gaussian processes equal in law. On the other hand since $(X_i(t_k) - X_i(u)) = (Z_i(t_k) - Z_i(u)) - Z_0(e^{-\theta t_k} - e^{-\theta u})$, $i = 1, 2$. Recall that $Z_i(r) := \int_{-\infty}^{r} e^{-\theta(r-t)}dW^i(t)$, $i = 1, 2$. We get

$$\begin{aligned}&\frac{2}{\theta}\frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}} \|X_1(t_{k_1}) - X_1(u)\|_{L^2}\|X_1(t_{k_2}) - X_1(v)\|_{L^2}dudv\\ &\quad \leqslant A_1(n) + A_2(n) + A_3(n) + A_4(n)\end{aligned}$$

where

$$A_1(n) := \frac{2}{\theta}\frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}} \|Z_1(t_{k_1}) - Z_1(u)\|_{L^2}\|Z_1(t_{k_2}) - Z_1(v)\|_{L^2}dudv,$$

$$A_2(n) := \frac{1}{T_n}\frac{\sqrt{2}}{\theta^{3/2}}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}} |e^{-\theta t_{k_2}} - e^{-\theta v}| \times \|Z_1(t_{k_1}) - Z_1(u)\|_{L^2}dudv,$$

$$A_3(n) := \frac{1}{T_n}\frac{\sqrt{2}}{\theta^{3/2}}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}} |e^{-\theta t_{k_1}} - e^{-\theta u}| \times \|Z_1(t_{k_2}) - Z_1(v)\|_{L^2}dudv,$$

$$A_4(n) := \frac{1}{T_n}\frac{1}{\theta^2}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}}|e^{-\theta t_{k_1}} - e^{-\theta u}| \times |e^{-\theta t_{k_2}} - e^{-\theta v}|dudv.$$

We will use in the sequel the fact that the increments of the process Z_1 satisfies $\mathbf{E}[(Z_1(t) - Z_1(s))^2] \leqslant |t-s|$, $t, s \geq 0$. For the first sequence $A_1(n)$ we have

$$\begin{aligned}
A_1(n) &\leqslant \frac{2}{\theta}\frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}}|t_{k_1}-u|^{1/2}|t_{k_2}-v|^{1/2}dudv,\\
&= \frac{2}{\theta}\frac{\Delta_n^3}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_0^1\int_0^1 t^{1/2}s^{1/2}dtds,\\
&= \frac{8}{9\theta}\times n\Delta_n^2.
\end{aligned}$$

where we used the change of variables $t = \frac{u-t_{k_1}}{\Delta_n}$, $s = \frac{v-t_{k_2}}{\Delta_n}$. For $A_2(n)$, we have

$$\begin{aligned}
A_2(n) &\leqslant \frac{\sqrt{2}}{\theta^{1/2}}\frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}}|t_{k_1}-u|^{1/2}|t_{k_2}-v|dudv,\\
&= \frac{\sqrt{2}}{\theta^{1/2}}\frac{1}{T_n}\Delta_n^{7/2}\sum_{k_1,k_2=0}^{n-1}\int_0^1\int_0^1 t^{1/2}sdtds\\
&\leqslant \frac{\sqrt{2}}{3}\frac{1}{\theta^{1/2}}\times n\Delta_n^{5/2}.
\end{aligned}$$

Similarly, we have

$$A_3(n) \leqslant \frac{\sqrt{2}}{3}\frac{1}{\theta^{1/2}}\times n\Delta_n^{5/2}.$$

For the last sequence $A_4(n)$, we get

$$\begin{aligned}
A_4(n) &\leqslant \frac{1}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_{t_{k_1}}^{t_{k_1+1}}\int_{t_{k_2}}^{t_{k_2+1}}|t_{k_1}-u||t_{k_2}-v|dudv\\
&\leqslant \frac{\Delta_n^4}{T_n}\sum_{k_1,k_2=0}^{n-1}\int_0^1\int_0^1 tsdtds\\
&\leqslant \frac{n\Delta_n^3}{4}.
\end{aligned}$$

The desired result following using the previous inequalities and the fact that $\Delta_n \to 0$ as $n \to +\infty$. □

Remark 4.2. One possible mesh that satisfies the assumptions of Lemma 4.1 is $\Delta_n = n^{-\lambda}$ with $1/2 < \lambda < 1$.

Proposition 4.3. *There exists a constant $C(\theta)$ such that*

$$d_{Kol}\left(A(n), \mathcal{N}\left(0, \frac{1}{4\theta^3}\right)\right) \leqslant C(\theta) \times \max\left((n\Delta_n)^{-1/2}, (n\Delta_n^2)^{\frac{1}{3}}\right)$$

In particular, if $n\Delta_n^2 \to 0$ as $n \to +\infty$, we get

$$A(n) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{4\theta^3}\right),$$

as $n \to +\infty$.

Proof. Since $A(n) = \delta(n) + F_{T_n}$, we have by Proposition 3.4 assertion 1. the following estimate

$$\begin{aligned} d_{Kol}\left(A(n), \mathcal{N}(0, \frac{1}{4\theta^3})\right) &\leqslant d_{Kol}\left(F_{T_n}, \mathcal{N}(0, \frac{1}{4\theta^3})\right) + \mathbf{P}\left(|2\theta^{3/2}\delta(n)| > \varepsilon\right) + \varepsilon \\ &\leqslant d_{Kol}\left(F_{T_n}, \mathcal{N}(0, \frac{1}{4\theta^3})\right) + 4\theta^3 \times \frac{\mathbf{E}[\delta^2(n)]}{\varepsilon^2} + \varepsilon \\ &\leqslant \frac{c(\theta)}{\sqrt{T_n}} + \inf_{\varepsilon > 0} g_n(\varepsilon) \end{aligned}$$

where $g_n(\varepsilon) = 4\theta^3 \times C_\theta \times \frac{n\Delta_n^2}{\varepsilon^2} + \varepsilon$, since g_n is convex on $\mathbb{R}_+$, $\arg\inf_{\varepsilon>0} g_n(\varepsilon) = \varepsilon^*(n) = \left(4\theta^3 \times C_\theta n\Delta_n^2\right)^{1/3}$, where C_θ is the constant from Lemma 4.1. Thus, we get the following estimate for the convergence in law of the random sequence $A(n)$:

$$d_{Kol}\left(A(n), \mathcal{N}(0, \frac{1}{4\theta^3})\right) \leqslant C(\theta) \times \max\left((n\Delta_n)^{-1/2}, (n\Delta_n^2)^{\frac{1}{3}}\right),$$

which ends the proof. □

On the other hand, from the decomposition (4.4), Corollary 3.6, Proposition 3.4, there exists a constant β such that $4\theta^{3/2}\frac{\sqrt{3}}{3}\sqrt{T_n}\mathbf{E}[\tilde{X}_1^2(n)] < \beta < (C \times \mathbf{E}[|B(n)|] \wedge 4)$, and such that

$$d_{Kol}\left(\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}, \mathcal{N}(0, \frac{1}{4\theta^3})\right) \leqslant d_{Kol}\left(A(n), \mathcal{N}(0, \frac{1}{4\theta^3})\right) + \beta\left(1 + \ln(\frac{4}{\beta})\right), \quad (4.5)$$

Let us now compute $\|\tilde{X}_1(n)\|_{L^2}^2$,

$$\begin{aligned} \tilde{X}_1(n) &= \frac{1}{n}\sum_{k=0}^{n-1} I_1^{W_1}(f_{t_k}) \\ &= I_1^{W_1}(g_n), \end{aligned}$$

with $g_n := \frac{1}{n}\sum_{k=0}^{n-1} f_{t_k}$. Thus similarly to a previous calculation, we have

$$\mathbf{E}[\tilde{X}_1^2(n)] = \|g_n\|_{L^2([0,T_n])}^2$$

$$
\begin{aligned}
&= \frac{1}{n^2} \sum_{k_1,k_2=0}^{n-1} \langle f_{t_{k_1}}, f_{t_{k_2}} \rangle_{L^2([0,T_n])} \\
&= \frac{1}{n^2} \sum_{k_1,k_2=0}^{n-1} \mathbf{E}[X_1(t_{k_1}) X_1(t_{k_2})] \\
&= \frac{2}{n^2} \sum_{k_1=0}^{n-1} \mathbf{E}[X_1(t_{k_1})^2] + \frac{2}{n^2} \sum_{\substack{k_1,k_2=0 \\ k_1 \neq k_2}}^{n-1} \mathbf{E}[X_1(t_{k_1}) X_1(t_{k_2})] \\
&= \frac{2}{n^2} \sum_{k_1=0}^{n-1} \left(\frac{1-e^{-2\theta t_{k_1}}}{2\theta} \right) + \frac{4}{n^2} \sum_{k_1=0}^{n-2} \sum_{k_2=k_1+1}^{n-1} \mathbf{E}[X_1(t_{k_1}) X_1(t_{k_2})] \\
&\leqslant \frac{1}{\theta}\frac{1}{n} + \frac{3}{\theta^2}\frac{1}{n\Delta_n} \leqslant \frac{1}{\theta}\left(1+\frac{3}{\theta}\right)(n\Delta_n)^{-1},
\end{aligned} \tag{4.6}
$$

where for the last equality we used the fact that

$$
\frac{4}{n^2} \sum_{k_1=0}^{n-2} \sum_{k_2=k_1+1}^{n-1} \mathbf{E}[X_1(t_{k_1}) X_1(t_{k_2})] \leqslant \frac{4}{n^2} \sum_{k_1=0}^{n-2} \sum_{k_2=k_1+1}^{n-1} \rho(t_{k_2} - t_{k_1})
$$

Recall that $\rho(r-s) := \mathbf{E}[Z_r^i Z_s^i]$, where $Z_r^i := \int_{-\infty}^r e^{-\theta(r-t)} dW^i(t),\ i = 1,2.$ Moreover,

$$
\begin{aligned}
\frac{4}{n^2} \sum_{k_1=0}^{n-2} \sum_{k_2=k_1+1}^{n-1} \rho(t_{k_2} - t_{k_1}) &= \frac{4}{n^2} \sum_{k_1=0}^{n-2} \sum_{k_2=k_1+1}^{n-1} \rho(\Delta_n(k_2-k_1)) \\
&= \frac{4}{n^2} \sum_{r=1}^{n-1} (n-r)\rho(\Delta_n r) \\
&= \frac{2}{\theta}\frac{1}{n^2} \sum_{r=1}^{n-1} (n-r) e^{-\theta\Delta_n r} \\
&\leqslant \frac{2}{\theta}\frac{1}{n}\frac{1}{(1-e^{-\Delta_n\theta})} \\
&\leqslant \frac{3}{\theta^2}(n\Delta_n)^{-1},
\end{aligned}
$$

for large n. We deduce that there exists a constant $c(\theta)$, such that for large n

$$
\begin{aligned}
\mathbf{E}[|B(n)|] &= \sqrt{T_n}\mathbf{E}[|\tilde{X}_1(n)|]\mathbf{E}[|\tilde{X}_2(n)|] \\
&\leqslant c(\theta)(n\Delta_n)^{-1/2}.
\end{aligned}
$$

Therefore, from equation (4.5) and the fact that the function $x \mapsto x\left(1+\ln(\frac{4}{x})\right)$ is increasing on $(0,4)$, we have for n large enough

$$
d_{Kol}\left(\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}, \mathcal{N}(0, \frac{1}{4\theta^3}) \right) \leqslant d_{Kol}\left(A(n), \mathcal{N}(0, \frac{1}{4\theta^3}) \right) + \frac{c(\theta)}{2}\frac{\ln(n\Delta_n)}{\sqrt{n\Delta_n}}.
$$

The following proposition follows.

Theorem 4.4. *There exists a constant $c(\theta)$ depending only on θ such that for large n*

$$d_{Kol}\left(\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}, \mathcal{N}(0, \frac{1}{4\theta^3})\right) \leqslant c(\theta) \times \ln(n\Delta_n) \max\left((n\Delta_n)^{-1/2}, (n\Delta_n^2)^{\frac{1}{3}}\right)$$

In particular, if $n\Delta_n^2 \to 0$ as $n \to +\infty$,

$$\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{4\theta^3}\right),$$

as $n \to +\infty$.

Example 1. If $\Delta_n = n^{-\lambda}$ with $\frac{1}{2} < \lambda < 1$, then we have

$$d_{Kol}\left(\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}, \mathcal{N}(0, \frac{1}{4\theta^3})\right) \leqslant C(\theta) \times (1-\lambda) \times \begin{cases} \ln(n) \times n^{\frac{1-2\lambda}{3}} & \text{if } \frac{1}{2} < \lambda \leqslant \frac{5}{7} \\ \ln(n) \times n^{\frac{\lambda-1}{2}} & \text{if } \frac{5}{7} \leqslant \lambda < 1. \end{cases}$$

Consequently,

$$\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{4\theta^3}\right)$$

as $n \to +\infty$.

4.2. The denominator term

We denote the denominator term in $\tilde{\rho}(n)$ by $\tilde{D}(n)$, i.e.

$$\tilde{D}(n) := 2\theta\sqrt{\frac{\tilde{Y}_{11}(n)}{T_n} \times \frac{\tilde{Y}_{22}(n)}{T_n}}$$

According to Proposition 3.4 assertion 2, we need to estimate $\mathbf{P}(|\tilde{D}(n) - 1| > \varepsilon)$ for instance for $0 < \varepsilon < 1$. Thus, denoting $\bar{Y}_{11}(n) := \frac{\tilde{Y}_{11}(n)}{T_n/2\theta}$, then similarly to the calculations performed in the continuous case, i.e. (3.15), we get

$$\mathbf{P}\left(|\tilde{D}(n) - 1| > \varepsilon\right) \leqslant 4\mathbf{P}\left(|\bar{Y}_{11}(n) - 1| > \frac{\varepsilon}{3}\right)$$

Writing ε instead of $\varepsilon/3$, by Proposition 3.4 assertion 2 applied to $\tilde{\rho}(n)$, we get

$$d_{Kol}\left(\sqrt{\theta}\sqrt{T_n}\tilde{\rho}(n), N\right) \leqslant d_{Kol}\left(2\theta^{3/2}\frac{\tilde{Y}_{12}(n)}{\sqrt{T_n}}, N\right) + 4 \times \mathbf{P}\left(|\bar{Y}_{11}(n) - 1| > \varepsilon\right) + 3\varepsilon. \tag{4.7}$$

where $N \sim \mathcal{N}(0,1)$. It remains to control the term $\mathbf{P}\left(|\bar{Y}_{11}(n)-1|>\varepsilon\right)$. we have

$$\bar{Y}_{11}(n) = D_{1,n} + D_{2,n} - 2\theta\tilde{X}_1^2(n), \tag{4.8}$$

where

$$D_{1,n} := \frac{2\theta}{n}\sum_{k=0}^{n-1}\left(X_1^2(t_k) - \mathbf{E}[X_1^2(t_k)]\right) \quad \text{and} \quad D_{2,n} := \frac{2\theta}{n}\sum_{k=0}^{n-1}\mathbf{E}[X_1^2(t_k)] \tag{4.9}$$

Lemma 4.5. *Consider $D_{1,n}$ defined in (4.9), then for every large n, we have*

$$\mathbf{E}\left[D_{1,n}^2\right] \leqslant 2\left(1+\frac{3}{2\theta}\right) \times (n\Delta_n)^{-1}.$$

Proof. The sequence $D_{1,n}$ can be written as follows:

$$\begin{aligned} D_{1,n} &= \frac{2\theta}{n}\sum_{k=0}^{n-1}\left(I_1^{W_1}(f_{t_k}) - \|f_{t_k}\|_{L^2([0,T_n])}\right) \\ &= \frac{2\theta}{n}\sum_{k=0}^{n-1} I_2^{W_1}(f_{t_k}^{\otimes 2}) \\ &= I_2^{W_1}(k_n), \end{aligned}$$

where $k_n := \frac{2\theta}{n}\sum_{k=0}^{n-1} f_{t_k}^{\otimes 2}$. Thus

$$\begin{aligned} \mathbf{E}[D_{1,n}^2] &= 2\|k_n\|^2_{L^2([0,T_n]^2)} \\ &= \frac{2\times(2\theta)^2}{n^2}\sum_{k_1,k_2=0}^{n-1}\langle f_{t_{k_1}}^{\otimes 2}, f_{t_{k_2}}^{\otimes 2}\rangle_{L^2([0,T_n]^2)} \\ &= \frac{2\times(2\theta)^2}{n^2}\sum_{k_1,k_2=0}^{n-1}\left(\langle f_{t_{k_1}}, f_{t_{k_2}}\rangle_{L^2([0,T_n])}\right)^2 \\ &= \frac{2\times(2\theta)^2}{n^2}\sum_{k_1,k_2=0}^{n-1}\left(\mathbf{E}[X_1(t_{k_1})X_1(t_{k_2})]\right)^2 \\ &= \frac{2\times(2\theta)^2}{n^2}\sum_{k_1=0}^{n-1}\mathbf{E}[X_1(t_{k_1})^2]^2 + \frac{2\times(2\theta)^2}{n^2}\sum\sum_{\substack{k_1,k_2=0\\k_1\neq k_2}}^{n-1}(\mathbf{E}[X_1(t_{k_1})X_1(t_{k_2})])^2 \\ &= \frac{2\times(2\theta)^2}{n^2}\sum_{k_1=0}^{n-1}\left(\frac{1-e^{-2\theta t_{k_1}}}{2\theta}\right)^2 \\ &\quad + \frac{4\times(2\theta)^2}{n^2}\sum_{k_1=0}^{n-2}\sum_{k_2=k_1+1}^{n-1}\left(\mathbf{E}[X_1(t_{k_1})X_1(t_{k_2})]\right)^2 \\ &\leqslant \frac{2}{n} + \frac{4\times(2\theta)^2}{n^2}\sum_{k_1=0}^{n-2}\sum_{k_2=k_1+1}^{n-1}\left(\mathbf{E}[X_1(t_{k_1})X_1(t_{k_2})]\right)^2. \end{aligned}$$

For the right hand partial sum, we can write

$$\frac{4}{n^2}\sum_{k_1=0}^{n-2}\sum_{k_2=k_1+1}^{n-1}\left(\mathbf{E}[X_1(t_{k_1})X_1(t_{k_2})]\right)^2 \leqslant \frac{4}{n^2}\sum_{k_1=0}^{n-2}\sum_{k_2=k_1+1}^{n-1}\rho^2(t_{k_2}-t_{k_1})$$

Moreover,

$$\begin{aligned}\frac{4}{n^2}\sum_{k_1=0}^{n-2}\sum_{k_2=k_1+1}^{n-1}\rho^2(t_{k_2}-t_{k_1}) &= \frac{4}{n^2}\sum_{k_1=0}^{n-2}\sum_{k_2=k_1+1}^{n-1}\rho^2(\Delta_n(k_2-k_1))\\ &= \frac{4}{n^2}\sum_{r=1}^{n-1}(n-r)\rho^2(\Delta_n r)\\ &= \frac{1}{\theta^2}\frac{1}{n^2}\sum_{r=1}^{n-1}(n-r)e^{-2\theta\Delta_n r}\\ &\leqslant \frac{1}{\theta^2}\frac{1}{n}\frac{1}{(1-e^{-2\Delta_n\theta})}\\ &\leqslant \frac{3}{4\theta^3}\frac{1}{n\Delta_n},\end{aligned}$$

for every large n. The desired result follows since $\max\left(n^{-1},(n\Delta_n)^{-1}\right)=(n\Delta_n)^{-1}$. □

For the sequence $D_{2,n}$, we have

$$\begin{aligned}|D_{2,n}-1| &\leqslant \frac{2\theta}{n}\sum_{k=0}^{n-1}\left|\mathbf{E}[X_1^2(t_k)]-\frac{1}{2\theta}\right|\\ &\leqslant \frac{2\theta}{n}\sum_{k=0}^{n-1}e^{-2\theta k\Delta_n} \leqslant \frac{2\theta}{n}\frac{1}{(1-e^{-2\theta\Delta_n})} \leqslant \frac{3}{2}\frac{1}{n\Delta_n},\end{aligned} \tag{4.10}$$

for every large n. It follows that $D_{2,n}=1+O\left((n\Delta_n)^{-1}\right)$ for n large. Therefore, following exactly the same analysis done for the denominator term in the continuous case and choosing $\varepsilon=\varepsilon(n)=c\times\frac{\ln(n\Delta_n)}{\sqrt{n\Delta_n}}$, we get the existence of a constant $c(\theta)$, such that for n large enough

$$\mathbf{P}\left(|\tilde{D}(n)-1|>\varepsilon(n)\right)\leqslant c(\theta)\frac{\ln(n\Delta_n)}{\sqrt{n\Delta_n}}, \tag{4.11}$$

The previous estimates will allow to prove first a Strong law result by showing that $\tilde{\rho}(n)$ converges to 0 almost surely as $n\to+\infty$, then the convergence in law of the statistic $\tilde{\rho}(n)$ given with its rate of convergence as $n\to+\infty$.

Proposition 4.6. *Assume that $\Delta_n=n^{-\lambda}$, for $\frac{1}{2}<\lambda<1$, then we have almost surely*

$$\tilde{\rho}(n)\longrightarrow 0 \quad as \quad n\to+\infty.$$

Proof. We can write $\tilde{\rho}(n)$ as follows

$$\tilde{\rho}(n) = \frac{\frac{\tilde{Y}_{12}(n)}{T_n}}{\sqrt{\frac{\tilde{Y}_{11}(n)}{T_n} \times \frac{\tilde{Y}_{22}(n)}{T_n}}}$$

For the numerator term, we have by the decomposition (4.4),

$$\frac{\tilde{Y}_{12}(n)}{T_n} = \frac{A(n)}{\sqrt{T_n}} - \frac{B(n)}{\sqrt{T_n}}$$

For the sequence $A(n)$, we have

$$\frac{A(n)}{\sqrt{T_n}} = \frac{\delta(n)}{\sqrt{T_n}} + \frac{F_{T_n}}{\sqrt{T_n}},$$

Then by Lemma 4.1, we get $\mathbf{E}\left[\left(\frac{\delta(n)}{\sqrt{T_n}}\right)^2\right] \leqslant C_\theta n^{-\lambda}$, and by the hypercontractivity property and Lemma 2.1, we obtain $\frac{\delta(n)}{\sqrt{T_n}} \to 0$, a.s. as $n \to +\infty$. By the same argument, and using Lemma 3.2, we obtain $\mathbf{E}\left[\left(\frac{F_{T_n}}{\sqrt{T_n}}\right)^2\right] \leqslant C_\theta n^{-(1-\lambda)}$, and we then get $\frac{F_{T_n}}{\sqrt{T_n}} \to 0$, a.s. as $n \to +\infty$. Consequently, we have a.s. as $n \to +\infty$, $\frac{A(n)}{\sqrt{T_n}} \to 0$. For the sequence $B(n)$, $\mathbf{E}\left[\left(\frac{B(n)}{\sqrt{T_n}}\right)^2\right] = \frac{1}{T_n}\mathbf{E}[\tilde{X}_1^2(n)]\mathbf{E}[\tilde{X}_2^2(n)] \leqslant C \times n^{-(1-\lambda)}$, thus $\frac{B(n)}{\sqrt{T_n}}$ a.s. as $n \to +\infty$. Finally,

$$\frac{\tilde{Y}_{12}(n)}{T_n} \longrightarrow 0,$$

a.s. as $n \to +\infty$. For the denominator term, we will need the following proposition

Proposition 4.7. *For every $p \geq 1$, there exists a constant $c(p,\theta)$ depending on p and θ, such that*

$$\mathbf{E}\left[\left|2\theta\sqrt{\frac{\tilde{Y}_{11}(n)}{T_n} \times \frac{\tilde{Y}_{22}(n)}{T_n}} - 1\right|^p\right]^{\frac{1}{p}} \leqslant c(p,\theta) \times T_n^{-\frac{1}{2}}$$

Proof. Using the fact that if $X \geq 0$ p.s. then we have $\mathbf{E}[|\sqrt{X}-1|^p] \leqslant \mathbf{E}[|X-1|^p]$ for every $p > 0$, then by the notation $\bar{Y}_{ii}(n) = \frac{\tilde{Y}_{ii}(n)}{T_n/2\theta}$, $i = 1,2$, we get

$$\mathbf{E}\left[\left|2\theta\sqrt{\frac{\tilde{Y}_{11}(n)}{T_n} \times \frac{\tilde{Y}_{22}(n)}{T_n}} - 1\right|^p\right]^{1/p} \leqslant \times \mathbf{E}\left[\left|\bar{Y}_{11}(n) \times \bar{Y}_{22}(n) - 1\right|^p\right]^{1/p}$$
$$\leqslant \mathbf{E}\left[|\bar{Y}_{11}(n)|^p\right]^{1/p} \times \mathbf{E}\left[|\bar{Y}_{22}(n) - 1|^p\right]^{1/p}$$

$$+\mathbf{E}\left[|\bar{Y}_{11}(n)-1|^p\right]^{1/p}$$

Moreover, since by the decomposition (4.8), $\bar{Y}_{11}(n) = D_{1,n} + D_{2,n} - 2\theta\bar{X}_1^2(n)$, where $D_{1,n}$ and $D_{2,n}$ are defined in (4.9), we get

$$\begin{aligned}\mathbf{E}\left[|\bar{Y}_{11}(n)-1|^p\right]^{1/p} &\leqslant \mathbf{E}\left[|D_{1,n}|^p\right]^{1/p} + |D_{2,n}-1| + 2\theta\mathbf{E}\left[\tilde{X}_1^2(n)^p\right]^{1/p} \\ &\leqslant c(p,\theta)\times T_n^{-1/2},\end{aligned}$$

where we used the hypercontractivity property (2.5), Lemma 4.5 and the estimate (4.10), with

$$\begin{aligned}&c(p,\theta) \\ &= 3\times\max\left(\left\{(p-1)\mathbf{1}_{\{p\geq 2\}}+\mathbf{1}_{\{p=1\}}\right\}\sqrt{2\left(1+\frac{3}{2\theta}\right)}, \frac{3}{2}, (2p-1)\frac{1}{\theta}\left(1+\frac{3}{\theta}\right)\right).\end{aligned}$$

The result of Proposition 4.7 is therefore established. □

By Proposition 4.7, we have for all $\eta > 0$,

$$\begin{aligned}&\sum_{n=1}^{+\infty}\mathbf{P}\left(\left|2\theta\sqrt{\frac{\tilde{Y}_{11}(n)}{T_n}\times\frac{\tilde{Y}_{22}(n)}{T_n}}-1\right|>\eta\right) \\ &\leqslant \frac{1}{\eta^p}\sum_{n=1}^{+\infty}\mathbf{E}\left[\left|2\theta\sqrt{\frac{\tilde{Y}_{11}(n)}{T_n}\times\frac{\tilde{Y}_{22}(n)}{T_n}}-1\right|^p\right] \\ &\leqslant \frac{1}{\eta^p}\sum_{n=1}^{+\infty}\frac{1}{n^{(1-\lambda)p/2}} < +\infty,\end{aligned}$$

for any $p > \frac{2}{1-\lambda}$, it follows from Borel-Cantelli's Lemma that

$$\sqrt{\frac{\tilde{Y}_{11}(n)}{T_n}\times\frac{\tilde{Y}_{22}(n)}{T_n}} \to \frac{1}{2\theta},$$

almost surely as $n\to+\infty$, which finishes the proof. □

Remark 4.8. Proposition 4.6 was proved in the scale $\Delta_n = n^{-\lambda}$, for $\frac{1}{2}<\lambda<1$, for ease of presentation, but one can also show that it holds for any mesh Δ_n satisfying:

- $\sum_{n=1}^{+\infty}\Delta_n^p < \infty$ for some p.
- $\sum_{n=1}^{+\infty}\frac{1}{n^p\Delta_n^p} < +\infty$ for some p.

Theorem 4.9. *There exists a constant $c(\theta)$ such that for n large enough, we have*

$$d_{Kol}\left(\sqrt{\theta}\sqrt{T_n}\tilde{\rho}(n), \mathcal{N}(0,1)\right) \leqslant c(\theta)\times\ln(n\Delta_n)\max\left((n\Delta_n)^{-1/2}, (n\Delta_n^2)^{\frac{1}{3}}\right)$$

In particular, if $n\Delta_n^2 \to 0$ *as* $n \to +\infty$,

$$\sqrt{T_n}\tilde{\rho}(n) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\theta}\right),$$

as $n \to +\infty$.

Remark 4.10. The results obtained in Theorem 4.9 can be as efficient as those of Theorem 3.8, as long as one picks a step size very precisely. In fact, Theorem 4.9 allows us to identify an optimal step size Δ_n immediately, by requiring that $\left(n\Delta_n^2\right)^{1/3}$ is of the same order as $(n\Delta_n)^{-1/2}$. By equating these two terms, we immediately find that it is optimal to choose Δ_n on the order of $n^{-5/7}$. When choosing $\Delta_n = n^{-5/7}$, one then immediately finds that $T_n = n\Delta_n = n^{2/7}$, which means that the speed in the Kolmogorov metric in Theorem 4.9 is bounded above by $\ln(n)n^{-1/7}$ up to a constant. Therefore, in terms of T_n, the rate of convergence is of the order of $\ln(T_n) \times T_n^{-1/2}$. We obtain therefore the same speed as in the continuous case in Theorem 3.8. It is in this sense that the convergence rate in Theorem 4.9 is as efficient in Theorem 3.8.

5. Numerical results

This section contains a numerical study of some of the properties of the discrete version of Yule's nonsense correlation statistic, which we denoted by $\tilde{\rho}(n)$ in (4.1). We first simulate the OU processes X_1 and X_2 according to the following steps:

1. Set the values of θ, the sample size n and the mesh $\Delta_n = n^{-\lambda}$, $\frac{1}{2} < \lambda < 1$.
2. Generate two independent Brownian motions W^1 and W^2.
3. Set $X_0^i = 0$, for $i = 1, 2$ and simulate the observations $X_{\Delta_n}^i$, $X_{2\Delta_n}^i$, ..., $X_{T_n}^i$, $i = 1, 2$ where $T_n = n \times \Delta_n$ following the Euler scheme:

$$X_{t_j}^i = (1 - \theta\Delta_n)X_{t_{j-1}}^i + \left(W_{t_j}^i - W_{t_{j-1}}^i\right) \quad j = 1, .., n, \quad i = 1, 2.$$

4. We obtain a simulation of the sample paths of X^1 and X^2 based on $X_{t_j}^i$, $j = 1, .., n$, $i = 1, 2$ by approximating X^1 and X^2 using the linear process linking the points $(t_j, X_{t_j}^i)_{\{1 \leqslant j \leqslant n\}}$ for $i = 1, 2$ as follows

$$X_t^i = \left(1 - \theta(t - t_{j-1})X_{t_{j-1}}^i\right) + \left(W_t^i - W_{t_{j-1}}^i\right), \quad \text{for all}$$
$$t \in [t_j, t_{j-1}], \quad j = 1, ..., n.$$

The figures below are an example of four sample paths of X^1 and X^2 for different values of the drift parameter θ.

The simulation of the Ornstein-Uhlenbeck sample paths X_1 and X_2 is done for different values of the parameter $\theta = \{0.5, 1, 5, 10\}$, for a sample size $n = 10000$ and a mesh $\Delta_n = 10000^{-0.6}$ ($\lambda = 0.6$) which corresponds to a time horizon

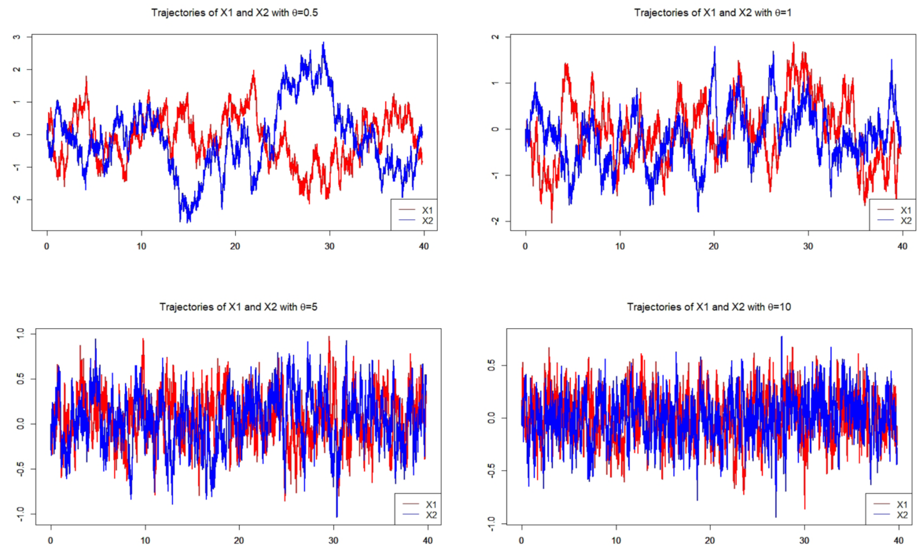


Fig 1. *Sample paths of X1 and X2 for different values of* θ*.*

$T_n \sim 40$. One can see from the figures above how the drift parameter value impacts the variability and raggedness of OU sample paths. The next step is to illustrate numerically Proposition 4.6. The table below shows the mean, the median and standard deviation values for $\tilde{\rho}(n)$ for different values of n, using 500 Monte-Carlo replications for three different values of the drift parameter $\theta = \{1, 5, 10\}$. Table 1 above shows that $\tilde{\rho}(n)$ approaches zero for large values of

Table 1
Estimation results for $n = \{10000, 50000, 100000\}$ *and* $\lambda = 0.6$*.*

		$n = 10000$ $T_n \sim 40$	$n = 50000$ $T_n \sim 76$	$n = 100000$ $T_n \sim 100$
	Mean	-0.01022	0.00667	0.00377
$\theta = 1$	Median	-0.00725	0.00873	0.00168
	S.Dev	0.14990	0.10162	0.10898
	Mean	0.00296	0.00130	0.00088
$\theta = 5$	Median	0.00136	0.00214	-0.0011
	S.Dev	0.05237	0.06892	0.04602
	Mean	-0.00258	-0.00038	0.00015
$\theta = 10$	Median	-0.00147	0.00036	0.00035
	S.Dev	0.04935	0.03641	0.03156

the sample size n which confirms Proposition 4.6, even for moderate θ, and even though $T = 40$ is not an inordinately large value. To investigate the asymptotic normal distribution of $\tilde{\rho}(n)$ empirically, we need to compare the distribution of the statistic

$$\psi(n, \theta) := \sqrt{T_n}\sqrt{\theta}\tilde{\rho}(n)$$

with the standard Gaussian distribution $\mathcal{N}(0,1)$. For this aim, we chose $\theta = 2$, $n = 100000$, $T_n = 100$ and based on 3000 replications, we obtained the following histogram:

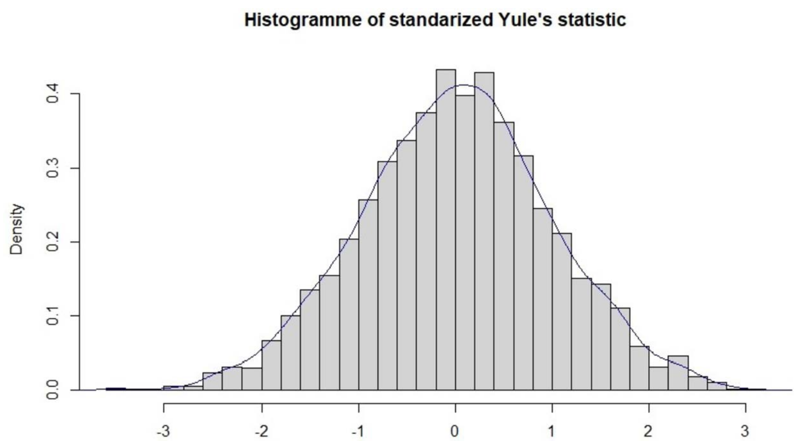


Fig 2. *Histogram of $\psi(n,\theta)$ for $n = 100000$, $T_n = 100$, $\theta = 2$*

The histogram (2) shows visually that the normal approximation of the distribution of the statistic $\psi(n,\theta)$ is reasonable for the sampling size $n = 100000$ and time horizon $T_n = 100$. Moreover, the results of the next table below show a comparison of statistical properties between $\psi(n,\theta)$ and $\mathcal{N}(0,1)$ with the same parameters as in Figure 2, we can see that the empirical mean, median and standard deviation of $\psi(n,\theta)$ and $\mathcal{N}(0,1)$ are quite close, which illustrates well our theoretical results.

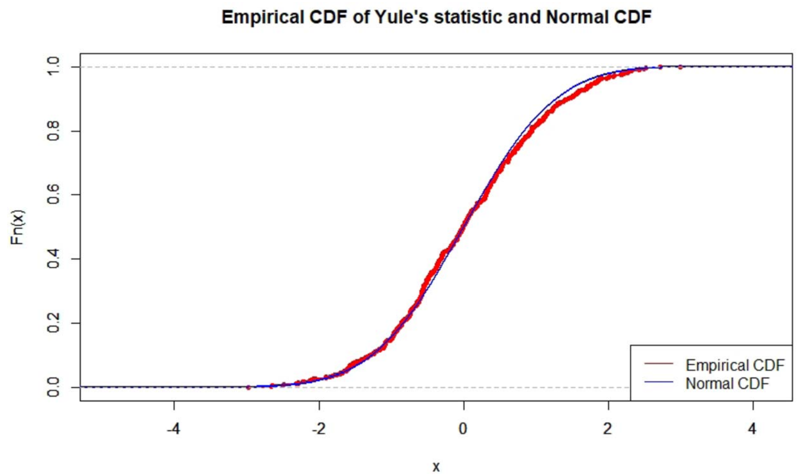
Statistics	Mean	Median	Standard Deviation
$\mathcal{N}(0,1)$	0	0	1
$\psi(n,\theta)$	0.00832	0.01206	0.99691

We can also illustrate numerically the rate of convergence in law of the statistic $\psi(n,\theta)$ to the standard Gaussian distribution, by approximately computing the Kolmogorov distance between $\psi(n,\theta)$ and $\mathcal{N}(0,1)$. For this aim, we approximate the cumulative distribution function using an empirical cumulative distribution function based on 500 replications of the computation of $\psi(n,\theta)$ for $n = 100000$. The figure below shows the empirical and standard normal cumulative distribution functions.

Based on Remark 4.10, when the mesh $\Delta_n = n^{-\lambda}$ for $\frac{1}{2} < \lambda < \frac{5}{7}$, we expect that

$$d_{Kol}\left(\sqrt{T_n}\sqrt{\theta}\tilde{\rho}(n), \mathcal{N}(0,1)\right) \leqslant c(\theta)(1-\lambda)\ln(n) \times n^{\frac{1-2\lambda}{3}}.$$

In fact, with our choice of $\lambda = 0.6$ and a sample size $n = 100000$, the time horizon $T_n = 100$. The mesh size $n^{-0.6}$ is larger than the optimal size $n^{-5/7}$, yielding a larger time horizon T_n than under the optimal observation frequency. The



Kolmogorov distance between the two laws, which equals the sup norm of the difference of these cumulative distribution functions, computes to approximately 0.01974, which implies that $c(\theta)$ is greater than 9.310^{-3}. We could have chosen the optimal mesh

$$\Delta_n = n^{-5/7},$$

yielding a rate of order $\ln(n) \times n^{-1/7}$, but in this case in order to have the same time horizon $T_n = 100$, we would have needed $n = 10^7$ data points, which is a large number. In practical applications, the cost of higher-frequency observations, if known, is to be balanced with desired precision on the Kolmogorov distance, which may well point to a lower frequency for a fixed time horizon.

Acknowledgements

We would like to thank the referee for her/his/their careful reading, constructive remarks and useful suggestions.

References

[1] Ernst, P. Shepp, L., Wyner, A. (2017). Yule's "nonsense correlation" solved! *The Annals of Statistics*, **45**(4), 1789–1809. MR3670196

[2] Ernst, P. Rogers, L.C.G., Zhou, Q. (2020). The distribution of Yule's "nonsense correlation". arXiv: 1909.02546

[3] Ernst, P. Huang, D., Viens, F. Yule's "nonsense correlation" for Gaussian random walks. (2021). arXiv: 2103.06176

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

[5] Phillips, P.C.B. (1986). Understanding spurious regressions in econometrics. *J. Econometrics* **33** 311–340. MR0867979

[6] Nourdin, I., Peccati, G. (2009). Stein's method and exact Berry Esseen asymptotics for functionals of Gaussian fields. *The Annals of Probability* **37**(6), 2231–2261. MR2573557

[7] Nourdin, I., Peccati, G. (2015). The optimal fourth moment theorem. *Proc. Amer. Math. Soc.* **143** 3123–3133. MR3336636

[8] Nourdin, I., Peccati, G. (2012). *Normal approximations with Malliavin calculus : from Stein's method to universality*. Cambridge Tracts in Mathematics 192. Cambridge University Press, Cambridge. MR2962301

[9] Nourdin, I., Viens, F. (2009). Density Formula and Concentration Inequalities with Malliavin Calculus. *Electron. J. Probab.* **14** 2287–2309. MR2556018

[10] Nualart, D. (2006). *The Malliavin calculus and related topics*. Springer-Verlag, Berlin. MR2200233

[11] Michel, R., Pfanzagl, J. (1971). The accuracy of the normal approximation for minimum contrast estimates. *Z. Wahrscheinlichkeitstheorie verw Gebiete* **18** 73–84. MR0288897

[12] Yule, G.U. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society* **89**(1), 1–63. https://doi.org/10.2307/2341482