J Stat Phys (2008) 132: 397–415
DOI 10.1007/s10955-008-9545-8

A Haar-like Construction for the Ornstein Uhlenbeck Process

Thibaud Taillefumier · Marcelo O. Magnasco

Received: 27 June 2007 / Accepted: 4 April 2008 / Published online: 25 April 2008

Abstract The classical Haar construction of Brownian motion uses a binary tree of triangular wedge-shaped functions. This basis has compactness properties which make it especially suited for certain classes of numerical algorithms. We present a similar basis for the Ornstein-Uhlenbeck process, in which the basis elements approach asymptotically the Haar functions as the index increases, and preserve the following properties of the Haar basis: all basis elements have compact support on an open interval with dyadic rational endpoints; these intervals are nested and become smaller for larger indices of the basis element, and for any dyadic rational, only a finite number of basis elements is nonzero at that number. Thus the expansion in our basis, when evaluated at a dyadic rational, terminates in a finite number of steps. We prove the covariance formulae for our expansion and discuss its statistical interpretation.

Keywords Ornstein-Uhlenbeck process · Brownian motion · Haar basis

1 Introduction

Random walks and continuous stochastic processes are of fundamental importance in a number of applied areas, including optics [1], chemical physics [2], biophysics [3, 4], biology [5] and finance [6]. The mathematical idealization of the one-dimensional continuous random walk, the Wiener process, can be expressed in—infinitely—many bases as a sum of random coefficients times basis elements. Unique among these bases, the Haar—or Schauder—basis has three properties that make it particularly suitable for certain numerical computations.

Electronic supplementary material The online version of this article (http://dx.doi.org/10.1007/s10955-008-9545-8) contains supplementary material, which is available to authorized users.

T. Taillefumier (✉) · M.O. Magnasco
Laboratory of Mathematical Physics, The Rockefeller University, 10021 New York, NY, USA
e-mail: ttaillefum@rockefeller.edu

M.O. Magnasco
e-mail: magnasco@rockefeller.edu

First, the basis elements all have *compact support*: the basis elements are nonzero only in open intervals. Second, the support is *increasingly compact*, i.e., the open intervals become smaller for higher indices of the basis elements; in fact, the intervals are nested in binary tree fashion, and have dyadic rational endpoints. Finally, given any dyadic rational, there is a finite number of basis elements which are nonzero at that number, so that evaluation of the Haar expansion at a dyadic rational *terminates* in a finite number of steps known beforehand. These properties can be used to great advantage in algorithms that construct the random walk in a "top-down" fashion, such as dychotomic search algorithms for first passage times.

However, the "plain" Wiener process has limited applicability in the areas mentioned above, so an extension of this construction to more complex stochastic processes is desirable. The naive generalization of the Haar basis construction to other stochastic processes fails to display our three properties. We present a method for constructing a Haar-like basis for the Ornstein-Uhlenbeck process which preserves these properties. The basis is therefore useful for advanced numerical computations: a fast dichotomic search algorithm for first passage time computations shall be presented elsewhere. The method we present is also amenable to further generalizations to other stochastic processes. We should emphasize that the advantages granted by use of this basis or its generalizations come at the price of generality: numerical integration of stochastic differential equations works for *any* stochastic differential equations, in particular many nonlinear vector fields of fundamental importance in physics and chemistry. Numerical methods tied to our basis construction work only for the specific stochastic process for which the basis is known; in the case presented here, only the Ornstein-Uhlenbeck process.

This paper is organized as follows. We first review some background in stochastic processes and basis expansions. Then we review the well-known Lévy construction of a Wiener process with the help of a basis of functions derived from the Haar system. In the third section, we give a statistical interpretation of such a construction, which leads us to propose a basis for the Ornstein-Uhlenbeck process. In the fourth section, we prove that the Ornstein-Uhlenbeck process is correctly represented as a discrete process in the proposed basis. In the last section, we extend the principle of the construction to build a bi-infinite representation.

2 Background on Stochastic Processes

2.1 General Definition for Stochastic Processes

The Wiener process and the Ornstein-Uhlenbeck process are continuous stochastic processes; we specify this class of process through the Langevin equation

$$\begin{cases} \dot{x} = f(x,t) + \eta(t), \\ x(t_0) = x_0, \quad t \in [t_0, T], \end{cases} \tag{1}$$

where f is a deterministic function and $\eta(t)$ describes the stochastic forcing. Equation (1) is a first order stochastic differential equation and its connection to the Fokker-Planck equation has been extensively studied [1, 2]. We only consider here the case where the noise is white and Gaussian: $\eta(t)$ are realizations of independent identically distributed Gaussian variables η_t, with time correlations satisfying

$$\langle \eta_t \cdot \eta_s \rangle = \Gamma \cdot \delta(s-t),$$

where δ is the Dirac distribution.

We denote by ω a given realization of the stochastic forcing: the collection of all the values $\{\eta(t)\}_{t\in[t_0,T]}$ in an interval. The set of ω values defines the sample space Ω and occurrence of a sequence ω in Ω is determined by the joint probability density of $\{\eta_t\}_{t\in[t_0,T]}$. With this notation, we can introduce the solution of the stochastic system as the stochastic process X. For a given realization of the noise ω, there is a unique solution to (1) called a sample path: neglecting to notate the dependence on the initial condition, we write $X_t(\omega)$ the value of this sample path at time t. The value $X_t(\omega)$ can be seen as the outcome of a random variable X_t defined on Ω. If X is the collection of random variable X_t for t in $[t_0, T]$, we say that X is the stochastic process solution of (1). Moreover, the process X has two important properties: it is a continuous process as it is defined for a continuous index set $[t_0, T]$; and, being a Markovian process, the value of X_t only depends on $\{\eta(u)\}_{u\in[0,t]}$, the sequence of realizations preceding t.

2.2 The Wiener Process and the Ornstein-Uhlenbeck Process

Two special forms of f shall concern us. When f is zero, the stochastic process solution of (1) is called the Wiener process W; when f is linear in x, the process is called the Ornstein-Uhlenbeck process U. These processes exhibit two interesting properties: first they are both Gaussian processes; second they both have continuous sample paths. Moreover, due to the relative simplicity of both situations, the probability laws of the processes (i.e., the Green functions of the associated Fokker-Planck equations) are known analytically. If a Wiener process is at x_0 at time $t = t_0$, the probability of finding the process between x and $x + dx$ at time t is

$$\mathbf{P}(W_t \in dx | W_{t_0} = x_0) = \frac{1}{\sqrt{2\pi}\cdot {}_W\sigma_t} \cdot \exp\left(-\frac{(x-x_0)^2}{2\cdot {}_W\sigma_t^2}\right)dx \tag{2}$$

with a variance ${}_W\sigma_t^2 = \Gamma\cdot(t-t_0)$. For the Ornstein-Uhlenbeck process where the function f is defined as $f(x) = -\alpha x$, a similar result holds

$$\mathbf{P}(U_t \in dx | U_{t_0} = x_0) = \frac{1}{\sqrt{2\pi}\cdot {}_U\sigma_t} \cdot \exp\left(-\frac{(x-x_0e^{-\alpha(t-t_0)})^2}{2\cdot {}_U\sigma_t^2}\right)dx \tag{3}$$

with a variance ${}_U\sigma_t^2 = \frac{\Gamma}{2\alpha}\cdot(1-e^{-\alpha(t-t_0)})$. The previous expressions describe the statistics of W and U, which will be called X when collectively designated.

2.3 Discrete Representation of Continuous Process

Continuous processes require an non-countable number of random variables, and establishing results about them is quite difficult. Consider for example the Ornstein-Uhlenbeck process, widely used from finance to neuroscience: finding analytically the first-passage times distribution with a fixed threshold proves a surprisingly intricate question [7–9]: only three representations of analytical nature have been derived in this situation, ranging from an expansion in terms of eigenfunctions, to an integral representation in terms of special functions through a functional of the three-dimensional Brownian bridge. These representations give rise to efficient numerical methods [10] but for a general form of threshold, sample paths are only approximated by stochastic Euler methods, or Runge-Kutta-like higher-order methods, with integration schema of low efficiency [11–13].

The mathematical difficulties in numerical integration of Langevin equations come from two roots. First, the stochastic term in the stochastic differential equations is $O(\sqrt{\Delta t})$ when integrating with a timestep Δt. Higher-order methods whose schema would yield $O(\Delta t^{\alpha})$ for an ordinary differential equation yield at the most $O(\Delta t^{\alpha/2})$ in a stochastic differential equation, and they have to be carefully constructed and evaluated [14–16]. Second, even this slower convergence should not be taken for granted: the Runge-Kutta integration schema assumes that the solution can be expanded in Taylor series in Δt, which is manifestly not the case for the solution of a stochastic differential equation. The correctness of the term-by-term expansion in Δt may be misleading: as the expansion itself is not a convergent series, equality of the expansion coefficients is not a strong guarantee, and some methods have been derived to sidestep integration order [17, 18]. These problems become manifest rather sporadically, and in general, numerical integration of stochastic differential equations is a phenomenally useful tool. However, at the purely practical level it is well-known that such integration is numerically costly.

To circumvent the problem, it is advantageous to represent a continuous process as a discrete process. Conspicuously enough, a discrete process has a countable index set of random variables. At stake is to write a Gaussian process X as a convergent series of random functions $f_n \cdot \xi_n$, where f_n is a deterministic function and ξ_n a Gaussian variable of law $\mathcal{N}(0, 1)$ (i.e. with null mean and unitary variance). Assuming the coefficients of the decomposition to be included in the definition of f_n, the identity

$$X_t = \sum_{n=0}^{\infty} f_n(t) \cdot \xi_n = \lim_{N\to\infty} \sum_{n=0}^{N} f_n(t) \cdot \xi_n$$

shows X_t as the limit of a sequence of finite processes $\sum_{n=0}^{N} f_n(t) \cdot \xi_n$. Depending on the nature of the convergence, this may result in two advantages. Analytically, it can be more tractable to prove mathematical results about continuous random processes by considering their discrete representation: it is noticeable when computing quantities such as the characteristic functional of random processes [19]. Numerically, the quantity $\sum_{n=0}^{N} f_n(t) \cdot \xi_n$ can be accurately computed: it provides us with an exact schema to simulate sample paths values at the points where the functions f_n becomes zero if n is large enough.

3 The Discrete Representation of the Wiener Process

3.1 The Haar-like Construction

In view of constructing a discrete representation of the Wiener process, we first introduce the Schauder functions [20, 21]. They are derived from the Haar system, which is the set of functions $h_{n,k}$ in $L^2([0, 1])$ defined as

$$h_{n,k}(t) = \begin{cases} 2^{\frac{n-1}{2}} & \text{if } (2k)2^{-n} \le t < (2k+1)2^{-n}, \\ -2^{\frac{n-1}{2}} & \text{if } (2k+1)2^{-n} \le t < 2(k+1)2^{-n}, \\ 0 & \text{otherwise} \end{cases}$$

for $n \ge 1$ with the addition of the function $h_{0,0}(t) = 1$ on $[0, 1]$. The Haar system has several interesting properties. First, the functions $h_{n,k}$ form a complete orthonormal basis of

$L^2([0,1])$ for the scalar product $(f,g) = \int_0^1 f(t)g(t)dt$. Second, each element $h_{n,k}$ has a compact support

$$S_{n,k} = [k \cdot 2^{-n+1}, (k+1)2^{-n+1}]$$

and, for a given n, the collection of supports $S_{n,k}$ represents a partition of $[0,1]$ up to the endpoints. Third, the functions $h_{n,k}$ build up a wavelet basis of $L^2([0,1])$, since we have the scale-invariant construction rule

$$h_{n,k}(t) = 2^{\frac{n-1}{2}} \cdot h_{1,0}(2^{n-1}t - k). \tag{4}$$

Such properties prove useful to decompose simple Gaussian processes as related in the following. The Schauder functions are defined as the indefinite integrals of the Haar functions $h_{n,k}$ and we denote them as

$$\Psi_{n,k}(t) = \sqrt{\Gamma} \cdot \int_0^1 \chi_{[0,t]}(u) h_{n,k}(u) du, \tag{5}$$

with the help of the indicator functions given by

$$\chi_{[0,t]}(u) = \begin{cases} 1 & \text{if } 0 \le u \le t, \\ 0 & \text{otherwise.} \end{cases}$$

For a fixed $n > 0$, the functions $\Psi_{n,k}$ are little tents of heights $\sqrt{\Gamma} \cdot 2^{-\frac{n+1}{2}}$ with non-overlapping support $S_{n,k}$ for different values of k. The first elements of the so-defined basis are shown on Fig. 1.

We can now proceed to the actual discrete construction of the Wiener process. For $N \ge 0$, we use the Schauder functions to form the process W_t^N defined on Ω as the finite sum

$$W_t^N = \sum_{n=0}^{N} \sum_{0 \le k < 2^{n-1}} \Psi_{n,k}(t) \cdot \xi_{n,k},$$

where $\xi_{n,k}$ are independent random variables of normal law $\mathcal{N}(0,1)$ on Ω. It is worth noticing that for every ω in Ω, the sample path $t \mapsto W_t^N(\omega)$ belongs to the set of continuous functions on $[0,1]$ denoted $C([0,1])$. Indeed, observing that $\sqrt{\Gamma} \cdot 2^{-\frac{n+1}{2}}$ is an upper bound of the Schauder functions $\Psi_{n,k}$, it can be shown [20, 21] that, for almost every ω in Ω, the sample path $t \mapsto W_t^N(\omega)$ converges normally and uniformly in t to a function $t \mapsto \overline{W}_t(\omega)$ when N goes to infinity. As every sample paths is continuous, the limit function $t \mapsto \overline{W}_t(\omega)$ results to be in $C([0,1])$. This allows to define on Ω a limit process $\overline{W} = \lim_{N \to \infty} W^N$ with continuous paths.

If $\overline{W}$ proves to be a Wiener process, we will have a valid discrete representation of the Wiener process. Being defined as the limit of Gaussian processes W^N, we know that $\overline{W}$ is also a Gaussian process. Therefore, showing that $\overline{W}$ is a Wiener process only amounts to demonstrate that it has the same law of covariance as a Wiener process [20, 21], i.e. $\langle \overline{W}_t \cdot \overline{W}_s \rangle = \Gamma \cdot \min(t,s)$, where $\min(t,s)$ is the minimum of t and s. The calculation of the covariance of the limit process $\overline{W}$ is the central point to validate the discrete representation of the Wiener process and needs to be carefully detailed.

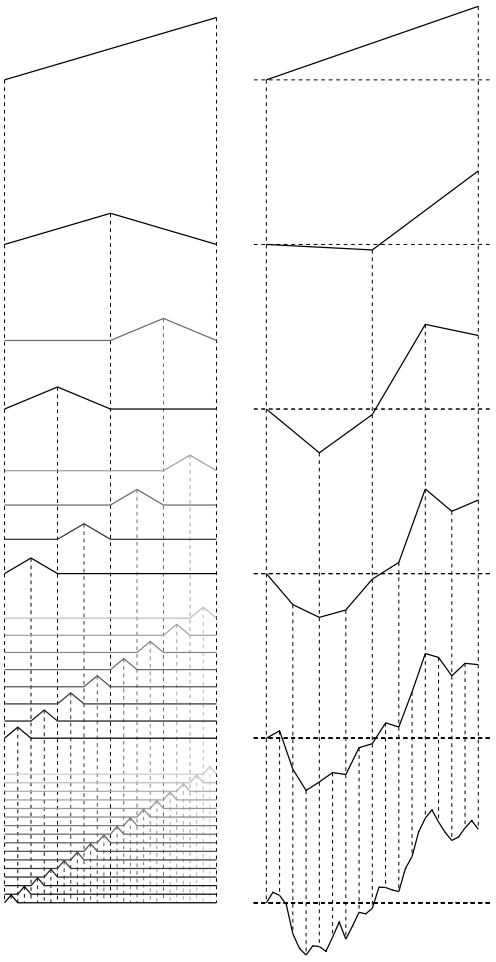

Fig. 1 In the *left column*, the elements of the basis $\Psi_{n,k}$ are represented for each rank n with $0 \leq n < 6$. In the *right column*, the partial sums $W^n(\omega)$ are shown for a given set of realizations ω. Note that each element $\Psi_{n,k}$ has a compact support delimited by dyadic numbers in $D_n = \{k2^{-n} | 0 \leq k \leq 2^n\}$ and that all $\Psi_{n',k}$ is zero on D_n for $n' > n$

3.2 Covariance Calculus

According to the preceding discussion, we need to evaluate the quantity

$$\langle \overline{W}_t \cdot \overline{W}_s \rangle = \lim_{N\to\infty} \langle W_t^N \cdot W_s^N \rangle = \lim_{N\to\infty} \sum_{n=0}^{N} \sum_{0\leq k<2^{n-1}} \Psi_{n,k}(t)\Psi_{n,k}(s)$$

which entails the calculation of a rather tedious series. Indeed, for a given t, at each step n, there is only one k for which $\Psi_{n,k}(t)$ is nonzero, and expressing the series analytically results in a complicated operation. One way to overcome the issue is to notice that the expected covariance result can be expressed in terms of

$$\min(t, s) = \int_0^1 \chi_{[0,t]}(u)\chi_{[0,s]}(u)du. \tag{6}$$

The right term of (6) is actually the scalar product of the functions $\chi_{[0,t]}$ and $\chi_{[0,s]}$ in the Hilbert space $L^2([0, 1])$. As the Haar functions form a Hilbert basis of $L^2([0, 1])$, we can

use the Parseval theorem to write the scalar product of two given functions as

$$\int_0^1 f(t)g(t)dt = \sum_{\substack{n\geq 0 \\ 0\leq k<2^{n-1}}} \int_0^1 f(u)h_{n,k}(u)du \int_0^1 g(u)h_{n,k}(u)du. \tag{7}$$

When applied to the indicator functions of interest $\chi_{[0,t]}$ and $\chi_{[0,s]}$, the relation (7) leads to

$$\begin{aligned} \Gamma \cdot \min(t,s) &= \Gamma \cdot \int_0^1 \chi_{[0,t]}(u)\chi_{[0,s]}(u)du \\ &= \sum_{n=0}^{\infty} \sum_{0\leq k<2^{n-1}} \Psi_{n,k}(t)\Psi_{n,k}(s), \end{aligned}$$

since the definition (5) describes $\Psi_{n,k}(t)$ as the coefficient relative to $h_{n,k}$ in the decomposition of $\chi_{[0,t]}$ on the Haar system. We can finally recap the result

$$\begin{aligned} \langle \overline{W}_t \cdot \overline{W}_s \rangle &= \lim_{N\to\infty} \langle W_t^N \cdot W_s^N \rangle \\ &= \lim_{N\to\infty} \sum_{n=0}^{N} \sum_{0\leq k<2^{n-1}} \Psi_{n,k}(t)\Psi_{n,k}(s) = \Gamma \cdot \min(t,s), \end{aligned}$$

establishing the discrete description of the Wiener process as a normally convergent series of terms $\Psi_{n,k} \cdot \xi_{n,k}$, where $\Psi_{n,k}$ is a Haar-derived function and $\xi_{n,k}$ a random variable of normal law $\mathcal{N}(0,1)$.

4 The Rationale of the Construction for the Ornstein-Uhlenbeck Process

4.1 Comparison of the Wiener Process and the Ornstein-Uhlenbeck Process

We recall that the Langevin equation (1) can be solved by quadratures in simple cases. In the framework of stochastic integration, we are provided with rigorous integral expressions for these solutions. If the process is at x_0 when $t = 0$, the Ornstein-Uhlenbeck process U_t is expressed

$$U_t = x_0 e^{-\alpha t} + \int_0^t e^{\alpha(u-t)} dW_u, \tag{8}$$

as opposed to the Wiener process W_t in the same conditions

$$W_t = x_0 + \int_0^t dW_u. \tag{9}$$

The comparison of definitions (8) and (9) explains why finding a basis of decomposition for U is a difficult task. Indeed, the process U is not anymore a simple integral of "white Gaussian noise elements" dW_u: the exponential factor in (8) indicates that the contribution of the dW_u in the integral expression of U depends on the position of u compared to t. It is also apparent that, due to the exponential modulation, the Ornstein-Uhlenbeck process does not exhibit scale-invariant properties. As a consequence, it is unlikely for the putative basis of decomposition to be build from an orthogonal wavelet basis.

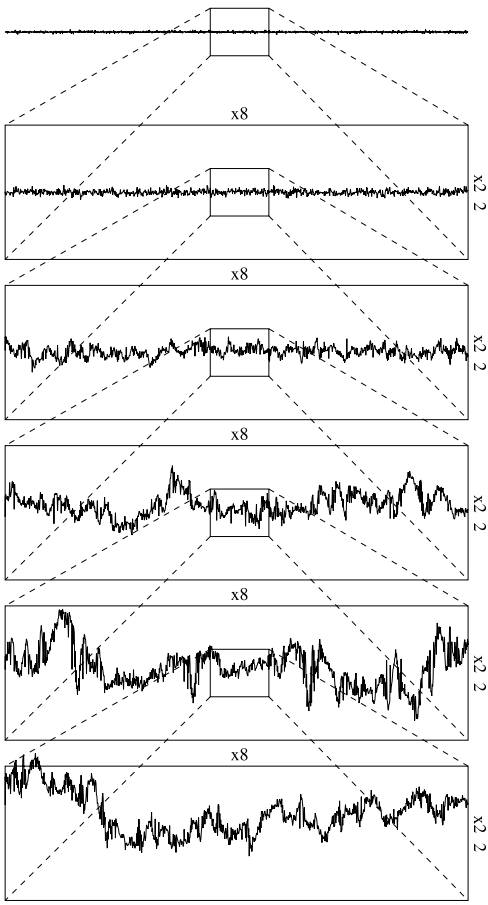

Fig. 2 A sample path $t \mapsto U_t(\omega)$ is represented at different magnifications following the scale invariance of a Wiener process: the vertical zooming factor is the square root of the horizontal factor. Note that the sample path $t \mapsto U_t(\omega)$ behaves as a Wiener process at small scales

Yet, as noticeable in Fig. 2, the examination of a sample path $t \mapsto U_t(\omega)$ reveals the scale-invariant behavior of a Wiener process for asymptotically small time scale as well as for asymptotically small α. It means that the basis of decomposition $\Psi_{n,k}$ for the Wiener process is asymptotically valid to describe U_t at fine scale. This observation suggests that, upon slight alteration of its analytical expression, the Haar derived basis $\Psi_{n,k}$ can give rise to a basis $\Phi_{n,k}$ adapted to the Ornstein-Uhlenbeck process. The change in the analytical expression of $\Psi_{n,k}$ should be consistent with the previously mentioned difficulties, preventing its formulation to be scale invariant or orthogonal. Under this restraint, the fundamental property that each element $\Psi_{n,k}$ exhibits a compact support of the form $S_{n,k}$ should be preserved in the expression of $\Phi_{n,k}$.

4.2 The Markov Property of the Wiener Process and the Ornstein-Uhlenbeck Process

To carry out this program, the key point is to consider $p(X_{t_y} = y | X_{t_x} = x, X_{t_z} = z)$ with $t_x < t_y < t_z$, the probability density of X_{t_y} knowing its values x and z at two framing times t_x and t_z. Because X is a Markovian process, a sample path $t \mapsto X_t(\omega)$ which originates from x and joins z through y is just the junction of two independent paths: a path originating in x going to y and a path originating from y going to z. Assuming conditional knowledge of its

origin x, the probability of such a compound path is the product of the probability of the two elementary paths with conditional knowledge of their respective origins x and y. Therefore, after normalization by the absolute probability for a path to go from x to y, the probability density $p(X_{t_y} = y|X_{t_x} = x, X_{t_z} = z)$ is expressed in the following expression

$$p(X_{t_y} = y|X_{t_x} = x, X_{t_z} = z) = \frac{p(X_{t_y} = y|X_{t_x} = x) \cdot p(X_{t_z} = z|X_{t_y} = y)}{p(X_{t_z} = z|X_{t_x} = x)}. \tag{10}$$

It is now a simple matter of calculation to compute the density of X_{t_y} knowing $X_{t_x} = x$ and $X_{t_z} = z$ with the analytical expression of the probability $p(X_{t_y} = y|X_{t_x} = x)$. In the case of a Gaussian process, it is expected to follow a normal law, which we refer to as $\mathcal{N}({}_X\mu(t_y), {}_X\sigma(t_y))$. For the Wiener process, using expression (2) for $p(X_{t_y} = y|X_{t_x} = x)$, the mean value ${}_W\mu(t_y)$ and the variance ${}_W\sigma^2(t_y)$ result in

$${}_W\mu(t_y) = {}_W\mu_{t_x,t_z}(t_y, x, z) = \frac{t_z - t_y}{t_z - t_x} \cdot x + \frac{t_y - t_x}{t_z - t_x} \cdot z, \tag{11}$$

$${}_W\sigma^2(t_y) = {}_W\sigma^2_{t_x,t_z}(t_y) = \Gamma \cdot \frac{(t_y - t_x)(t_z - t_y)}{t_z - t_x}. \tag{12}$$

For the Ornstein-Uhlenbeck process, using expression (3) for $\mathbf{P}(X_{t_y} = y|X_{t_x} = x)$ similarly yields the mean ${}_U\mu(t_y)$ and the variance ${}_U\sigma(t_y)^2$ as shown in supplementary materials:

$${}_U\mu(t_y) = {}_U\mu_{t_x,t_z}(t_y, x, z) = \frac{\sinh(\alpha(t_z - t_y))}{\sinh(\alpha(t_z - t_x))} \cdot x + \frac{\sinh(\alpha(t_y - t_x))}{\sinh(\alpha(t_z - t_x))} \cdot z, \tag{13}$$

$${}_U\sigma^2(t_y) = {}_U\sigma^2_{t_x,t_z}(t_y) = \frac{\Gamma}{2\alpha} \cdot \frac{2 \cdot \sinh(\alpha(t_y - t_x)) \cdot \sinh(\alpha(t_z - t_y))}{\sinh(\alpha(t_z - t_x))}. \tag{14}$$

In the limit of very short time scale or vanishing α, we notice that ${}_U\mu(t_y)$ and ${}_U\sigma^2(t_y)$ approximate ${}_W\mu(t_y)$ and ${}_W\sigma^2(t_y)$.

4.3 The Conditional Averages of the Processes on Dyadic Ensembles

We note D_N the set of reals $\{k2^{-N} \mid 0 \leq k \leq 2^N\}$ and we have $\{0, 1\} = D_0 \subset D_1 \subset \cdots \subset D_N$ a growing sequence of sets with limit ensemble $\mathcal{D}$ the set of dyadic points in $[0, 1]$. Let us consider $\langle X_t \rangle_{D_N}$ the conditional expectation of the random variable X_t given $\{X_t\}_{t \in D_N}$. The collection of random variables $\langle X_t \rangle_{D_N}$ defined on Ω specify a continuous stochastic process $\langle X \rangle_{D_N}$ on Ω. For a Wiener process, (11) shows that a sample path $t \mapsto \langle W_t \rangle_{D_N}(\omega)$ is a piece-wise linear function of t interpolating each points of D_N; whereas for an Ornstein-Uhlenbeck process, (13) depicts a sample path $t \mapsto \langle U_t \rangle_{D_N}$ as a succession of catenaries joining successive points of D_N. With $0 \leq k < 2^{-N}$, if $t_x = k2^{-N}$ and $t_z = (k + 1)2^{-N}$ are the two successive points of D_N framing t, the average $\langle X_t \rangle_{D_N}$ is only conditioned by the random variables X_{t_x} and X_{t_z}. For the sake of simplicity, we write for a given realization ω in Ω

$$\langle X_t \rangle_{D_N}(\omega) = \langle X_t \rangle_{x,z} = {}_X\mu_{t_x,t_z}(t, x, z) \stackrel{\text{def}}{=} {}_X\mu^N(t), \tag{15}$$

where the conditional dependency upon the outcomes $X_{t_x} = x$ and $X_{t_z} = z$ is implicit in ${}_X\mu^N$.

We want to investigate the change in the estimation of X_t due to the conditional knowledge of its value on the dyadic set D_{N+1}. In that perspective, we exemplified the conditional

expectation $\langle X_t \rangle_{D_{N+1}}$ on $[t_x, t_z]$ where the estimation of X_t is now dependent upon the outcome of X_{t_y} with t_y the midpoint of t_x and t_z: for every ω in Ω, we write

$$\langle X_t \rangle_{D_{N+1}}(\omega) = \langle X_t \rangle_{x,y,z} = \begin{cases} \langle X_t \rangle_{x,y} & \text{if } t_x \le t \le t_y, \\ \langle X_t \rangle_{y,z} & \text{if } t_y \le t \le t_z \end{cases}$$
$$\stackrel{\text{def}}{=} {}_X\nu^N(t, y), \tag{16}$$

where we underline the dependency on y, due to the additional conditioning on $D_{N+1} \setminus D_N$.

Now, let us assume that we only have conditional knowledge of the process on D_N. By the Markov property, the value y in (16) is the outcome of X_{t_y} given X_{t_x} and X_{t_z}, whose statistics is described by the law $\mathcal{N}({}_X\mu(t_y), {}_X\sigma(t_y))$. We use the short notation $Y_{N,k}$ to designate a Gaussian variable following such a law. Remembering that we only have knowledge of the process on D_N, we remark that the conditional law of $\langle X_t \rangle_{D_{N+1}}$ is the same as the law of the random variable ${}_X\nu^N(t, Y_{N,k})$. It represents, through the function ${}_X\nu^N$, the random contribution of ignoring $X_{t_y} = y$ when one estimates the process knowing its values on t_x and t_z.

4.4 Identification of the Conditional Averages with the Partial Sums

The results above allows to gain insight in the building of a Wiener process W as the converging series of random functions $\Psi_{n,k} \cdot \xi_{n,k}$. It is easy to see from the definition (4) that $\Psi_{n,k}$ is linear between any two points in D_n for $n \le N$ and that $\Psi_{n,k}$ is zero on D_n for every $n > N$. In other words, the partial sum

$$W_t^N = \sum_{n=0}^{N} \sum_{0 \le k < 2^n - 1} \Psi_{n,k}(t) \cdot \xi_{n,k} \quad \text{for } t \in D_N$$

coincide with W_t on D_N and more generally with $\langle X_t \rangle_{D_N}$ on $[0, 1]$. Identifying partial sums with conditional averages, it is then straightforward to express for every ω in Ω the component $\Psi_{N+1,k}(t) \cdot \xi_{N+1,k}$ in the decomposition of W_t

$$\begin{aligned} \Psi_{N+1,k}(t) \cdot \xi_{N+1,k}(\omega) &= W_t^{N+1}(\omega) - W_t^N(\omega) \\ &= \langle W_t \rangle_{D_{N+1}}(\omega) - \langle W_t \rangle_{D_N}(\omega) \\ &= {}_W\nu^N(t, y) - {}_W\mu^N(t). \end{aligned} \tag{17}$$

We bear in mind the previous definitions for which $[t_x, t_z] = [k2^{-N}, (k+1)2^{-N}]$ is the support $S_{N+1,k}$ of $\Psi_{N+1,k}$ and y is the outcome of X_{t_y} with t_y the midpoint of t_x and t_z. If we only assume conditional knowledge on D_N, ${}_W\nu^N(t, y) - {}_W\mu^N(t)$ is distributed according to the same law as of the law of the random variable ${}_W\nu^N(t, Y_{N,k}) - {}_W\mu^N(t)$. We deduce that, conditionally to the values on D_N, $\Psi_{N+1,k}(t) \cdot \xi_{N+1,k}$ has the same law as ${}_W\nu^N(t, Y_{N,k}) - {}_W\mu^N(t)$.

The tight connection between $\xi_{N+1,k}$ and $Y_{N,k}$ is made obvious: if one knows the values of the process on D_N, the random contribution of $\sum_k \Psi_{N+1,k}(t) \cdot \xi_{N+1,k}$ conveys the uncertainty about W_t that is discarded by the knowledge of its values on $D_{N+1} \setminus D_N$. In this regard, we stress the fact that the conditional law of $\xi_{N+1,k}$ knowing the values of the $\xi_{n,k}$ for $n \le N$ is again $\mathcal{N}(0, 1)$ by independence of the $\xi_{n,k}$.

We are now in a position to complete our program by continuing the identification of partial sums and conditional average for the Ornstein-Uhlenbeck process U. As a basis of decomposition, we propose the set of functions $\Phi_{n,k}$ defined on $S_{n,k}$ with $n \geq 0$ and $0 < k \leq 2^{n-1}$ by the following criteria: for a given N, the function $\Phi_{N+1,k}$ is the continuous positive function on $S_{N+1,k}$ such that the random variable $\Phi_{N+1,k}(t) \cdot \xi_{N+1,k}$ has the same law as the law of ${}_U\nu^N(t, Y_{N,k}) - {}_U\mu^N(t)$. The previous criteria assumes the adapted definitions of ${}_U\mu^N$ and ${}_U\nu^N$ on $S_{N+1,k}$, the support of the investigated functions $\Phi_{N+1,k}$. We underline that the notation $Y_{N,k}$ refers here to the random variable U_t at the midpoint of the support $t = (2k+1)2^{-(N+1)}$ knowing its values on the extremities.

5 A Discrete Construction for the Ornstein-Uhlenbeck Process

5.1 The Candidate Basis for a Discrete Representation

In view of representing an Ornstein-Uhlenbeck process as a discrete process, the comparison with a Wiener process suggests a candidate basis of decomposition of the form $\Phi_{n,k} \cdot \xi_{n,k}$, the variable $\xi_{n,k}$ following the law $\mathcal{N}(0, 1)$. The deterministic function $\Phi_{n,k}$ is defined with support $S_{n,k} = [k \cdot 2^{-n+1}, (k+1)2^{-n+1}]$ for $n > 0$ with $0 \leq 2k < 2^n$. We use expressions (13) and (14) to make explicit the formulation of $\Phi_{n,k}$ and we obtain

$$\Phi_{n,k}(t) = \begin{cases} \frac{\sqrt{\Gamma}\cdot\sinh(|\alpha|(t-2k\cdot 2^{-n}))}{\sqrt{\alpha\cdot\sinh(\alpha 2^{-n+1})}} & \text{if } (2k)2^{-n} \leq t < (2k+1)2^{-n}, \\ \frac{\sqrt{\Gamma}\cdot\sinh(|\alpha|(2(k+1)2^{-n}-t))}{\sqrt{\alpha\cdot\sinh(\alpha 2^{-n+1})}} & \text{if } (2k+1)2^{-n} \leq t < 2(k+1)2^{-n}, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Without any further comment, the element $\Phi_{0,0}$ is defined as

$$\Phi_{0,0}(t) = \frac{\sqrt{\Gamma}\cdot e^{-\frac{\alpha}{2}}\sinh(|\alpha|t)}{\sqrt{\alpha\cdot\sinh\alpha}}, \tag{19}$$

a choice we will explain in Sect. 6.1. The first elements $\Phi_{n,k}$ are shown in Fig. 3. As expected, they are only asymptotically scale-invariant but they exhibit the desirable property of being compactly supported on $S_{n,k}$, the interval between two dyadic points $k \cdot 2^{-n+1}$ and $(k+1)2^{-n+1}$.

To validate the decomposition of an Ornstein-Uhlenbeck process U_t on the set of functions $\Phi_{n,k}$, we need to study the convergence of the partial sums

$$U_t^N = \sum_{n=0}^{N} \sum_{0 \leq k < 2^{n-1}} \Phi_{n,k}(t) \cdot \xi_{n,k}.$$

As each function $\Phi_{n,k}$ is dominated by $\sqrt{\Gamma} \cdot e^{|\alpha|} 2^{-\frac{n+1}{2}}$, the same argument as for the Wiener process entails the normal and uniform convergence in t of the sample path $t \mapsto U_t^N(\omega)$ almost surely on the sample space Ω. We then denote $\overline{U}$ the limit process defined almost surely on Ω as $\overline{U} = \lim_{N\to\infty} U_t^N$. The normal convergence causes $\overline{U}$ to have continuous paths; being a sum of Gaussian variables, it is also a Gaussian process. Therefore, proving

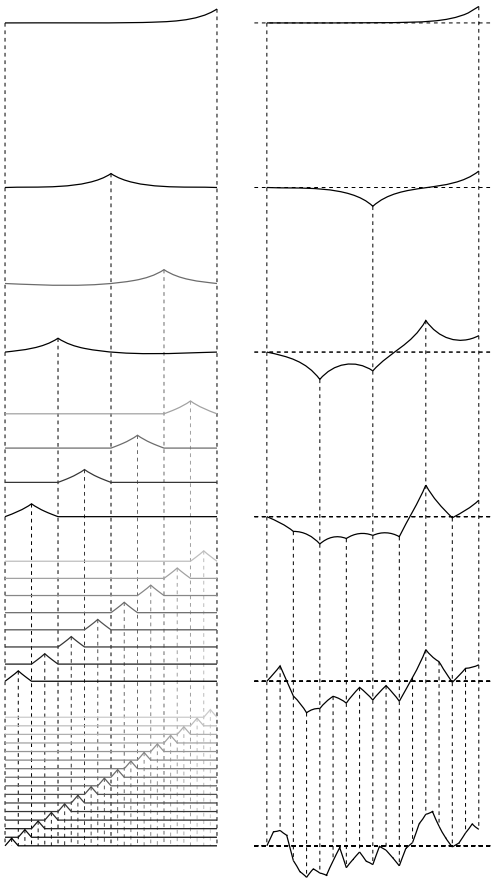

Fig. 3 In the *left column*, the elements of the basis $\Phi_{n,k}$ are represented for each rank n with $0 \leq n < 6$. In the *right column*, the conditional Ornstein-Uhlenbeck process $\langle U_t \rangle_{D_n}$ is shown for a given set of realizations on D_n. Once more, note that each element $\Psi_{n,k}$ has a compact support delimited by dyadic numbers in $D_n = \{k2^{-n} \mid 0 \leq k \leq 2^n\}$ and that all $\Phi_{n',k}$ is zero on D_n for $n' > n$

that $\overline{U}$ is an Ornstein-Uhlenbeck process just requires us to show that the covariance of $\overline{U}$ satisfies

$$\langle \overline{U}_t \cdot \overline{U}_s \rangle = \frac{\Gamma}{2\alpha} \cdot e^{-\alpha(t+s)}(e^{2\alpha \min(t,s)} - 1). \tag{20}$$

To establish this relation, we need to evaluate the covariance of $\overline{U}$ as the limit covariance of the partial sums

$$\langle U_t^N \cdot U_s^N \rangle = \sum_{n=0}^{N} \sum_{0 \leq k < 2^{n-1}} \Phi_{n,k}(t)\Phi_{n,k}(s).$$

5.2 Covariance Calculus

It is possible to simplify the above expression, even though the functions $\Phi_{n,k}$ are not orthogonal. For each given n, the disjoint supports of $\Phi_{n,k}$ form a partition of $[0, 1]$ as a collection of segments $S_{n,k}$ of equal length 2^{-n+1}. Considering a real t in $[0, 1]$, there is only one sequence of indices k_n such that t belongs to each support of S_{n,k_n}. The succession of k_n represents t as the intersection of decreasing dyadic segments $\bigcap_{n=0}^{\infty} S_{n,k_n}$, which can be

explained in terms of the binary representation $t = \sum_1^\infty a_i 2^{-i}$, $a_i \in \{0, 1\}$, if we exclude inappropriate infinite developments. Bearing in mind the system of indexing for $S_{n,k}$, a simple recurrence argument leads to the expression of k_n corresponding to a given t in its binary representation

$$k_n = \frac{1}{2} \cdot \sum_{i=1}^{n-1} a_i 2^{n-i}. \tag{21}$$

We are now in a position to write the reduced expression of the partial sums

$$U_t^N = \sum_{n=0}^{N} \Phi_{n,k_n}(t) \cdot \xi_{n,k_n}$$

where the terms $\Phi_{k_n,n}(t)$ is made explicit using the previous formulation of k_n in the definition (18)

$$\Phi_{n,k_n}(t) = \begin{cases} \frac{\sqrt{\Gamma} \cdot \sinh\left(|\alpha| \sum_{n+1}^{\infty} a_i 2^{-i}\right)}{\sqrt{\alpha \cdot \sinh\left(\alpha 2^{-n+1}\right)}} & \text{if } a_n = 0, \\ \frac{\sqrt{\Gamma} \cdot \sinh\left(|\alpha| \sum_{n+1}^{\infty} (1-a_i) 2^{-i}\right)}{\sqrt{\alpha \cdot \sinh\left(\alpha 2^{-n+1}\right)}} & \text{if } a_n = 1. \end{cases}$$

Informed by these preliminaries, we shall carry out the calculation of the covariance. The reduced formulation of partial sums allows us to write

$$\langle U_t^N \cdot U_s^N \rangle = \sum_{n=1}^{N} \Phi_{n,k_n}(t) \Phi_{n,l_n}(s) + \Phi_{0,0}(t) \Phi_{0,0}(s) \tag{22}$$

where the indices k_n and l_n designate the sequence of functions Φ_{n,k_n} and Φ_{n,l_n} whose supports contain t and s respectively. When t and s are distinct, we notice that for $n > 1 - \log_2 |t - s|$, the supports $S_{k_n,n}$ and $S_{l_n,n}$ containing t and s respectively are disjoint, so that the cross-products $\Phi_{k_n,n}(t)\Phi_{l_n,n}(s)$ cancel out if n is large enough. It is then possible to write expression (22) as a finite sum where the terms $\Phi_{k_n,n}(t)$ and $\Phi_{l_n,n}(s)$ are specified due to the binary representations $t = \sum_1^\infty a_i 2^{-i}$ and $s = \sum_1^\infty b_i 2^{-i}$. We specify that we only consider proper binary representations, that is, the binary representation of dyadic points is chosen in its finite form. For the sake of simplicity, we assume that $t < s$. Formulated in the binary representation, the order $t < s$ is equivalent to the existence of a natural $N_0 > 0$ such that $a_n = b_n$ as long as $n < N_0$ and $a_{N_0} < b_{N_0}$, that is $a_{N_0} = 0$ and $b_{N_0} = 1$. With the preceding remarks, it is clear that $S_{k_n,n}$ and $S_{l_n,n}$ are disjoint for $n > N_0$ and we can write the covariance of U_t^N for $N > N_0$ in the explicit form

$$\langle U_t^N \cdot U_s^N \rangle_{N>N_0} = \frac{\Gamma}{2\alpha} \left(\sum_{n=1}^{N_0} \frac{2 \cdot u_n}{\sinh\left(\alpha 2^{-n+1}\right)} + e^{-\alpha} \frac{2 \sinh(\alpha t) \sinh(\alpha s)}{\sinh \alpha} \right). \tag{23}$$

The variable u_n apparent in (23) represents for $n < N_0$ the numerator of the cross-products $\Phi_{k_n,n}(t)\Phi_{l_n,n}(s)$ when the extension a_n and b_n coincide

$$u_n = \begin{cases} \sinh\left(\alpha \sum_{n+1}^{\infty} a_i 2^{-i}\right) \sinh\left(\alpha \sum_{n+1}^{\infty} b_i 2^{-i}\right) & \text{if } a_n = b_n = 0, \\ \sinh\left(\alpha \sum_{n+1}^{\infty} (1-a_i) 2^{-i}\right) \sinh\left(\alpha \sum_{n+1}^{\infty} (1-b_i) 2^{-i}\right) & \text{if } a_n = b_n = 1. \end{cases}$$

As for the limit case $n = N_0$, u_{N_0} expresses the numerator of the cross-product $\Phi_{k_{N_0},N_0}(t)\Phi_{l_{N_0},N_0}(s)$ with $a_{N_0} = 0$ and $b_{N_0} = 1$

$$u_{N_0} = \sinh\left(\alpha \sum_{N_0+1}^{\infty} a_i 2^{-i}\right) \sinh\left(\alpha \sum_{N_0+1}^{\infty} (1 - b_i) 2^{-i}\right).$$

At that point, the explicit form of the covariance (23) results in a rather complicated combination of hyperbolic functions. Fortunately enough, we can resort to using remarkable identities to simplify its expression. The solution actually lies in the consideration of the quantity

$$v_n = \sinh\left(\alpha \sum_{n}^{\infty} a_i 2^{-i}\right) \sinh\left(\alpha \sum_{n}^{\infty} (1 - b_i) 2^{-i}\right). \tag{24}$$

We show in the supplementary materials that, as long as $n < N_0$, v_n verifies the recurrence relation

$$v_n = 2\cosh(\alpha 2^{-n}) \cdot v_{n+1} + u_n. \tag{25}$$

We can express u_n in terms of v_n and v_{n+1} to compute the following series by cancellation term by term

$$\sum_{n=1}^{N_0-1} \frac{u_n}{\sinh(\alpha 2^{-n+1})} = \frac{v_1}{\sinh\alpha} - \frac{v_{N_0}}{\sinh(\alpha 2^{-N_0+1})}. \tag{26}$$

Remembering that $a_{N_0} = 0$ and $b_{N_0} = 1$, we remark that $v_{N_0} = v_{N_0+1} = u_{N_0}$ so that the insertion of (26) in expression (23) caused the remaining terms in u_{N_0} to cancel out. It is then straightforward to write the covariance

$$\langle U_t^N \cdot U_s^N \rangle_{N>N_0} = \frac{\Gamma}{2\alpha}\left(\frac{2 \cdot v_1}{\sinh\alpha} + e^{-\alpha}\frac{2\sinh(\alpha t)\sinh(\alpha s)}{\sinh\alpha}\right). \tag{27}$$

We observed that the definition of v_1 invokes the full binary representations of t and s so that we have $v_1 = \sinh(\alpha t)\sinh(\alpha(1-s))$. After a several manipulations, expression (27) finally yields

$$\langle U_t^N \cdot U_s^N \rangle_{N>N_0} = \frac{\Gamma}{2\alpha} \cdot e^{-\alpha(t+s)}(e^{2\alpha t} - 1),$$

which is the expected result for the covariance of an Ornstein-Uhlenbeck process (20) given that $t = \min(t, s)$ as $t < s$.

Regarding the calculation of the variance when $t = s$, the series of cross-products $\Phi_{k_n,n}(t)\Phi_{l_n,n}(s) = \Phi^2_{k_n,n}(t)$ becomes infinite, but fortunately the recurrence relation (25) is then valid for every $n > 0$. As the quantity v_n vanishes when n grows to infinity, the cancellation term by term is still effective to compute the series in (23). It leads to the expected variance expression for an Ornstein-Uhlenbeck process ${}_U\sigma_t^2 = \frac{\Gamma}{2\alpha} \cdot (1 - e^{-\alpha t})$.

We finally recap the result for any t and s without assuming any order

$$\langle \overline{U}_t \cdot \overline{U}_s \rangle = \lim_{N\to\infty} \langle U_t^N \cdot U_s^N \rangle = \frac{\Gamma}{2\alpha} \cdot e^{-\alpha(t+s)}(e^{2\alpha \min(t,s)} - 1).$$

It proves the discrete description of an Ornstein-Uhlenbeck processes as the normally convergent series of random functions $\Phi_{n,k} \cdot \xi_{n,k}$, where $\Phi_{n,k}$ is a deterministic function defined in (5) and $\xi_{n,k}$ a random variable of normal law $\mathcal{N}(0, 1)$.

6 Bi-infinite Representation on the Set of Positive Reals

6.1 The Choice of the Basis Function

Whether standing for a Wiener process or an Ornstein-Uhlenbeck process, X can be decomposed in a discrete basis of functions $f_{n,k}$, where $f_{n,k}$ is a generic notation for the deterministic functions $\Psi_{n,k}$ and $\Phi_{n,k}$. The identification of the partial sums

$$X_t^N = \sum_{n=0}^{n=N} \sum_{0 \le k < 2^{n-1}} f_{n,k}(t) \cdot \xi_{n,k}$$

with the conditional average $\langle X_t \rangle_{D_N}$ suggests to consider the discrete representation of X as a recurrence construction, a view that explains how to chose the first element of the basis $f_{0,0}$.

First, we formulate the inductive step of the recurrence argument. Suppose we know at rank N that X_t^N and $\langle X_t \rangle_{D_N}$ have the same probability distribution on Ω, we want to show that X_t^{N+1} and $\langle X_t \rangle_{D_{N+1}}$ follows the same probability. Let us assume prior knowledge of X_t on D_N. If we consider a given time t, there exists a unique k such that $k2^{-N} \le t < (k+1)2^{-N}$ and we know that the collection of segments $S_{N+1,k} = [k2^{-N}, (k+1)2^{-N}]$ for $0 \le k < 2^N$ defines a partition of $[0, 1]$ up to the endpoints. We also remark that $t_{N+1,k} = (2k+1)2^{-(N+1)}$ is the only point of $D_{N+1} \setminus D_N$ in $S_{N+1,k}$. According to the results exposed in the second section, we know that the outcome of $X_{t_{N+1,k}}$ given $X_{t_{N,k}}$ and $X_{t_{N,k+1}}$ follows the same law as

$$_X\sigma(t_{N+1,k}) \cdot \xi_{N+1,k} + {}_X\mu(t_{N+1,k})$$

where $\xi_{N+1,k}$ is of normal law $\mathcal{N}(0, 1)$. As X is a Markov process, the outcome of $\langle X_t \rangle_{D_{N+1}} - \langle X_t \rangle_{D_N}$ only depends on the outcome of $X_{t_{N+1,k}}$ given $X_{t_{N,k}}$ and $X_{t_{N,k+1}}$ when restricted on the support $S_{n+1,k}$. Due to the simplicity of the situation, it is possible to find on $S_{N+1,k}$ an analytical expression of the form $f_{N+1,k} \cdot \xi_{n,k}$ to describe the probability law of $\langle X_t \rangle_{D_{N+1}} - \langle X_t \rangle_{D_N}$ given $\{X_t\}_{t \in D_N}$. Then, by construction of f_{N+1}, if X_t^N and $\langle X_t \rangle_{D_N}$ follows the same distribution, the probability distributions of X_t^{N+1} and $\langle X_t \rangle_{D_{N+1}}$ also coincide on Ω, which finishes the inductive step. Incidentally, we have an interpretation for the statistical contribution of a component $f_{n,k}(t) \cdot \xi_{n,k}$. At each step N,the function $\sum_k f_{N+1,k}(t) \cdot \xi_{N+,k}$ represents the uncertainty about X_t that is discarded by the knowledge of its values on $D_{N+1} \setminus D_N$.

It now remains to verify the basis statement to validate the recurrence argument, that is:

$$X_t^0 = f_{0,0}(t) \cdot \xi_{0,0} = \langle X_t \rangle_{D_0}.$$

Actually, the need to satisfy this prerequisite enforces how to set the expression of $f_{0,0}$. The conditional average $\langle X_t \rangle_{D_0}$ is a function of the value of X_t on $D_0 = \{0, 1\}$. By construction the value of X_t in 0 is assumed to be zero. We note ${}_XZ$ the random function X_1 knowing $X_0 = 0$, and we recall that its statistics is given by relations (2) for a Wiener process and (3) for an Ornstein-Uhlenbeck process respectively. With the notation of Sect. 4.2, we write $\langle X_t \rangle_{D_0}$ as a function of ${}_XZ$

$$\langle X_t \rangle_{D_0} = {}_X\mu_{0,1}(t, 0, {}_XZ). \tag{28}$$

It defines a Gaussian random function $\langle X_t \rangle_{D_0}$ of the form $f_{0,0} \cdot \xi_{0,0}$. The dependency of its variance upon t yields the expression of the deterministic part $f_{0,0}$

$$f_{0,0}(t) = \sqrt{\langle \langle X_t \rangle_{D_0}^2 \rangle}. \tag{29}$$

When applied to the Wiener process, relation (28) reads

$$\langle W_t \rangle_{D_0} = t \cdot {}_W Z$$

and relation (29) gives the right expression for $\Psi_{0,0}$

$$\Psi_{0,0}(t) = \sqrt{\langle {}_W Z \rangle^2 \cdot t^2} = \sqrt{\Gamma} \cdot t.$$

When applied to the Ornstein-Uhlenbeck process, relation (28) yields

$$\langle U_t \rangle_{D_0} = \frac{\sinh(\alpha t)}{\sinh \alpha} \cdot {}_U Z$$

and relation (29) gives the already mentioned expression of $\Phi_{0,0}$

$$\Psi_{0,0}(t) = \sqrt{\langle {}_U Z \rangle^2 \cdot \left(\frac{\sinh(\alpha t)}{\sinh \alpha} \right)^2} = \frac{\sqrt{\Gamma} \cdot e^{-\frac{\alpha}{2}} |\sinh(\alpha t)|}{\sqrt{\alpha \cdot \sinh \alpha}}.$$

6.2 Representation as a Bi-infinite Sum

Now, we further this recurrence description to show that X can be naturally represented as a bi-infinite series of random functions. In that perspective, we extend the definition of the dyadic sets to $D_N = \{k2^{-N} \,|\, k \in \mathbb{Z}\}$ and we have the increasing sequence of sets $D_{-N} = 2^N \mathbb{Z} \subset \cdots \subset D_0 = \mathbb{Z} \subset \cdots \subset D_N = 2^{-N}\mathbb{Z}$. We can easily adapt the rationale of the recurrence construction on $S_{-M,0} = [0, 2^M]$ to build a limit process by summation of the random contributions of $f_{n,k} \cdot \xi_{n,k}$. In order to distinguish this situation from the previous case, we designate ${}_M X^N$ the partial sum at rank $N \geq -M$ when the building is initiated at rank $-M$. The usual definitions are still valid to explicit the basis of functions $f_{n,k}$ for $n > -M$ but the index k now runs satisfying $0 \leq k < 2^{M+n-1}$. As previously, the expression of the basis element denoted $\bar{f}_{-M,0}$ satisfies on $S_{-M,0}$

$${}_M X^{-M} = \bar{f}_{-M,0}(t) \cdot \xi_{-M,0} = \langle X_t \rangle_{D_{-M}}, \tag{30}$$

where the first partial sum is denoted ${}_M X^{-M}$. We posit the general formulation for the partial sums ${}_M X^N$ on $S_{-M,0}$ for $N \geq -M$

$${}_M X_t^N = \sum_{n=-M+1}^{N} \sum_{0 \leq k < 2^{M+n-1}} f_{n,k}(t) \cdot \xi_{n,k} + \bar{f}_{-M,0}(t) \cdot \xi_{-M,0}. \tag{31}$$

We specify these results for our two cases of interest. In the case of a Wiener process, we defined the functions $\Psi_{n,k}$

$$\Psi_{n,k}(t) = \begin{cases} \sqrt{\frac{\Gamma}{2^{-n+1}}} \cdot (t - 2k \cdot 2^{-n}) & \text{if } (2k)2^{-n} \leq t < (2k+1)2^{-n}, \\ \sqrt{\frac{\Gamma}{2^{-n+1}}} \cdot (2(k+1)2^{-n} - t) & \text{if } (2k+1)2^{-n} \leq t < 2(k+1)2^{-n}, \\ 0 & \text{otherwise,} \end{cases}$$

for $n > -M$, and the first element $\bar{\Psi}_{-M,0}$

$$\bar{\Psi}_{-M,0}(t) = \sqrt{\frac{\Gamma}{2^M}} \cdot t. \tag{32}$$

In the case of an Ornstein-Uhlenbeck process, we similarly write the functions $\Phi_{n,k}$

$$\Phi_{n,k}(t) = \begin{cases} \frac{\sqrt{\Gamma}\cdot\sinh(|\alpha|(t-2k\cdot 2^{-n}))}{\sqrt{\alpha\cdot\sinh(\alpha 2^{-n+1})}} & \text{if } (2k)2^{-n} \le t < (2k+1)2^{-n}, \\ \frac{\sqrt{\Gamma}\cdot\sinh(|\alpha|((2(k+1)2^{-n}-t))}{\sqrt{\alpha\cdot\sinh(\alpha 2^{-n+1})}} & \text{if } (2k+1)2^{-n} \le t < 2(k+1)2^{-n}, \\ 0 & \text{otherwise,} \end{cases}$$

for $n > -M$, and the first element $\bar{\Phi}_{-M,0}$

$$\bar{\Phi}_{-M,0}(t) = \frac{\sqrt{\Gamma}\cdot e^{-\alpha 2^{M-1}}\sinh(|\alpha|t)}{\sqrt{\alpha\cdot\sinh\alpha 2^M}}. \tag{33}$$

Expression (31) suggests that the process X admits a natural discrete representation as a bi-infinite series on $\lim_{M\to\infty}\cup_M S_{-M,0} = \mathbb{R}^+$. To this end, we extend the definition of $f_{n,k}$ on $\mathbb{R}^+$ by setting its value to zero outside $S_{-M,0}$ and we form the quantity

$$\widehat{X}_t^N = \sum_{n=-N+1}^{N} \sum_{0\le k<2^{N+n-1}} f_{n,k}(t)\cdot\xi_{n,k}.$$

We want to prove the normal convergence of the sample path $t \mapsto \widehat{X}_t^N(\omega)$ on every compact A in $\mathbb{R}^+$ and almost surely on Ω. Let us chose M so that $A \subset S_{-M,0} = [0, 2^M]$. For $n > 0$, we already know the following inequalities

$$\sup_{0\le k<2^{N+n-1}} \sup_{0\le t\le 2^M} |\Phi_{n,k}| \le \sqrt{\Gamma}\cdot 2^{-\frac{n+1}{2}},$$

$$\sup_{0\le k<2^{N+n-1}} \sup_{0\le t\le 2^M} |\Psi_{n,k}| \le \sqrt{\Gamma}\cdot e^{|\alpha|}2^{-\frac{n+1}{2}}$$

and we can show for $n < 0$ the new inequalities

$$\sup_{0\le k<2^{N+n-1}} \sup_{0\le t\le 2^M} |\Phi_{n,k}| \le \sqrt{\Gamma}2^M\cdot 2^{-\frac{|n|+1}{2}},$$

$$\sup_{0\le k<2^{N+n-1}} \sup_{0\le t\le 2^M} |\Psi_{n,k}| \le \sqrt{\frac{\Gamma 2^{M-1}}{|\alpha|}}\cdot 2^{-\frac{|n|+1}{2}}.$$

The speed of convergence of the upper bounds when n goes to infinity by negative or positive values is in $2^{-\frac{|n|+1}{2}}$, which justifies by the usual argument that $t \mapsto \widehat{X}^N(\omega)$ converges normally and uniformly in t and almost surely on Ω. We have the normal convergence of almost every sample path on any compact A and it causes the limit process $\lim_{N\to\infty}\widehat{X}^N$ to have continuous paths on $\mathbb{R}^+$. Moreover, as a sum of Gaussian variables, it is also a Gaussian process. Therefore, proving that $\lim_{N\to\infty}\widehat{X}^N$ is a discrete representation of the process X just requires us to show that for every t, s in $\mathbb{R}^+$ its covariance satisfies

$$\lim_{N\to\infty}\langle\widehat{X}_t^N\cdot\widehat{X}_s^N\rangle = \langle\widehat{X}_t\cdot\widehat{X}_s\rangle.$$

6.3 Covariance Calculus

To verify the cogency of the bi-infinite decomposition, it is enough to demonstrate that the covariance

$$\begin{aligned}\langle \widehat{X}_t^N \cdot \widehat{X}_s^N \rangle &= \sum_{n=-N+1}^{N} \sum_{0 \le k < 2^{N+n-1}} f_{n,k}(t) f_{n,l}(s) \\ &= \langle {}_N X_t^N \cdot {}_N X_s^N \rangle - \bar{f}_{-N,0}(t) \bar{f}_{-N,0}(s)\end{aligned}$$

converges toward the expected covariance of the process X for every t, s in $\mathbb{R}^+$. We chose M satisfying $2^M \ge \max(t, s)$ and we have t, s elements of $S_{-M,0}$. We introduce for $n \ge -M$ the indices k_n and l_n to designate the sequence of supports S_{n,k_n} and S_{n,l_n}, which contain t and s respectively. We can write the reduced expression of the partial sums ${}_M X_t^N$ defined in (31)

$${}_M X_t^N = \sum_{n=-M}^{N} f_{n,k_n}(t) \cdot \xi_{n,k_n} + \bar{f}_{-M,0}(t) \cdot \xi_{-M,0}.$$

We remark that, for $P > M$, the covariance of the partial sums ${}_P X^N$ reads on $S_{-M,0}$

$$\begin{aligned}\langle {}_P X_t^N \cdot {}_P X_s^N \rangle &= \langle {}_M X_t^N \cdot {}_M X_s^N \rangle + \bar{f}_{-P,0}(t) \bar{f}_{-P,0}(s) \\ &\quad - \bar{f}_{-M,0}(t) \bar{f}_{-M,0}(s) + \sum_{n=-P+1}^{-M} f_{n,k_n}(t) f_{n,l_n}(s). \end{aligned} \tag{34}$$

In the previous relation, we have extended the definition of the indices k_n and l_n for $n \ge -P$ on $S_{-P,0}$. If t and s are elements of $S_{-M,0}$, it is easy to see that we necessarily have $k_n = 0$ and $l_n = 0$ for $n < -M$. We then realize that the bottom line of (34) cancels out since we show in supplementary materials that

$$\begin{aligned}\sum_{n=-P+1}^{-M} f_{n,k_n}(t) f_{n,l_n}(s) &= \sum_{n=-P+1}^{-M} f_{n,0}(t) f_{n,0}(s) \\ &= \bar{f}_{-M,0}(t) \bar{f}_{-M,0}(s) - \bar{f}_{-P,0}(t) \bar{f}_{P,0}(s).\end{aligned}$$

We are in a position to identify the covariance of ${}_N X^N$ and ${}_M X^N$ for t, s in $S_{-M,0}$ and $N \ge M$

$$\langle {}_N X_t^N \cdot {}_N X_s^N \rangle = \langle {}_M X_t^N \cdot {}_M X_s^N \rangle.$$

We now need to study the convergence of the covariance of $\widehat{X}^N$ expressed on $S_{-M,0}$ under the new form

$$\langle \widehat{X}_t^N \cdot \widehat{X}_s^N \rangle = \langle {}_M X_t^N \cdot {}_M X_s^N \rangle - \bar{f}_{-N,0}(t) \bar{f}_{-N,0}(s).$$

The exact same calculation as in Sects. 3.2 and 5.2 proves that we have

$$\lim_{N \to \infty} \langle {}_M X_t^N \cdot {}_M X_s^N \rangle = \langle X_t \cdot X_s \rangle,$$

whether X designates a Wiener process or an Ornstein-Uhlenbeck process. Therefore, the covariance of $\widehat{X}^N$ converges toward the expected expression for the covariance of the process

X as soon as we have

$$\lim_{N\to\infty} \bar{f}_{-N,0}(t)\, \bar{f}_{-N,0}(s) = 0.$$

For the Wiener process the examination of expression (32) reveals that $\bar{\Psi}_{-N,0}(t)$ tends to zero for every $t \geq 0$ when N goes to infinity. For the Ornstein-Uhlenbeck process, the examination of expression (33) shows that $\bar{\Phi}_{-N,0}(t)$ tends to zero for every $t \geq 0$ when N goes to infinity if and only if $\alpha \geq 0$. With the restriction that α should be positive, this validates our claim to represent the Wiener process and the Ornstein-Uhlenbeck process as a bi-infinite sum of random functions $f_{n,k} \cdot \xi_{n,k}$ on $\mathbb{R}^+$.

Acknowledgement We thank Mariela Sued and Daniel Andor. This work has been supported in part by NIH under grant R01-DC07294.

References

1. Risken, H.: The Fokker-Planck Equation. Springer, Berlin (1989)
2. Van Kampen: Stochastic Processes in Physics and Chemistry. Elsevier, Amsterdam (2007)
3. Gerstner, W., Kistler, W.: Spiking Neuron Models. Cambridge University Press, Cambridge (2002)
4. Koch, C.: Biophysics of Computation: Information Processing in Single Neurons. Oxford University Press, London (1998)
5. Berg, H.: Random Walks in Biology. Princeton University Press, Princeton (1993)
6. Rolski, T., Schmidli, H., Schmidt, V., Teugels, J.: Stochastic Processes for Insurance and Finance. Wiley Series in Probability and Statistics. Wiley, New York (2001)
7. Siegert, A.J.F.: On the first passage time probability problem. Phys. Rev. **81**, 617–623 (1951)
8. Ricciardi, L.M., Sato, S.: First-passage time density and moments of the Ornstein-Uhlenbeck process. J. Appl. Probab. **25**, 43 (1988)
9. Leblanc, B., Renault, O., Scaillet, O.: A correction note on the first passage time of an Ornstein-Uhlenbeck process to a boundary. Finance Stoch. **4**(1), 109–111 (2000)
10. Alili, L., Patie, P., Pedersen, J.L.: Representations of the first hitting time density of an Ornstein-Uhlenbeck process. Stoch. Models **21**, 967–980 (2005)
11. Giraudo, M.T., Sacerdote, L., Zucca, C.: A Monte Carlo method for the simulation of first passage times for diffusion processes. Methodol. Comput. Appl. Probab. **3**, 215–231 (2001)
12. Bouleau, N., Lepingle, D.: Numerical Methods for Stochastic Processes. Wiley Series in Probability and Statistics. Wiley, New York (1994)
13. Gardiner, C.W.: Handbook of Stochastic Methods. Springer, Berlin (2004)
14. Klauder, J.L., Petersen, W.P.: Numerical integration of multiplicative-noise stochastic differential equations. SIAM J. Numer. Anal. **22**, 1153–1166 (1985)
15. Fox, R.F.: Second-order algorithm for the numerical integration of colored-noise problems. Phys. Rev. A **46**, 2649–2654 (1991)
16. Honeycutt, R.L.: Stochastic Runge-Kutta algorithms. I. White noise. Phys. Rev. A **45**, 600–603 (1992)
17. Cecchi, G.A., Magnasco, M.O.: Negative resistance and rectification in Brownian transport. Phys. Rev. Lett. **76**, 1968–1971 (1996)
18. Bao, J., Abe, Y., Zhuo, Y.: An integral algorithm for numerical integration of one-dimensional additive colored noise problems. J. Stat. Phys. **90**, 1037–1045 (1997)
19. Hida, T.: Brownian Motion. Springer, Berlin (1980)
20. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, Berlin (2000)
21. Pyke, R.: The Haar Function Construction of Brownian Motion Indexed by Sets. Springer, Berlin (1983)