



Parameter estimation of Ornstein-Uhlenbeck process generating a stochastic graph

Emmanuel Gobet, Gustaw Matulewicz

► To cite this version:

Emmanuel Gobet, Gustaw Matulewicz. Parameter estimation of Ornstein-Uhlenbeck process generating a stochastic graph. *Statistical Inference for Stochastic Processes*, 2016, 20 (2), pp.211-235. 10.1007/s11203-016-9142-4 . hal-01271994

HAL Id: hal-01271994

<https://polytechnique.hal.science/hal-01271994>

Submitted on 9 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Parameter estimation of Ornstein-Uhlenbeck process generating a stochastic graph

Emmanuel Gobet*, Gustaw Matulewicz†

February 9, 2016

Abstract

Given Y a graph process defined by an incomplete information observation of a multivariate Ornstein-Uhlenbeck process X , we investigate whether we can estimate the parameters of X . We define two statistics of Y . We prove convergence properties and show how these can be used for parameter inference. Finally, numerical tests illustrate our results and indicate possible extensions and applications.

Keywords: stochastic graph process, inference for stochastic process, incomplete information, asymptotic properties of estimators.

MSC 2010: 62Mxx; 05C80; 62F12.

1 Introduction

1.1 Statement of the problem

Take an Ornstein-Uhlenbeck process $X = (X_t : t \geq 0)$ with values in \mathbb{R}^d ($d \in \mathbb{N}^+$), solution to the stochastic differential equation:

$$dX_t = -AX_t dt + \Sigma dW_t, \quad X_0 \text{ given.} \quad (1.1)$$

We consider the model of stochastic graphs generated as follows: the adjacency value between vertices i and j is

$$Y_t^{ij} = \mathbb{1}_{X_t \in S^{ij}}$$

where $(S^{ij})_{i,j}$ are subsets of \mathbb{R}^d .

The topic of random graphs is a well-developed research area. Since the Erdős–Rényi model, many other ways of generating a random graph have been proposed, most notably the preferential attachment model, the Chung-Lu model or the Kronecker graph model [Bol01, MX07]. Most models have the goal to create a single instance of a random graph.

*CMAP, Ecole Polytechnique and CNRS, Université Paris Saclay, Route de Saclay, 91128 Palaiseau cedex, France. Email: emmanuel.gobet@polytechnique.edu.

†CMAP, Ecole Polytechnique and CNRS, Université Paris Saclay, Route de Saclay, 91128 Palaiseau cedex, France. Email: gustaw.matulewicz@polytechnique.edu. This work was funded jointly by *Chaire Risques Financiers* of the *Risk Fondation*, the *Finance for Energy Market Research Centre* and the Natixis Foundation for Quantitative Research.

Some proceed by "growing" the graph, i.e. by successively adding nodes and edges, as in the preferential attachment model. Other models also enable deleting nodes and edges.

In contrast, in our model the nodes are fixed, hence it is not a "growing" graph, but the edges are evolving continuously in time. One could in principle fix a T and consider the random graph Y_T , but the real richness of our model resides in the evolution of the graph in continuous time. For instance, it adds correlation between graphs at arbitrary time-scales.

Y gives only partial information about X . Hence, usual results on inference for stochastic processes can not be applied. We therefore aim at extending these results to our setting and ask then the question of finding A, Σ from Equation (1.1), given the sole observation of Y .

We will consider that we have access to one realization of the process Y observed at discrete times $(k\Delta_n)_{0 \leq k \leq n}$, $n \geq 1$. Therefore, we hope to get results in the long time limit $n\Delta_n \rightarrow +\infty$ as $n \rightarrow +\infty$, in which we can expect to use ergodic properties of the process X . Intuitively, doing so we will estimate parameters arising in the stationary distribution. Also, to estimate parameters related to local fluctuations (i.e. Σ), we are interested in the high frequency limit $\Delta_n \rightarrow 0$.

1.2 Applications in systemic risk modeling

In [CFS15, FI13], authors present a model for inter-bank lending in which d bank reserves are modeled through real-valued random processes X^i . Whenever bank i has more reserves than bank j , i lends money to j , thus reducing reserve X^i and increasing X^j . Gaussian noise is added in order to model random variations of the reserves. We will refer to this model through Equation (1.2) (which is exactly the model from [CFS15] without a central bank):

$$dX_t^i = -\frac{a}{d} \sum_{j=1}^d (X_t^i - X_t^j) dt + \sigma dW_t^i. \quad (1.2)$$

In this model, authors define a systemic event when the mean reserve falls below some predetermined value. The analysis of the model shows that the probability of the systemic event can be computed and the result depends explicitly on the values of the parameters, especially on the correlation of the Brownian motions with a common noise.

Therefore, it is crucial to know how to estimate the values of the parameters of the equation, even in the realistic situation where one wouldn't have complete information of the banks' reserves. We consider for instance that the regulator, whose perspective we are analysing here, would fix a regulatory threshold r and would observe all variables of the form

$$Y_t^{ij} := \mathbb{1}_{X_t^i - X_t^j > r}.$$

1.3 Summary of results

In this article, we define two statistics and show the following convergences.

- An "occupation time" statistic OT_n , which counts the number of times the process is present in a given set. Normalized by n , this number converges to the stationary measure of that set, in the following ways:
 - in L^2 , with a speed of convergence bounded by $\sqrt{n\Delta_n}$, which is the square root of the time horizon of the estimation;
 - with the right normalization, convergence in law to a Gaussian variable, for one-dimensional processes.
- A "crossings" statistic \mathcal{C}_n , which counts the number of times the process X goes in or out of a given set S^{ij} , i.e. the number of changes of the $Y^{i,j}$ -value. We show a convergence in L^2 of this - suitably normalized - statistic.

We show then how to use this in order to estimate the parameters of model (1.1).

1.4 Related work

The question that we investigate in this article is on the recovery of the parameters of an Ornstein-Uhlenbeck equation given the n observations of a single realization of the process. It relates therefore to the largely developed field of inference for stochastic processes. Many results exist on this subject: for instance, see [Kut04] for continuous-time observations and [KLS12] for discrete-time ones.

In this work, we are specifically interested in discrete-time observation schemes. We observe three distinct discrete-time settings. First, the low-frequency long-time (LF-LT) setting consists in fixing a time step Δ and observing at times $(i\Delta)_{i \leq n}$ with $n \rightarrow +\infty$ [Yos92]. Second, the high-frequency fixed-time (HF-FT) setting, where a time horizon T is fixed and observations are taken at $(i\Delta_n)_{i \leq n}$ with $\Delta_n = T/n \rightarrow 0$ [GJ93]. Third, the high-frequency long-time (HF-LT) setting where one assumes observations at $(i\Delta_n)_{i \leq n}$ with the time step $\Delta_n \rightarrow 0$ and the time horizon $n\Delta_n \rightarrow +\infty$ [Kes97, Gob02, ASM04]. Our work is placed in the latter HF-LT setting.

Some results already exist on problems with observation of crossings of a given threshold. For instance, [Flo87] considers the estimation using only the observation of the sign of the process. However this is done in the LF-LT setting and in dimension 1. The same remark applies to [Flo89, Flo91]. We will extend her CLT results to the HF limit.

The HF-LT setting is combined with partial information observation in [IUY09]. The authors consider, for $\epsilon > 0$, the observation of $\mathbb{1}_{|X_t| \geq \epsilon} X_t$. Our assumption of a binary observation leaves us with even scarcer information, thus making the inference problem more delicate.

To sum up, the closest work to ours is seemingly [IUY09] and [Flo87] but our main original contribution concerns the multi-dimensional scope and the case of binary observation in the HF-LT setting.

Organisation of the paper. In the next subsection, we define the notations and assumptions used throughout this work. Then in Section 2, we study the first statistic based on occupation time. We first prove general convergence results useful to analyse the convergence of all estimators (Theorem 2.1). L^2 and CLT results are proved (Theorems 2.2 and 2.3). Section 3 is devoted to the study of the second estimator based on crossings

(Theorem 3.1). Last in Section 4, we report numerical experiments. Applications to parameters inference are discussed along the sections. Technical results are postponed to Appendix.

1.5 Notations and assumptions

1.5.1 Notations

For $d \in \mathbb{N}^+$, take $m \in \mathbb{R}^d$ and V a symmetric positive definite $d \times d$ matrix. We call $\mathcal{N}(m, V)$ the law of a Gaussian r.v. with mean m and covariance matrix V . For $m = 0$, this centered Gaussian distribution is denoted by ν_V and its density by μ_V :

$$\mu_V(x) = (2\pi)^{-d/2} \det(V)^{-1/2} \exp\left(-\frac{1}{2}x^*V^{-1}x\right), \quad x \in \mathbb{R}^d,$$

where x^* is the transpose of x . In dimension $d = 1$, we introduce additionally the CDF of $\mathcal{N}(0, 1)$:

$$N(x) = \int_{-\infty}^x \mu_1(s) ds, \quad x \in \mathbb{R}.$$

Given a measurable function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and a probability measure ν on \mathbb{R}^d , we denote $\nu(f) = \int f(x) \nu(dx)$. For a measurable set $S \subset \mathbb{R}^d$, we write $\nu(S) = \nu(\mathbb{1}_S)$ by a slight abuse of notation.

1.5.2 Restatement of the model and standing assumptions

Consider two matrices $A \in \mathcal{M}_{d,d}(\mathbb{R})$ and $\Sigma \in \mathcal{M}_{d,q}(\mathbb{R})$ where $d, q \in \mathbb{N}^+$, which serves to model (1.1). The standing assumptions on A and Σ are the following.

(H) The matrix $\Sigma\Sigma^*$ is invertible and the spectrum of A has strictly positive real parts:

$$a_0 := \min_{\lambda \in \text{Sp}(A)} \text{Re}(\lambda) > 0. \quad (1.3)$$

We define an important class of covariance matrices:

$$V_t = \int_0^t e^{-Au} \Sigma \Sigma^* e^{-A^*u} du, \quad V_\infty = \int_0^{+\infty} e^{-Au} \Sigma \Sigma^* e^{-A^*u} du.$$

We easily check that V_∞ is well defined, symmetric positive definite. For one-dimensional processes, we simply have:

$$v_t = \frac{\sigma^2}{2a} (1 - e^{-2at}), \quad v_\infty = \frac{\sigma^2}{2a}.$$

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}^+}, \mathbb{P})$ be a filtered space and $(W_t)_{t \in \mathbb{R}^+}$ a q -dimensional Brownian motion with respect to \mathcal{F} . In this setting, we consider the multi-dimensional Ornstein-Uhlenbeck equation for X as introduced in (1.1):

$$dX_t = -AX_t + \Sigma dW_t, \quad X_0 \stackrel{d}{=} \mathcal{N}(0, V_\infty), \quad (1.4)$$

where X_0 is a r.v. independent of W . In the following $X = (X_t : t \geq 0)$ stands for the \mathbb{R}^d -valued solution of (1.4). We recall some properties from [KS91, Chapter 5.6]. First X is stationary:

$$\forall t \in \mathbb{R}^+, X_t \stackrel{d}{=} \mathcal{N}(0, V_\infty). \quad (1.5)$$

To simplify we denote by ν_∞ the Gaussian distribution $\mathcal{N}(0, V_\infty)$ and by μ_∞ its density. In the subsequent analysis, the initial distribution could be different from ν_∞ , it would not change significantly the analysis since the OU-process converges exponentially fast to its stationary regime.

Second, X is Markovian and ergodic. Take $t > s$, we can write:

$$X_t = e^{-A(t-s)} X_s + \int_s^t e^{-A(t-u)} \Sigma dW_u, \quad (1.6)$$

from which we deduce

$$X_t \mid X_s \stackrel{d}{=} \mathcal{N}\left(e^{-A(t-s)} X_s, V_{t-s}\right). \quad (1.7)$$

Equality (1.6) gives also an important insight on decorrelation of the process:

$$\text{Cov}(X_t, X_s) = e^{-A(t-s)} \text{Var}(X_s) = e^{-A(t-s)} V_\infty, \quad t \geq s. \quad (1.8)$$

In the following, all the limits will be considered as $n \rightarrow +\infty$, under the asymptotics of high frequency data ($\Delta_n \rightarrow 0$) on a long-time interval ($n\Delta_n \rightarrow +\infty$). Also, for simplicity, we assume $\Delta_n \leq 1$.

Remark 1.1. In Equations (1.4) and (1.7), we see that the distribution of X_0 and $X_t \mid X_s$ depend on Σ only through $\Sigma\Sigma^*$. Hence we shall restate our inference problem as the estimation of $(A, \Sigma\Sigma^*)$.

2 Occupation time statistic

Consider the d -dimensional process governed by Equation (1.4). We define the first statistic:

Definition 2.1. Let S be a measurable subset of \mathbb{R}^d . Define:

$$Y_t^S = \mathbb{1}_{X_t \in S}.$$

The occupation time statistic is defined as:

$$\text{OT}_n^S = \frac{1}{n} \sum_{k=0}^{n-1} Y_{k\Delta_n}^S = \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{X_{k\Delta_n} \in S}.$$

This statistic gives the frequency of occupation of S by the process X , hence the name.

2.1 Preliminary tools

The study of convergence of OT_n^S and further statistics will be made possible by using some tight controls related to the mixing properties of X at different times. Note that we cannot directly invoke general mixing properties of Markov chains since here the Markov chain $(X_{k\Delta_n} : k \geq 0)$ depends on n through Δ_n : this is the main difficulty. The current estimates are made possible using the Gebelein inequality (a.k.a. Lancaster inequality) about maximal correlation between Gaussian spaces.

Theorem 2.1 (Mixing properties). *Assume that X solves the Ornstein-Uhlenbeck Equation (1.4), and recall the definition of a_0 in (1.3). There exists a finite constant $C_{(2.1)}$, depending only on the stationary distribution covariance matrix V_∞ , such that for any $t \geq s \geq 0$ and for any functions $\varphi : \mathcal{C}^0([0, s], \mathbb{R}^d) \rightarrow \mathbb{R}$, $\phi : \mathcal{C}^0([t, +\infty), \mathbb{R}^d) \rightarrow \mathbb{R}$ such that φ, ϕ are square-integrable w.r.t. the law of X , we have*

$$|\text{Cov}(\varphi((X_u)_{u \leq s}), \phi((X_v)_{v \geq t}))| \leq C_{(2.1)} e^{-a_0|t-s|} \sqrt{\text{Var}(\varphi((X_u)_{u \leq s})) \text{Var}(\phi((X_v)_{v \geq t}))}. \quad (2.1)$$

The proof is done in Appendix B. A very useful corollary is related to the convergence study of sum of general local functionals of X .

Corollary 2.1. *Consider a measurable function $g : \mathbb{N} \times \mathbb{N} \times \mathcal{C}^0([0, 1], \mathbb{R}^d) \rightarrow \mathbb{R}$ such that $\mathbb{E}[g(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n})^2] < +\infty$ for any $k, n \in \mathbb{N}$. For $n \in \mathbb{N}$ define*

$$v_n^2 = \sup_{k < n} \text{Var}(g(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n})),$$

$$\xi_k^{(n)} = \sqrt{\frac{\Delta_n}{n}} g(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n}).$$

Then, there is a finite constant $C_{(2.2)}$, dependent only on the parameters A, Σ of the model, such that:

$$\text{Var}\left(\sum_{k=0}^{n-1} \xi_k^{(n)}\right) \leq C_{(2.2)} v_n^2. \quad (2.2)$$

Remark 2.1. *If $\mathbb{E}\left[\sum_{k=0}^{n-1} \xi_k^{(n)}\right] \rightarrow l$ for some $l \in \mathbb{R}$, then $v_n \rightarrow 0$ implies*

$$\sum_{k=0}^{n-1} \xi_k^{(n)} \xrightarrow{L^2} l.$$

Proof of Corollary 2.1. Denote $g_k = g(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n})$; without loss of generality, we can assume that $\mathbb{E}[g_k] = 0$. We have

$$\text{Var}\left(\sum_{k=0}^{n-1} \xi_k^{(n)}\right) = \frac{\Delta_n}{n} \sum_{k=0}^{n-1} \text{Var}(g_k) + \frac{2\Delta_n}{n} \sum_{k=0}^{n-1} \sum_{l=k+1}^{n-1} \text{Cov}(g_k, g_l).$$

For $l > k$, we have $[k\Delta_n, (k+1)\Delta_n] \subset [0, (k+1)\Delta_n]$ and $[l\Delta_n, (l+1)\Delta_n] \subset [l\Delta_n, +\infty[$. Apply Theorem 2.1:

$$\text{Cov}(g_k, g_l) \leq C_{(2.1)} e^{-a_0|k+1-l|\Delta_n} \sqrt{\text{Var}(g_k) \text{Var}(g_l)}.$$

Then we deduce

$$\begin{aligned} \mathbb{V}\text{ar} \left(\sum_{k=0}^{n-1} \xi_k^{(n)} \right) &\leq \frac{\Delta_n}{n} n v_n^2 + \frac{2\Delta_n}{n} n \sum_{m \geq 0} C_{(2.1)} v_n^2 e^{-a_0 m \Delta_n} \\ &\leq v_n^2 \left(\Delta_n + 2C_{(2.1)} \frac{\Delta_n}{1 - e^{-a_0 \Delta_n}} \right) \\ &\leq C_{(2.2)} v_n^2, \end{aligned}$$

where we set $C_{(2.2)} = \sup_{x \in [0,1]} \left(x + 2C_{(2.1)} \frac{x}{1 - e^{-a_0 x}} \right)$. □

2.2 L^2 convergence of occupation time statistics

Theorem 2.2. *For any measurable set $S \subset \mathbb{R}^d$, OT_n^S converges to $\nu_\infty(S)$ in L^2 and*

$$\mathbb{E} [(\text{OT}_n^S - \nu_\infty(S))^2] = O \left(\frac{1}{n\Delta_n} \right).$$

Proof. As the process is stationary, $\mathbb{E} [\text{OT}_n^S] = \mathbb{E} [\mathbb{1}_{X_0 \in S}] = \nu_\infty(S)$. Next, we apply Corollary 2.1 to $\text{OT}_n^S = \sum_{k=0}^{n-1} \xi_k^{(n)}$ with

$$\begin{aligned} \xi_k^{(n)} &= \frac{1}{n} \mathbb{1}_{X_{k\Delta_n} \in S} = \sqrt{\frac{\Delta_n}{n}} g(k, n, X_{k\Delta_n}), \\ g(k, n, X_{k\Delta_n}) &= \frac{1}{\sqrt{n\Delta_n}} \mathbb{1}_{X_{k\Delta_n} \in S}, \\ \mathbb{V}\text{ar}(g(k, n, X_{k\Delta_n})) &= \frac{1}{n\Delta_n} \mathbb{V}\text{ar}(\mathbb{1}_{X_{k\Delta_n} \in S}) = \frac{\nu_\infty(S)(1 - \nu_\infty(S))}{n\Delta_n}. \end{aligned}$$

Therefore, we get $\mathbb{E} [(\text{OT}_n^S - \nu_\infty(S))^2] = \mathbb{V}\text{ar}(\text{OT}_n^S) \leq \frac{C_{(2.2)}}{n\Delta_n}$. □

2.3 Central Limit Theorem for one-dimensional processes

Here we restrict the study to the one-dimensional situation. There are two technical reasons for this: we solve explicitly the Poisson equation (see Lemma C.2) and derive tractable bounds on it. Additionally, we take advantage of the one-dimensional situation to handle explicit computations. The validity of a Central Limit Theorem in the multi-dimensional setting remains an open question to us.

For $d = 1$, the model becomes

$$dX_t = -aX_t dt + \sigma dW_t. \tag{2.3}$$

Assumption **(H)** reads $a > 0$ and $\sigma \neq 0$. We consider the case $S = [1, +\infty[$. The extension of the following results to the case where S is a finite union of intervals is straightforward, and it is left to the reader.

Theorem 2.3. *As $n \rightarrow +\infty$, we have*

$$\sqrt{n\Delta_n} \left(\text{OT}_n^{[1,+\infty[} - \nu_\infty([1,+\infty[) \right) \xrightarrow{d} \mathcal{N}(0, \nu_\infty(\sigma^2 F'^2))$$

where F is defined in (C.2) and is such that

$$F'(x) = \frac{2}{\sigma^2} \frac{N\left(\frac{x \wedge 1}{\sqrt{\nu_\infty}}\right) - N\left(\frac{1}{\sqrt{\nu_\infty}}\right) N\left(\frac{x}{\sqrt{\nu_\infty}}\right)}{\mu_\infty(x)} \in \left[0, 2\sqrt{\frac{\pi}{a\sigma^2}}\right].$$

Proof. A simple inspection on F' shows that it is non-negative. The upper bound is proved in Lemma C.3 (see inequality (C.7)). This proves the inclusion of $F'(x)$.

We now prove the Central Limit Theorem. We follow the approach by [Flo84]. The main difference is that the function $x \mapsto \mathbb{1}_{x \geq 1}$ is non continuous, which raises technical issues.

Consider first the continuous time extension of $\text{OT}_n^{[1,+\infty[}$, i.e.

$$\text{OT}_t^c = \frac{1}{t} \int_0^t \mathbb{1}_{X_s \geq 1} ds.$$

Denote in this proof $f(x) = \mathbb{1}_{x \geq 1}$ and $\hat{f}(x) = f(x) - \nu_\infty([1,+\infty[)$, so that

$$\int_0^t \hat{f}(X_s) ds = t(\text{OT}_t^c - \nu_\infty([1,+\infty[)).$$

Introduce then $L = -ax \frac{\partial}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2}$ the infinitesimal generator of X : Lemma C.2 in Appendix C ensures that F defined in (C.2) verifies the Poisson equation

$$LF = -\hat{f}.$$

Introduce $M_t = F(X_t) - F(X_0) + \int_0^t \hat{f}(X_s) ds$. F is twice differentiable but F'' has a single point of discontinuity at 1. However, we can still apply Itô's formula in that case (see Lemma C.1). We get:

$$M_t = \int_0^t \sigma F'(X_s) dW_s, \quad \langle M \rangle_t = \int_0^t \sigma^2 F'(X_s)^2 ds.$$

F' being bounded, M is a martingale. As we have $t^{-1} \langle M \rangle_t \rightarrow \nu_\infty(\sigma^2 F'^2)$ in probability (ergodic theorem) as $t \rightarrow +\infty$, we can use a CLT for martingales (see Lemma C.4 with $K_t = t^{-1/2}$) to get

$$\frac{M_t}{\sqrt{t}} = \frac{F(X_t) - F(X_0) + (\text{OT}_t^c - \nu_\infty([1,+\infty[)) t}{\sqrt{t}} \xrightarrow{d} \mathcal{N}(0, \nu_\infty(\sigma^2 F'^2)).$$

Finally, F is sublinear (F' bounded), thus $\frac{1}{\sqrt{t}} (F(X_t) - F(X_0)) \xrightarrow{L^2} 0$. Consequently we have proved

$$\sqrt{t}(\text{OT}_t^c - \nu_\infty([1,+\infty[)) \xrightarrow{d} \mathcal{N}(0, \nu_\infty(\sigma^2 F'^2)).$$

We now aim at proving that the above result extends to the discrete version $\text{OT}_n^{[1, +\infty[}$. For this, define

$$\begin{aligned} D_n &:= \sqrt{n\Delta_n} \left(\text{OT}_n^{[1, +\infty[} - \text{OT}_{n\Delta_n}^c \right) \\ &= \sqrt{\frac{\Delta_n}{n}} \sum_{k=0}^{n-1} \int_0^{\Delta_n} \frac{f(X_{k\Delta_n}) - f(X_{k\Delta_n+u})}{\Delta_n} du \\ &:= \sqrt{\frac{\Delta_n}{n}} \sum_{k=0}^{n-1} g(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n}). \end{aligned}$$

Observe that it remains to prove that $D_n \xrightarrow{\mathbb{P}} 0$. In view of Corollary 2.1 and since

$$\mathbb{E} [g(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n})] = 0,$$

it is enough to prove that

$$v_n^2 := \sup_{k < n} \mathbb{E} [g(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n})^2] \rightarrow 0.$$

Actually, by Jensen inequality, the stationarity property and since f takes values in $\{0, 1\}$, we have

$$v_n^2 \leq \frac{1}{\Delta_n} \int_0^{\Delta_n} \mathbb{E} [|f(X_0) - f(X_u)|^2] du = \int_0^1 \mathbb{E} [|f(X_0) - f(X_{t\Delta_n})|] dt.$$

With probability 1, $f(X_{t\Delta_n}) \rightarrow f(X_0)$, since f is continuous except on a set of zero ν_∞ -measure: by the dominated convergence theorem, we obtain $v_n \rightarrow 0$, then $D_n \xrightarrow{\mathbb{P}} 0$. From this we have:

$$\sqrt{n\Delta_n} \left(\text{OT}_n^{[1, +\infty[} - \nu_\infty([1, +\infty[) \right) \xrightarrow{d} \mathcal{N}(0, \nu_\infty(\sigma^2 F'^2)).$$

□

2.4 Application to parameter inference

Lemma 2.1. *Fix $S \subset \mathbb{R}^d$ and recall (H). Then $\nu_\infty(S)$ is a continuous function of V_∞ .*

Proof. Write

$$\nu_\infty(S) = \int_S \mu_\infty(x) dx = \int_S (2\pi)^{-d/2} \det(V_\infty)^{-1/2} \exp\left(-\frac{1}{2} x^* V_\infty^{-1} x\right) dx.$$

As the determinant and the inverse are continuous functions, $\mu_\infty(x)$ is continuous in V_∞ for any x . We also have:

$$\mu_\infty(x) \leq (2\pi)^{-d/2} v_m^{-d/2} \exp\left(-\frac{v_M^{-1}}{2} |x|^2\right)$$

where $v_M = \max_{\lambda \in \text{Sp}(V_\infty)} \lambda$, $v_m = \min_{\lambda \in \text{Sp}(V_\infty)} \lambda$.

Applying the Hoffman-Wielandt theorem ([HJ86, Theorem 6.3.5]) we know that v_M and v_m are continuous functions of V_∞ , which are also non-zero in the neighborhood of invertible V_∞ . From this follows that there is a local bound (in the neighbourhood of every invertible V_∞) by an integrable function of the form:

$$\mu_\infty(x) \leq \text{Cst} \exp(-\text{Cst}|x|^2)$$

with a positive constant Cst. Conclude using the dominated convergence theorem. \square

For one-dimensional processes. We consider here Equation (2.3). The limit value of OT_n depends on the stationary distribution of the process, which is a centered Gaussian r.v. with variance $v_\infty = \sigma^2/2a$ (see Section 1.5). If $\nu_\infty(S)$ is monotonous with respect to v_∞ , then we can construct an estimator of v_∞ .

For instance, if $S = [1, +\infty[$, then $\nu_\infty(S) = N(-1/\sqrt{v_\infty})$ which is strictly increasing with v_∞ .

However, v_∞ is not a one-to-one function of a and σ hence we need more information to find the parameters of the process, using for instance the crossings statistic of Section 3.

For multi-dimensional processes. Here again the limit value of OT_n^S is the measure of S under the stationary distribution. This distribution depends on the value of the matrix V_∞ . Without further information or assumptions, V_∞ is a symmetric $d \times d$ matrix, representing $d(d+1)/2$ unknowns. We can expect to be able to find these unknowns only if we consider more than one set S and the corresponding statistics.

In the following, we will use the fact that the covariance matrices of the marginals of a Gaussian variable are the restrictions of its covariance matrix to the relevant spaces.

Consider first for $i \leq d$ the set $S^i = \{x : x^i \geq 1\}$. Then $\nu_\infty(S^i)$ depends only on the value of $(V_\infty)^{ii}$. Applying the result from the preceding paragraph, we can construct an estimator of that value.

Consider then for $i \neq j$ the set $S^{ij} = \{x : x^i \geq 1, x^j \geq 1\}$. Then $\nu_\infty(S^{ij})$ depends only on the values of $(V_\infty)^{ii}, (V_\infty)^{jj}, (V_\infty)^{ij}$. From the previous point, we know we can construct estimators of $(V_\infty)^{ii}, (V_\infty)^{jj}$. For the last parameter, we use the following result.

Proposition 2.1. *Take (G_1, G_2) a non-degenerate centered Gaussian vector. Denote ρ the correlation between G_1 and G_2 . Denote $S = [1, +\infty[^2$ and μ_ρ the density of the distribution of (G_1, G_2) . Then $\mu_\rho(S)$ is a strictly increasing function of ρ .*

Proof. Denote $\sigma_1 = \sqrt{\text{Var}(G_1)}$, $\sigma_2 = \sqrt{\text{Var}(G_2)}$. By symmetry between G_1 and G_2 , we can safely assume $\sigma_1 \geq \sigma_2$. Introduce a standard centered Gaussian G . We can write

$$\begin{aligned} (G_1, G_2) &\stackrel{d}{=} \left(\rho \frac{\sigma_1}{\sigma_2} G_2 + \sigma_1 \sqrt{1 - \rho^2} G, G_2 \right), \\ \mu_\rho(S) &= \mathbb{E} \left[\mathbb{1}_{G_2 \geq 1} \mathbb{E} \left[\mathbb{1}_{\rho \frac{\sigma_1}{\sigma_2} G_2 + \sigma_1 \sqrt{1 - \rho^2} G \geq 1} \mid G_2 \right] \right] \\ &= \int_{\mathbb{R}} \mathbb{1}_{y \geq 1} N \left(\frac{\rho \frac{\sigma_1}{\sigma_2} y - 1}{\sigma_1 \sqrt{1 - \rho^2}} \right) \mu_{\sigma_2^2}(y) dy. \end{aligned}$$

Set $g(\rho, y) = \frac{\rho \frac{\sigma_1}{\sigma_2} y - 1}{\sigma_1 \sqrt{1 - \rho^2}}$: then $\frac{dg}{d\rho}(\rho, y) = \frac{\frac{\sigma_1}{\sigma_2} y - \rho}{\sigma_1 (1 - \rho^2)^{3/2}}$. This is strictly positive for $y \geq 1$ since $\frac{\sigma_1}{\sigma_2} \geq 1$ and $\rho < 1$. Therefore $N(g(\rho, y))$ is strictly increasing in ρ , and so is $\mu_\rho(S)$ on $] -1, 1[$.

□

This shows that we can construct an estimator of $(V_\infty)^{ij}$ given the knowledge of $(V_\infty)^{ii}, (V_\infty)^{jj}$, which we have as noted before. Therefore, using $d(d+1)/2$ estimators, we can recover the whole matrix V_∞ .

3 Crossings statistic

Given a binary observation Y_t , a function of X_t , we will count how many times Y goes from 0 to 1 and vice-versa. The following defines a statistic counting the number of jumps between 0 and 1 of the discretization of Y .

Definition 3.1. *We define the crossings statistic by:*

$$\mathcal{C}_n^S = \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{Y_{k\Delta_n}^S \neq Y_{(k+1)\Delta_n}^S}.$$

In the following, we restrict the convergence analysis to sets which are half-spaces $S^i = \{x : x^i \geq 1\}$; by symmetry, we assume that $i = 1$. Therefore we drop the superscript S and

$$Y_t = \mathbb{1}_{X_t^1 \geq 1}.$$

\mathcal{C}_n counts the number of times the discretised projection of X on the first coordinate crosses 1.

3.1 L^2 convergence

Theorem 3.1. *Assume that $n\Delta_n^{3/2} \rightarrow +\infty$. We have the following convergence:*

$$\begin{aligned} \frac{\mathcal{C}_n}{\sqrt{\Delta_n}} &\xrightarrow{L^2} 2\sqrt{\frac{(\Sigma\Sigma^*)^{11}}{2\pi}} \mu_{V_\infty^{11}}(1), \\ \text{Var}\left(\frac{\mathcal{C}_n}{\sqrt{\Delta_n}}\right) &= O\left(\frac{1}{n\Delta_n^{3/2}}\right). \end{aligned}$$

Proof. For ease of writing, introduce a new notation:

$$Z_k^{(n)} = \mathbb{1}_{X_{k\Delta_n}^1 < 1} \mathbb{1}_{X_{(k+1)\Delta_n}^1 \geq 1}.$$

Now, divide the sum in two similar parts:

$$\begin{aligned} \mathcal{C}_n &= \mathcal{C}_n^{+-} + \mathcal{C}_n^{-+}, \\ \mathcal{C}_n^{+-} &= \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{X_{k\Delta_n}^1 \geq 1} \mathbb{1}_{X_{(k+1)\Delta_n}^1 < 1}, \quad \mathcal{C}_n^{-+} = \frac{1}{n} \sum_{k=0}^{n-1} Z_k^{(n)}. \end{aligned}$$

Although the two sums aren't perfectly symmetric, we will show that our reasoning will apply to both sums. Concentrate then on the second sum. In order to apply Corollary 2.1, we introduce the following notations:

$$\begin{aligned}\frac{\mathcal{C}_n^{-+}}{\sqrt{\Delta_n}} &:= \sum_{k=0}^{n-1} \xi_k^{(n)}, \\ \xi_k^{(n)} &= \frac{1}{n\sqrt{\Delta_n}} Z_k^{(n)} := \sqrt{\frac{\Delta_n}{n}} g\left(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n}\right), \\ v_n^2 &:= \sup_{k < n} \text{Var}\left(g\left(k, n, (X_s)_{k\Delta_n \leq s \leq (k+1)\Delta_n}\right)\right).\end{aligned}$$

We obviously have $v_n^2 = \frac{1}{n\Delta_n^2} \sup_{k < n} \text{Var}\left(Z_k^{(n)}\right)$ and Corollary A.1 we have $\text{Var}\left(Z_k^{(n)}\right) = O(\sqrt{\Delta_n})$. Therefore $v_n^2 = O\left(\frac{1}{n\Delta_n^{3/2}}\right) = o(1)$ by assumption. Thus, Corollary 2.1 gives us the convergence to 0 of the variance.

Using again Corollary A.1 we get

$$\mathbb{E}\left[\frac{\mathcal{C}_n^{-+}}{\sqrt{\Delta_n}}\right] = \frac{\mathbb{E}\left[Z_k^{(n)}\right]}{\sqrt{\Delta_n}} \sim \sqrt{\frac{(\Sigma\Sigma^*)^{11}}{2\pi}} \mu_{V_\infty^{11}}(1).$$

Therefore, recalling Remark 2.1 gives

$$\frac{\mathcal{C}_n^{-+}}{\sqrt{\Delta_n}} \xrightarrow{L^2} \sqrt{\frac{(\Sigma\Sigma^*)^{11}}{2\pi}} \mu_{V_\infty^{11}}(1).$$

The same reasoning would apply to \mathcal{C}_n^{+-} if only we replaced $Z_k^{(n)}$ by $\tilde{Z}_k^{(n)} = \mathbb{1}_{X_{k\Delta_n}^1 \geq 1} \mathbb{1}_{X_{(k+1)\Delta_n}^1 < 1}$. Observe

$$\begin{aligned}\mathbb{E}\left[\tilde{Z}_{\Delta_n}\right] &= \mathbb{E}\left[\left(1 - \mathbb{1}_{X_0^1 < 1}\right) \left(1 - \mathbb{1}_{X_{\Delta_n}^1 \geq 1}\right)\right] \\ &= 1 - \mathbb{E}\left[\mathbb{1}_{X_0^1 < 1}\right] - \mathbb{E}\left[\mathbb{1}_{X_{\Delta_n}^1 \geq 1}\right] + \mathbb{E}\left[Z_{\Delta_n}\right] \\ &= \mathbb{E}\left[Z_{\Delta_n}\right]\end{aligned}$$

using stationarity. Thus, we can transfer the estimate on $Z_k^{(n)}$ to $\tilde{Z}_k^{(n)}$, i.e. on \mathcal{C}_n^{-+} to \mathcal{C}_n^{+-} , which gives the final result. \square

3.2 Application to parameter inference

For one-dimensional processes. For $d = 1$, Theorem 3.1 simplifies to

$$\frac{\mathcal{C}_n}{\sqrt{\Delta_n}} \xrightarrow{L^2} \frac{2\sigma}{\sqrt{2\pi}} \mu_\infty(1).$$

Thus the renormalized \mathcal{C}_n converges to a value that depends on σ and the density $\mu_\infty(1)$, which itself depends on $v_\infty = \sigma^2/2a$. In Section 2.4, we show how to estimate the value of v_∞ . Using this estimate, we can compute $\mu_\infty(1)$ and construct simply an estimator of σ . Finally, as we estimate both σ and μ_∞ , we also estimate a . Therefore, the two statistics OT_n and \mathcal{C}_n are sufficient to estimate the parameters of the model (2.3).

For multi-dimensional processes. As we show above, we can estimate parameters of one-dimensional processes. The extension to multi-dimensional processes is not obvious. However, in specific cases, we can leverage Theorem 3.1. Consider specifically that A is a diagonal matrix with diagonal terms (a_1, \dots, a_d) . Then

$$(V_\infty)^{ij} = \int_0^{+\infty} e^{-a_i u} (\Sigma \Sigma^*)^{ij} e^{-a_j u} du = \frac{(\Sigma \Sigma^*)^{ij}}{a_i + a_j}.$$

Assuming we know V_∞ , which we can estimate using results from Section 2.4, and as we can estimate $(\Sigma \Sigma^*)^{ii}$ using Theorem 3.1, we can estimate the values of a_i :

$$a_i = \frac{(\Sigma \Sigma^*)^{ii}}{2(V_\infty)^{ii}}.$$

We complete the inference by using the relation $(\Sigma \Sigma^*)^{ij} = (a_i + a_j)(V_\infty)^{ij}$.

4 Numerical tests

In the following, we present some inference results in the case of one-dimensional Ornstein-Uhlenbeck process. The observations are obtained via simulation and we consider simulation lengths n between 10 and $10 \cdot 2^{14}$ and discretization time-steps Δ between 0.1 and $0.1 \cdot 2^{-7}$.

We simulate trajectories using $\sigma = a = 1$:

$$dX_t = -X_t dt + dW_t.$$

For each data point characterized by (n, Δ) , we compute the expectation and standard deviation of OT_n and \mathcal{C}_n . These are empirically computed using 50000 simulated trajectories, independently initialized in the stationary distribution. Each figure will show these values plotted against either n or Δ in blue, and regression lines are added to each series in red. The figures are plotted in log-log scale, in order to show dependence of the results as powers of n and Δ .

For some plots, the regressions don't give a clear answer on the power dependence. Because we know many of our results come from the long-time limit, we privilege data series that maximize $n\Delta_n$, the horizon of the simulation.

Plots. For programming and plotting purposes, it is clearer to define statistics OT_n and \mathcal{C}_n omitting normalization. Therefore, only in this paragraph, we use

$$\widetilde{\text{OT}}_n = \sum_{k=0}^{n-1} \mathbb{1}_{X_{k\Delta_n} > 1}, \quad \widetilde{\mathcal{C}}_n = \sum_{k=0}^{n-1} \mathbb{1}_{X_{k\Delta_n} \neq X_{(k+1)\Delta_n}}.$$

Additionally, we use ts as a notation for Δ_n .

Figures 1 and 2 show a linear dependence in n and no dependence in Δ . This is in agreement with the stationarity of the process.

The lines fitted to the scatter in Figure 3 are somewhat misleading. Their slopes vary from 0.52 to 0.75. However, given our analysis in the paper, it is $n\Delta$, the inferring

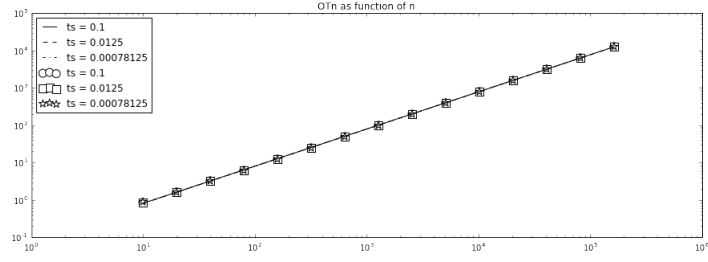


Figure 1: Expectation of $\widetilde{\text{OT}}_n$ versus n for different values of time steps $\Delta_n (ts)$

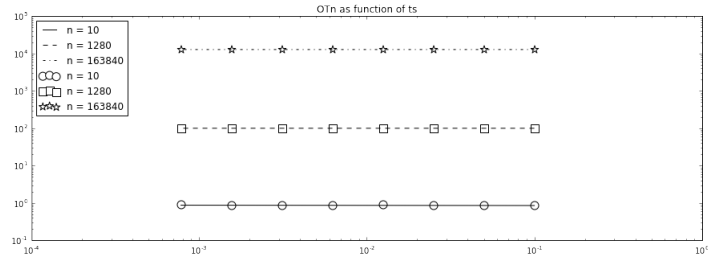


Figure 2: Expectation of $\widetilde{\text{OT}}_n$ versus $\Delta_n (ts)$ for different values of n

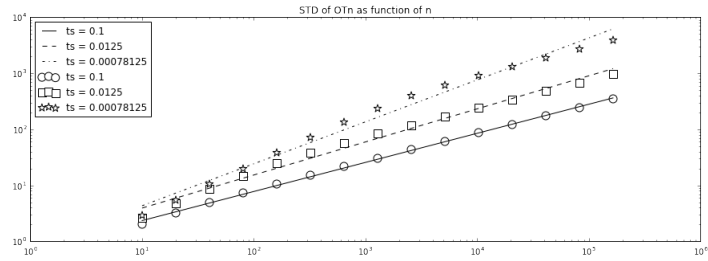


Figure 3: Standard deviation of $\widetilde{\text{OT}}_n$ versus n for different values of time steps $\Delta_n (ts)$

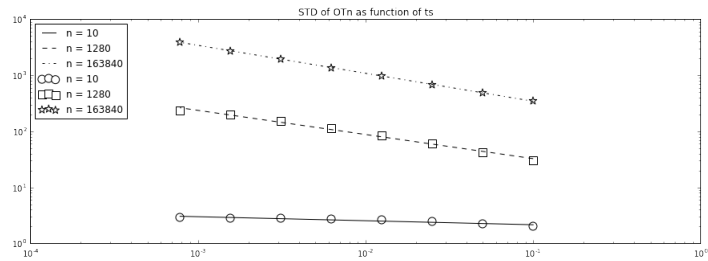


Figure 4: Standard deviation of $\widetilde{\text{OT}}_n$ versus $\Delta_n (ts)$ for different values of n

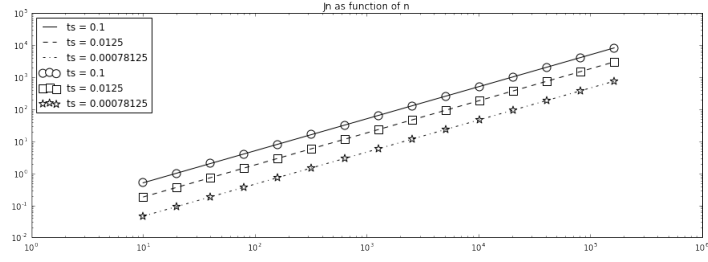


Figure 5: Expectation of $\widetilde{\mathcal{C}}_n$ versus n for different values of time steps Δ_n (ts)

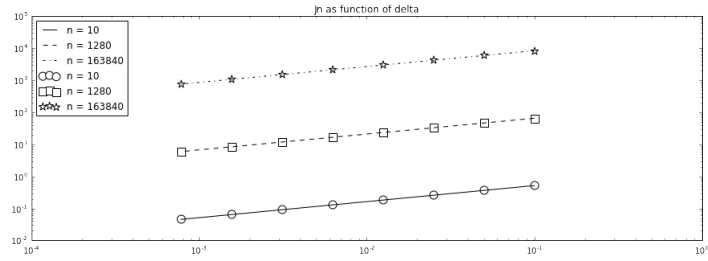


Figure 6: Expectation of $\widetilde{\mathcal{C}}_n$ versus Δ_n (ts) for different values of n

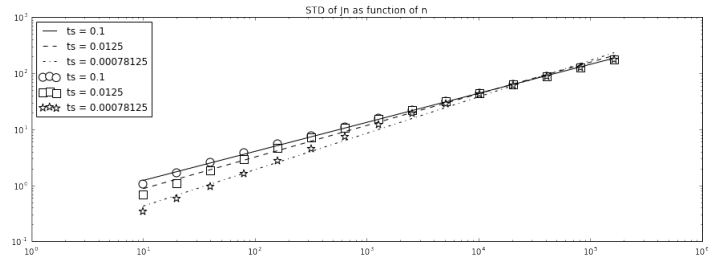


Figure 7: Standard deviation of $\widetilde{\mathcal{C}}_n$ versus n for different values of time steps Δ_n (ts)

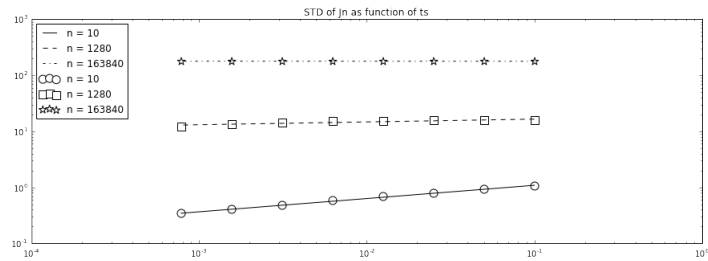


Figure 8: Standard deviation of $\widetilde{\mathcal{C}}_n$ versus Δ_n (ts) for different values of n

horizon, that has the largest impact on the quality of the convergence. For this reason, we should concentrate on high values of n . Further work on the last 5 points of the scatters concludes with a consistent slope of 0.5. The same applies to Figure 4: although slopes vary from -0.07 to -0.5 , we have the most confidence in the line corresponding to the highest value of n . We retain then the value of -0.5 .

Thus, we numerically observe (up to constants)

$$\mathbb{E} [\widetilde{\text{OT}}_n] \sim n, \quad \sqrt{\text{Var} (\widetilde{\text{OT}}_n)} \sim n^{1/2} \Delta^{-1/2}.$$

Regarding $\widetilde{\mathcal{C}}_n$, Figures 5 and 6 show clear power dependencies of 1 and 0.5 with respect to n and Δ . Again, Figures 7 and 8 have to be observed only at the largest values of n . We conclude with slopes of 0.5 and 0.

$$\mathbb{E} [\widetilde{\mathcal{C}}_n] \sim n \Delta^{1/2}, \quad \sqrt{\text{Var} (\widetilde{\mathcal{C}}_n)} \sim n^{1/2}.$$

Summary. Taking the results from the preceding paragraph and rewriting them using our regular expressions of OT_n and \mathcal{C}_n , as in Definitions 2.1 and 3.1, we get from these numerical tests:

- For OT_n :

$$\begin{aligned} & - \mathbb{E} [\text{OT}_n] \propto 1, \\ & - \text{Var} (\text{OT}_n) \propto \frac{1}{n \Delta_n}. \end{aligned}$$

This is in agreement with Theorem 2.2.

- For \mathcal{C}_n :

$$\begin{aligned} & - \mathbb{E} [\mathcal{C}_n] \propto \Delta_n^{1/2}, \\ & - \text{Var} \left(\frac{\mathcal{C}_n}{\sqrt{\Delta_n}} \right) \propto \frac{1}{n \Delta_n}. \end{aligned}$$

The expectation estimate is in agreement with Theorem 3.1. However, our variance estimate is seemingly not optimal: the missing factor $\Delta_n^{1/2}$ may come from subtle cancellations in small time, in conjunction with the low regularity of indicator function. This issue is left to future research.

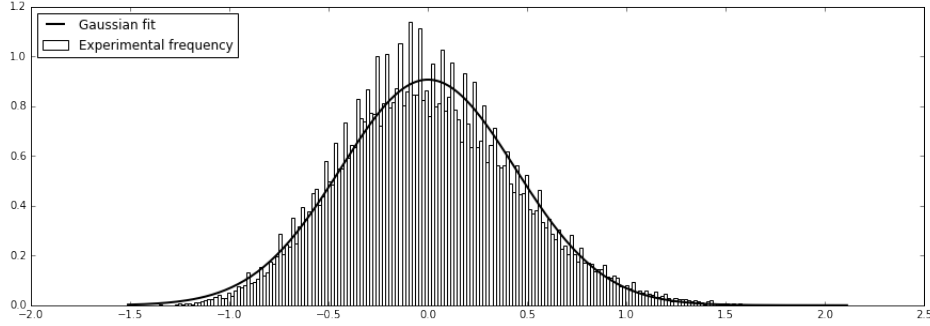
In Table 1 we can compare our theoretical limits with the estimates from simulation.

Numerical investigation regarding a central limit theorem for \mathcal{C}_n . Our experimental observation of $\text{Var} \left(\frac{\mathcal{C}_n}{\sqrt{\Delta_n}} \right) \propto \frac{1}{n \Delta_n}$ suggests we may expect a central limit theorem for $\sqrt{n \Delta_n} \left(\frac{\mathcal{C}_n}{\sqrt{\Delta_n}} - m \right)$ where $m = \lim_{n \rightarrow +\infty} \mathbb{E} \left[\frac{\mathcal{C}_n}{\Delta_n} \right]$. However, this result is out of the scope of the present paper.

Nonetheless we use the results of our simulations in order to see whether this conjecture is likely. In Figure 9, we compare the normalized histogram of \mathcal{C}_n to the probability density function of a fitted Gaussian. The agreement of the two seems to show that the validity of a central limit theorem is likely.

	Simulation result	Theoretical value
$\mathbb{E}[\text{OT}_n]$	0.07977	0.07865
$\mathbb{E}\left[\frac{\mathcal{C}_n}{\sqrt{\Delta_n}}\right]$	0.16418	0.16560
$n\Delta_n \text{Var}(\text{OT}_n)$	0.25902	N/A
$n\Delta_n \text{Var}\left(\frac{\mathcal{C}_n}{\sqrt{\Delta_n}}\right)$	0.43324	N/A

Table 1: Observed versus theoretical values of the limits of the expressions on the left

Figure 9: Normalized histogram of \mathcal{C}_n , for $n = 10 \cdot 2^{14}$ and $\Delta_n = 0.1 \cdot 2^{-7}$ and plot of PDF of Gaussian fit

A Expectation of threshold crossing for OU processes in small time

In this section, X is the d -dimensional Ornstein-Uhlenbeck process of Equation (1.4). Define the indicator of the crossing of the threshold 1 by the first coordinate of X :

$$Z_\Delta = \mathbb{1}_{X_0^1 < 1} \mathbb{1}_{X_\Delta^1 \geq 1}.$$

From Section 1.5.2, the first coordinate of $X_\Delta | X_0$ is a Gaussian vector centered at $(e^{-A\Delta} X_0)^1$ with variance V_Δ^{11} . Therefore $\mathbb{E} \left[\mathbb{1}_{X_\Delta^1 \geq 1} \mid X_0 \right] = N \left(\frac{(e^{-A\Delta} X_0)^1 - 1}{\sqrt{V_\Delta^{11}}} \right)$ and

$$\mathbb{E}[Z_\Delta] = \int_{\mathbb{R}^d} \mathbb{1}_{x^1 < 1} N \left(\frac{(e^{-A\Delta} x)^1 - 1}{\sqrt{V_\Delta^{11}}} \right) \mu_\infty(x) dx. \quad (\text{A.1})$$

Lemma A.1. *Let h be a bounded continuous increasing non-negative function, integrable in $-\infty$ and define:*

$$I_h(\Delta) := \int_{\mathbb{R}^d} \mathbb{1}_{x^1 < 1} h \left(\frac{(e^{-A\Delta} x)^1 - 1}{\sqrt{V_\Delta^{11}}} \right) \mu_\infty(x) dx.$$

Then when $\Delta \rightarrow 0$,

$$I_h(\Delta) \sim \sqrt{\Delta} \sqrt{(\Sigma \Sigma^*)^{11}} \mu_{V_\infty^{11}}(1) \int_{-\infty}^0 h(x) dx.$$

Proof. We use the following change of variables:

$$(y^1, y^2, \dots, y^d) = \left(\frac{x^1 - 1}{\sqrt{V_\Delta^{11}}}, x^2, \dots, x^d \right),$$

$$\frac{I_h(\Delta)}{\sqrt{V_\Delta^{11}}} = \int_{\mathbb{R}^d} \mathbb{1}_{y^1 < 0} h \left((e^{-A\Delta})^{11} y^1 + \frac{(e^{-A\Delta})^{11} - 1 + \sum_{i \geq 2} (e^{-A\Delta})^{1i} y^i}{\sqrt{V_\Delta^{11}}} \right) \times \mu_\infty \left(1 + \sqrt{V_\Delta^{11}} y^1, y^2, \dots, y^d \right) dy. \quad (\text{A.2})$$

Using the power expansion of the matrix exponential $e^{-A\Delta} = \text{Id} - \Delta A + O(\Delta^2)$, we have:

$$(e^{-A\Delta})^{11} = 1 + O(\Delta),$$

$$(e^{-A\Delta})^{1i} = \Delta A^{1i} + O(\Delta^2), \quad i \geq 2.$$

Simultaneously, we know that $V_\Delta \sim \Delta(\Sigma\Sigma^*)$ as $\Delta \rightarrow 0$. Hence in the integral (A.2), $h(\dots)$ and $\mu_\infty(\dots)$ converge pointwise to $h(y^1)$ and $\mu_\infty(1, y^2, \dots, y^d)$ respectively, for any y . To pass to the limit for (A.2), it remains to dominate $h(\dots)\mu_\infty(\dots)$ uniformly in Δ , by an integrable function on \mathbb{R}^d .

- On the one hand, using the monotone and non-negative properties of h , we observe that $h(\dots)$ is bounded by $h(\frac{1}{2}y^1 + 2\sum_{j \geq 2} |y^j|)$ for any y such that $y^1 < 0$, provided that Δ is small enough.
- On the other hand, since V_∞^{-1} is symmetric definite positive, $\mu_\infty(\dots)$ is bounded by $\text{Cst} \exp(-\sum_{j \geq 2} |y^j|^2 / \text{Cst})$ for some positive constant Cst .

Finally, we easily check that the product of the two bounds is integrable, using in particular that h is integrable on \mathbb{R}^- and bounded over \mathbb{R} . Thus the dominated convergence theorem yields

$$\frac{I_h(\Delta)}{\sqrt{V_\Delta^{11}}} \rightarrow \int_{-\infty}^0 \int_{\mathbb{R}^{d-1}} h(y^1) \mu_\infty(1, y^2, \dots, y^d) dy,$$

$$I_h(\Delta) \sim \sqrt{\Delta} \sqrt{(\Sigma\Sigma^*)^{11}} \int_{-\infty}^0 h(x) dx \int_{\mathbb{R}^{d-1}} \mu_\infty(1, y^2, \dots, y^d) dy.$$

We can simply rewrite the last integral. Let $G \stackrel{\text{d}}{=} \mathcal{N}(0, V_\infty)$:

$$\int_{-\infty}^x \int_{\mathbb{R}^{d-1}} \mu_\infty(y^1, z) dy^1 dz = \mathbb{P}[G^1 \leq x] = \int_{-\infty}^x \mu_{V_\infty^{11}}(y^1) dy^1$$

where we have used that $G^1 \stackrel{\text{d}}{=} \mathcal{N}(0, V_\infty^{11})$ at the second equality. It gives, for any $x \in \mathbb{R}$,

$$\int_{\mathbb{R}^{d-1}} \mu_\infty(x, y^2, \dots, y^d) dy = \mu_{V_\infty^{11}}(x),$$

and therefore the announced result. \square

Corollary A.1. *In the limit $\Delta \rightarrow 0$, we have*

$$\mathbb{E}[Z_\Delta] \sim \text{Var}(Z_\Delta) \sim \sqrt{\Delta} \sqrt{\frac{(\Sigma \Sigma^*)^{11}}{2\pi}} \mu_{V_\infty^{11}}(1).$$

Proof. Since Z_Δ takes values in $\{0, 1\}$, if $\mathbb{E}[Z_\Delta] \rightarrow 0$, we have $\text{Var}(Z_\Delta) = \mathbb{E}[Z_\Delta](1 - \mathbb{E}[Z_\Delta]) \sim \mathbb{E}[Z_\Delta]$. Thus, it remains to show the estimate on $\mathbb{E}[Z_\Delta]$.

Start from (A.1) and apply Lemma A.1 with $h(\cdot) = N(\cdot)$. We have $\int_{-\infty}^0 N(x) dx = \frac{1}{\sqrt{2\pi}}$, hence the result. \square

B Maximal correlation inequality

In this section, we aim at proving the very useful Theorem 2.1.

B.1 Gebelein's inequality

We start by an abstract version.

Theorem B.1 ([Jan97, Theorem 10.11]). *Take H, K two closed subspaces of some Gaussian Hilbert space. Define P_{HK} the restriction to H of the orthogonal projection onto K . Define the maximal correlation coefficient between variables A, B respectively measurable w.r.t. the sigma field generated by H and K :*

$$\rho(H, K) = \sup_{A \in L^2(H), B \in L^2(K)} |\text{Cor}(A, B)|.$$

Then we have:

$$\rho(H, K) = \|P_{HK}\|$$

where $\|\cdot\|$ is the operator norm.

We now restate in a more convenient way the above result in a finite dimensional case and for given Gaussian vectors: we believe such a statement may exist in the literature but we could not find a ready reference. Consider a Gaussian Hilbert space and two Gaussian vectors X, Y in this space. Define then H_X and H_Y the subspaces spanned respectively by X and Y .

Denoting P_{XY} the orthogonal projection on H_Y restricted to H_X , from Theorem B.1 we have that $\rho(H_X, H_Y) = \|P_{XY}\|$. This value is independent of the enclosing Gaussian Hilbert Space.

Using the notations stated above, set

$$\rho(X, Y) = \rho(H_X, H_Y) = \|P_{XY}\|.$$

Corollary B.1. *Take X, Y two Gaussian vectors in the same Gaussian Hilbert space and f, g two functions such that $f(X), g(Y)$ are square-integrable. Then we have:*

$$|\text{Cov}(f(X), g(Y))| \leq |\rho(X, Y)| \sqrt{\text{Var}(f(X)) \text{Var}(g(Y))}.$$

In preparation of Theorem 2.1, we now aim at making more explicit the coefficients $\rho(X, Y)$ in terms of the correlations of the components of X and Y , or of their transforms.

B.2 Finite-dimensional Gaussian vectors

Notation. We denote $\text{Cov}(X)$ the covariance matrix of X and $\text{Cov}(X, Y)$ the covariance matrix of X and Y , which is also the upper-right part of the covariance matrix of the vector (X, Y) . We reserve the notation $\text{Var}(X)$ for the variance of a real-valued X .

We set $K_{XX} = \text{Cov}(X)$, $K_{YY} = \text{Cov}(Y)$ and $K_{XY} = \text{Cov}(X, Y)$. As K_{XX} and K_{YY} are symmetric non-negative definite matrices, there exist O_X, O_Y orthogonal matrices and D_X, D_Y diagonal non-negative definite matrices such that we have:

$$\begin{aligned} K_{XX} &= O_X^* D_X O_X = \left(O_X^* D_X^{1/2} O_X \right) \left(O_X^* D_X^{1/2} O_X \right)^*, \\ K_{YY} &= O_Y^* D_Y O_Y = \left(O_Y^* D_Y^{1/2} O_Y \right) \left(O_Y^* D_Y^{1/2} O_Y \right)^*. \end{aligned}$$

Note that $\left(O_X^* D_X^{1/2} O_X \right)$ and $\left(O_Y^* D_Y^{1/2} O_Y \right)$ are symmetric.

Take now X' and Y' orthonormal basis of respectively H_X and H_Y . It is easy to see that the covariance matrix R of X' and Y' is the projection matrix from H_X on H_Y written in the basis X' and Y' and therefore $\rho(X, Y) = \|R\|$.

B.2.1 Non-degenerate case

Assume, in this paragraph, D_X and D_Y are non-degenerate or, equivalently, that K_{XX} and K_{YY} are non-degenerate.

We can choose $X' = \left(O_X^* D_X^{1/2} O_X \right)^{-1} X$ and $Y' = \left(O_Y^* D_Y^{1/2} O_Y \right)^{-1} Y$. Our calculation shows then that

$$R = \left(O_X^* D_X^{1/2} O_X \right)^{-1} K_{XY} \left(O_Y^* D_Y^{1/2} O_Y \right)^{-1}.$$

From the other side, $O_X^* D_X^{1/2} O_X$ and $O_Y^* D_Y^{1/2} O_Y$ are symmetric matrices which square to K_{XX} and K_{YY} . Therefore we have

$$(K_{XX})^{1/2} = O_X^* D_X^{1/2} O_X, \quad (K_{YY})^{1/2} = O_Y^* D_Y^{1/2} O_Y.$$

From this we have the following proposition.

Proposition B.1. *Take (X, Y) a Gaussian vector. Assume that $\text{Cov}(X)$, $\text{Cov}(Y)$ are non-degenerate. Then we have*

$$\rho(X, Y) = \|\text{Cov}(X)^{-1/2} \text{Cov}(X, Y) \text{Cov}(Y)^{-1/2}\|.$$

Corollary B.2. *If X and Y are orthogonal with non degenerate components ($\text{Var}(X^i) \neq 0$, $\text{Var}(Y^j) \neq 0$, for any i, j), then we have $\rho(X, Y) = \|R\|$ with*

$$R^{ij} = \text{Cor}(X^i, Y^j) = \frac{\text{Cov}(X^i, Y^j)}{\sqrt{\text{Var}(X^i) \text{Var}(Y^j)}}.$$

B.2.2 Degenerate case

In this part, we consider that X or Y is degenerate (or both). We know there are orthonormal matrices O_X and O_Y such that $O_X X$ and $O_Y Y$ both have diagonal covariances. These diagonals can have zero values; taking only the variables of $O_X X$ and $O_Y Y$ that have non-zero variances, we get a couple of orthonormal families. Applying results from previous paragraph and using the fact that extending a matrix with zeroes doesn't change its operator norm, we have the following.

Proposition B.2. *Let (X, Y) be a Gaussian vector and O_X, O_Y two orthogonal matrices such that $X_O := O_X X$ and $Y_O := O_Y Y$ are respectively orthogonal families. Then the maximal correlation coefficient verifies $\rho(X, Y) = \|R\|$ with*

$$R^{ij} = \text{Cor}(X_O^i, Y_O^j) = \begin{cases} \frac{\text{Cov}(X_O^i, Y_O^j)}{\sqrt{\text{Var}(X_O^i) \text{Var}(Y_O^j)}} & \text{if } \text{Var}(X_O^i) \text{Var}(Y_O^j) \neq 0, \\ 0 & \text{if } \text{Var}(X_O^i) \text{Var}(Y_O^j) = 0. \end{cases}$$

B.3 Application to functions of Gaussian processes

We are now in a position to give the maximal correlation between X_s and X_t for $t, s \in \mathbb{R}^+$, in terms of the OU parameters (A, Σ) (Assumption **(H)**).

Proposition B.3. *Using the previous notation, we have for any $s, t \in \mathbb{R}^+$*

$$\rho(X_s, X_t) \leq \sqrt{\frac{v_M}{v_m}} e^{-a_0|t-s|},$$

where $a_0 := \min_{\lambda \in \text{Sp}(A)} \text{Re}(\lambda)$, $v_M = \max_{\lambda \in \text{Sp}(V_\infty)} \lambda$, $v_m = \min_{\lambda \in \text{Sp}(V_\infty)} \lambda$.

Proof. Let $t \geq s \geq 0$. From (1.5) and (1.8), we have

$$\text{Cov}(X_s) = \text{Cov}(X_t) = V_\infty, \quad \text{Cov}(X_t, X_s) = e^{-A(t-s)} V_\infty.$$

Since V_∞ is non degenerate (owing to Assumption **(H)**), we can apply proposition B.1, to get

$$\rho(X_s, X_t) = \|V_\infty^{-1/2} e^{-A(t-s)} V_\infty^{1/2}\|.$$

The bound on $\rho(X_s, X_t)$ is a consequence of sub-multiplicativity of the operator norm. \square

We immediately deduce the following.

Corollary B.3. *For any $\varphi, \phi : \mathbb{R}^d \rightarrow \mathbb{R}$ square-integrable w.r.t. ν_∞ , we have*

$$|\text{Cov}(\varphi(X_s), \phi(X_t))| \leq \sqrt{\frac{v_M}{v_m}} e^{-a_0|t-s|} \sqrt{\text{Var}(\varphi(X_s)) \text{Var}(\phi(X_t))}.$$

We can proceed to the proof of Theorem 2.1. In its setting we have $t > s \geq 0$ and φ, ϕ take as arguments respectively $(X_u)_{0 \leq u \leq s}$ and $(X_v)_{v \geq t}$.

Proof. For ease of writing, denote $\varphi_s = \varphi\left((X_u)_{0 \leq u \leq s}\right)$ and $\phi_t = \phi\left((X_v)_{v \geq t}\right)$. Without loss of generality, we can assume $\mathbb{E}[\varphi_s] = \mathbb{E}[\phi_t] = 0$. We repeatedly use the Markov property and the tower property of conditional expectation to write

$$\begin{aligned} \text{Cov}(\varphi_s, \phi_t) &= \mathbb{E}[\varphi_s \phi_t] \\ &= \mathbb{E}[\varphi_s \mathbb{E}[\phi_t \mid \mathcal{F}_s]] \\ &= \mathbb{E}[\varphi_s \mathbb{E}[\phi_t \mid X_s]] \\ &= \mathbb{E}[\mathbb{E}[\varphi_s \mid X_s] \mathbb{E}[\phi_t \mid X_s]] \\ &= \mathbb{E}[\mathbb{E}[\mathbb{E}[\varphi_s \mid X_s] \phi_t \mid X_s]] \\ &= \mathbb{E}[\mathbb{E}[\varphi_s \mid X_s] \phi_t] \\ &= \mathbb{E}[\mathbb{E}[\varphi_s \mid X_s] \mathbb{E}[\phi_t \mid X_t]] \\ &= \text{Cov}(\mathbb{E}[\varphi_s \mid X_s], \mathbb{E}[\phi_t \mid X_t]). \end{aligned}$$

We define now $f(X_s) = \mathbb{E}[\varphi_s \mid X_s]$, $g(X_t) = \mathbb{E}[\phi_t \mid X_t]$ and apply Corollary B.3:

$$\text{Cov}(f(X_s), g(X_t)) \leq \sqrt{\frac{v_M}{v_m}} e^{-a_0|t-s|} \sqrt{\text{Var}(f(X_s)) \text{Var}(g(X_t))}.$$

Then, the announced inequality of Theorem 2.1 stems from the standard decomposition $\text{Var}(h) = \text{Var}(\mathbb{E}[h \mid \mathcal{G}]) + \mathbb{E}[\text{Var}(h \mid \mathcal{G})] \geq \text{Var}(\mathbb{E}[h \mid \mathcal{G}])$ for any sigma-field \mathcal{G} and any square-integrable variable h . \square

C Central limit theorem for discontinuous functions of OU processes

In this section, we broaden the domain of application of the properties from [Flo84], precisely extending it to the case of a bounded function with a single point of irregularity, as $x \mapsto \mathbb{1}_{x \geq 1}$. In this section, the process X is one-dimensional.

C.1 Itô formula for piecewise C^2 function

First, we recall a generalization of Itô's lemma to functions that are not C^2 .

Lemma C.1. *Let g be a function $g : \mathbb{R} \rightarrow \mathbb{R}$, twice differentiable and g'' is continuous except at a single point z . Assume also that $\forall x \neq z, |g''(x)| \leq K$. Then Itô's formula applies to g , i.e.:*

$$dg(X_t) = g'(X_t)dX_t + \frac{\sigma^2}{2}g''(X_t)dt.$$

Proof. As we have $g \in C^1$ and g'' is integrable on any interval, we can apply [RW87, Lemma 45.9]. \square

C.2 Solution to Poisson equation $LF = -f$

Consider now a function f smooth with the exception of a single point. Consider also L the infinitesimal generator associated to X verifying (2.3):

$$L\phi(x) = -ax \frac{\partial \phi}{\partial x}(x) + \frac{\sigma^2}{2} \frac{\partial^2 \phi}{\partial x^2}(x).$$

The next lemma gives a solution to the Poisson equation $LF = -f$.

Lemma C.2. *Let f be a bounded function. Then*

$$F(x) := -\frac{2}{\sigma^2} \int_0^x \frac{dy}{\mu_\infty(y)} \int_{-\infty}^y f(u) \mu_\infty(u) du \quad (\text{C.1})$$

is a solution to $LF = -f$.

With $f = \mathbf{1}_{[1, +\infty[} - N(-1/\sqrt{v_\infty})$, we have:

$$F(x) = \frac{2}{\sigma^2} \int_0^x \frac{N\left(\frac{u \wedge 1}{\sqrt{v_\infty}}\right) - N\left(\frac{1}{\sqrt{v_\infty}}\right) N\left(\frac{u}{\sqrt{v_\infty}}\right)}{\mu_\infty(u)} du. \quad (\text{C.2})$$

Proof. Simple computations give:

$$\begin{aligned} \frac{\partial F}{\partial x}(x) &= -\frac{2}{\sigma^2 \mu_\infty(x)} \int_{-\infty}^x f(u) \mu_\infty(u) du, \\ \frac{\partial^2 F}{\partial x^2}(x) &= -2 \frac{2ax}{\sigma^2 \mu_\infty(x)} \int_{-\infty}^x \frac{f(u)}{\sigma^2} \mu_\infty(u) du - 2 \frac{f(x)}{\sigma^2}. \end{aligned}$$

We then deduce $LF = -f$.

Choosing $f = \mathbf{1}_{[1, +\infty[} - N(-1/\sqrt{v_\infty}) = 1 - \mathbf{1}_{]-\infty, 1[} - N(-1/\sqrt{v_\infty}) = -\mathbf{1}_{]-\infty, 1[} + N(1/\sqrt{v_\infty})$, we get:

$$\begin{aligned} F(x) &= \frac{2}{\sigma^2} \left(\int_0^x \frac{dy}{\mu_\infty(y)} \int_{-\infty}^{y \wedge 1} \mu_\infty(u) du - \int_0^x \frac{N\left(\frac{1}{\sqrt{v_\infty}}\right)}{\mu_\infty(y)} dy \int_{-\infty}^y \mu_\infty(u) du \right) \\ &= \frac{2}{\sigma^2} \int_0^x \frac{N\left(\frac{u \wedge 1}{\sqrt{v_\infty}}\right) - N\left(\frac{1}{\sqrt{v_\infty}}\right) N\left(\frac{u}{\sqrt{v_\infty}}\right)}{\mu_\infty(u)} du. \end{aligned}$$

□

We now establish bounds on F and its derivatives.

Lemma C.3. *Assume that f is bounded and such that $\nu_\infty(f) = 0$. Define F as in (C.1). Then there exist finite constants $C_{(\text{C.3})}, C_{(\text{C.4})}, C_{(\text{C.5})}$ (depending only on the model) such that, for any $x \in \mathbb{R}$,*

$$|F(x)| \leq C_{(\text{C.3})} |f|_\infty |x|, \quad (\text{C.3})$$

$$|F'(x)| \leq C_{(\text{C.4})} |f|_\infty \left(1 \wedge \frac{1}{|x|} \right), \quad (\text{C.4})$$

$$|F''(x)| \leq C_{(\text{C.5})} |f|_\infty. \quad (\text{C.5})$$

Proof. Using the assumption that $\nu_\infty(f) = \int_{\mathbb{R}} f(x) \mu_\infty(x) dx = 0$, we can write:

$$\begin{aligned} F'(x) &= -\frac{2}{\sigma^2 \mu_\infty(x)} \int_{-\infty}^x f(u) \mu_\infty(u) du \\ &= \frac{2}{\sigma^2 \mu_\infty(x)} \int_x^{+\infty} f(u) \mu_\infty(u) du, \\ |F'(x)| &\leq \frac{2}{\sigma^2 \mu_\infty(x)} \left(\int_{-\infty}^x |f(u)| \mu_\infty(u) du \wedge \int_x^{+\infty} |f(u)| \mu_\infty(u) du \right). \end{aligned}$$

Applying now the assumption that f is bounded, we have:

$$|F'(x)| \leq \frac{2|f|_\infty}{\sigma^2} \frac{N\left(-\frac{|x|}{\sqrt{v_\infty}}\right)}{\mu_\infty(x)}. \quad (\text{C.6})$$

Using the classic inequality $N(-|x|) \leq e^{-x^2/2}$ yields

$$|F'(x)| \leq \frac{2|f|_\infty}{\sigma^2} \frac{\exp\left(-\frac{x^2}{2v_\infty}\right)}{\frac{1}{\sqrt{2\pi v_\infty}} \exp\left(-\frac{x^2}{2v_\infty}\right)} \leq 2\sqrt{\frac{\pi}{a\sigma^2}} |f|_\infty. \quad (\text{C.7})$$

By integrating, we complete the proof of Equation (C.3). Next, we use the Mills inequality,

$$N(-|x|) \leq \frac{e^{-x^2/2}}{\sqrt{2\pi}|x|},$$

which combined with (C.6) gives

$$|x| |F'(x)| \leq \frac{2|f|_\infty}{\sigma^2} |x| \frac{N\left(-\frac{|x|}{\sqrt{v_\infty}}\right)}{\mu_\infty(x)} \leq \frac{2|f|_\infty}{\sigma^2} |x| \frac{\frac{1}{\sqrt{2\pi} \frac{|x|}{\sqrt{v_\infty}}}}{\frac{1}{\sqrt{2\pi v_\infty}}} \leq \frac{2|f|_\infty}{\sigma^2} v_\infty.$$

Thus, joined with (C.7), Inequality (C.4) is proved. Last, as $LF = -f$,

$$|F''(x)| = \frac{2}{\sigma^2} |axF'(x) - f(x)| \leq \frac{2}{\sigma^2} \left(a \frac{2v_\infty}{\sigma^2} + 1 \right) |f|_\infty = \frac{4}{\sigma^2} |f|_\infty,$$

which proves (C.5). \square

C.3 CLT for multi-dimensional continuous-time martingales

Lemma C.4 ([van00, Theorem 4.1]). *Let $(M_t; \mathcal{F}_t : t \geq 0)$ be a d -dimensional continuous local martingale. If there exist invertible, non-random $d \times d$ -matrices $(K_t : t \geq 0)$ such that as $t \rightarrow \infty$*

- $K_t \langle M \rangle_t K_t^* \xrightarrow{\mathbb{P}} \eta \eta^*$ where η is a random $d \times d$ -matrix;
- $|K_t| \rightarrow 0$;

then, for each \mathbb{R}^k -valued random vector X defined on the same probability space as M , we have

$$(K_t M_t, X) \xrightarrow{d} (\eta Z, X) \quad \text{as } t \rightarrow \infty,$$

where $Z \stackrel{d}{=} \mathcal{N}(0, \text{Id})$ and Z is independent of (η, X) .

References

- [ASM04] Y. Ait-Sahalia and P.A. Mykland. Estimators of diffusions with randomly spaced discrete observations: A general theory. *Ann. Statist.*, 32(5):2186–2222, 10 2004.
- [Bol01] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics, 2nd edition, 2001.
- [CFS15] R. Carmona, J.-P. Fouque, and L.-H. Sun. Mean field games and systemic risk. *Commun. Math. Sci.*, 13(4):911–933, 2015.
- [FI13] J.-P. Fouque and T. Ichiba. Stability in a model of interbank lending. *SIAM J. Financial Math.*, 4(1):784–803, 2013.
- [Flo84] D. Florens-Zmirou. Statistics on crossings of discretized diffusions and local time. *Stochastic Processes and their Applications*, 39(1):139 – 151, 1984.
- [Flo87] D. Florens-Zmirou. Estimation du paramètre d’une diffusion par les changements de signe de sa discrétisée. *Comptes Rendus de l’Académie des Sciences Paris*, 305:661 – 664, 1987.
- [Flo89] D. Florens-Zmirou. Approximate discrete-time schemes for statistics of diffusion processes. *Statistics*, 20(4):547–557, 1989.
- [Flo91] D. Florens-Zmirou. Statistics on crossings of discretized diffusions and local time. *Stochastic Processes and their Applications*, 39(1):139 – 151, 1991.
- [GJ93] V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l’institut Henri Poincaré (B) Probabilités et Statistiques*, 29(1):119–151, 1993.
- [Gob02] E. Gobet. LAN property for ergodic diffusions with discrete observations. *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, 38(5):711–737, 2002.
- [HJ86] R.A. Horn and C.R. Johnson, editors. *Matrix Analysis*. Cambridge University Press, New York, NY, USA, 1986.
- [IUY09] S.M. Iacus, M. Uchida, and N. Yoshida. Parametric estimation for partially hidden diffusion processes sampled at discrete times. *Stochastic Processes and their Applications*, 119(5):1580 – 1600, 2009.
- [Jan97] S. Janson. *Gaussian Hilbert Spaces*. Cambridge Tracts in Mathematics. Cambridge University Press, 1997.
- [Kes97] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24(2):211–229, 1997.
- [KLS12] M. Kessler, A. Lindner, and M. Sørensen, editors. *Statistical methods for stochastic differential equations*, volume 124 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2012. Revised papers from the 7th Séminaire Européen de Statistique on Statistics for Stochastic Differential Equations Models held in Cartagena, May 7–12, 2007.

- [KS91] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 1991.
- [Kut04] Y.A. Kutoyants. *Statistical inference for ergodic diffusion processes*. Springer Series in Statistics. Springer-Verlag London, Ltd., London, 2004.
- [MX07] M. Mahdian and Y. Xu. Stochastic Kronecker graphs. In Anthony Bonato and Fan R. K. Chung, editors, *WAW*, volume 4863 of *Lecture Notes in Computer Science*, pages 179–186. Springer, 2007.
- [RW87] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes and Martingales, vol 2: Ito calculus*. John Wiley, 1987.
- [van00] H. van Zanten. A multivariate central limit theorem for continuous local martingales. *Statistics and Probability Letters*, 50(3):229 – 235, 2000.
- [Yos92] N. Yoshida. Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis*, 41(2):220 – 242, 1992.