HAL open science

Functional quantization based stratified sampling methods

Sylvain Corlay, Gilles Pagès

To cite this version:

Sylvain Corlay, Gilles Pagès. Functional quantization based stratified sampling methods. 2010. hal-00464088v1

HAL Id: hal-00464088

https://hal.science/hal-00464088v1

Preprint submitted on 15 Mar 2010 (v1), last revised 4 Oct 2014 (v3)

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.

Functional quantization based stratified sampling methods

Sylvain Corlay* † Gilles Pagès †

March 15, 2010

Abstract

In this article, we propose several quantization based stratified sampling methods to reduce the variance of a Monte-Carlo simulation.

Theoretical aspects of stratification lead to a strong link between the problem of optimal L^2-quantization of a random variable and the variance reduction that can be achieved. We first emphasize on the consistency of quantization for designing strata in stratified sampling methods in both finite dimensional and infinite dimensional frameworks. We show that this strata design has a uniform efficiency among the class of Lipschitz continuous functionals.

Then a stratified sampling algorithm based on product functional quantization is proposed for path-dependent functionals of multi-factor diffusions. The method is also available for processes made of other Gaussian processes as the Brownian bridge or an Ornstein-Uhlenbeck process.

The balance between the algorithmic complexity of the simulation and the variance reduction factor has also been studied.

Keywords: functional quantization, vector quantization, stratification, variance reduction, Monte-Carlo simulation, Karhunen-Loève basis, Gaussian process, Brownian motion, Brownian bridge, fractional Brownian motion, principal component analysis, numerical integration, option pricing, Voronoi diagram, product quantizer.

*Natixis, Equity Derivatives and Arbitrage. E-mail: sylvain.corlay@natixis.com. The authors would like to thank the members of Natixis' equity derivatives quantitative R&D team for fruitful discussions.

†Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: gilles.pages@upmc.fr

Introduction

The quantization of a random variable X consists of its approximation by a random variable Y taking at most N values. This problem has been initially investigated for its applications to signal transmission and for compression issues [8]. In this context, quantization was a process of signal discretization. The point of interest was to choose the random variable Y in order to minimize the resulting error. This led to the concept of optimal quantization.

More recently, quantization has been introduced in numerical probability to device numerical integration methods [17] and for solving multi-dimensional stochastic control problems such as American options pricing [1] and swing options pricing [2, 3]. Optimal quantization has many other applications and extensions in various fields like automatic classification (quantization of empirical measures) and pattern recognition.

In the early 2000's, the infinite dimensional setting has been extensively investigated from a theoretical and numerical viewpoints with a special attention paid to functional quantization [13, 14]. Stochastic processes are viewed as random vectors taking values in their path spaces such as $L^2_T := L^2([0,T], dt)$.

Still the Monte-Carlo simulation remains the most common numerical methods in the field of numerical probability. One reason is that it is easy to implement in an industrial configuration. In the industry of derivative pricing, banks implement generic Monte-Carlo frameworks for pricing numerous payoffs with a wide variety of models. Another advantage is that the Monte-Carlo method can be parallelized.

Variance reduction methods can be used to reduce dramatically the computation time of a Monte-Carlo simulation. Main variance reduction methods are (adaptive) control variate, pre-conditioning, importance sampling and stratification [9]. The problem is that these methods may strongly depend on the payoff or the model and imply wide changes in the practical implementation of the Monte-Carlo method. Thus, most institutions do not implement the most advanced methods in practice except for marginal cases.

In this paper, we point out theoretical aspects of quantization that lead to a strong link between the problem of optimal L^2-quantization of a random variable and the variance reduction that can be achieved by stratification. We emphasize on the consistency of quantization for designing strata in stratified sampling methods in both finite dimensional and infinite dimensional frameworks. Then a stratified sampling algorithm based on product functional quantization is proposed for path-dependent functionals of multi-factor Brownian diffusions. The simulation cost of the conditional path is $O(n)$ where n is the number of discretization dates, as in naive Monte-Carlo simulations. In this context, this stratification based variance reduction method can be considered as a guided Monte-Carlo simulation. (See figure 6.) The method is in fact available for any Gaussian processe as soon as its Karhunen-Loève decomposition is explicitly known. It is the case for the Brownian bridge or an Ornstein-Uhlenbeck process. The special case of the Ornstein-Uhlenbeck process is derived in annex 7.

One very common situation is the case of Monte-Carlo implementations that are based on multi-dimensional and multi-factor Brownian diffusions approximated by their Euler scheme. The presented method is particulary adapted to this situation. Even in the multi-dimensonal case, no matter how the inde-

pendent Brownian motions are correlated or used afterwards ; no matter if it is used for diffusing the underlying stock, a stochastic volatility process or an actualization factor. Functional stratification can be used as a generic variance reduction method. The point is that it is used upstream in the Monte-Carlo framework. One does not need to re-implement the whole framework but only the way it is alimented by Brownian motions. Thus quantization-based functional stratifications can come along on the top of a production context. In the last section, numerical tests are provided with a benchmark with a Up-In-Call pricing in the Black and Scholes model.

1 Optimal quantization, the abstract framework

1.1 Introduction to quantization of random variables

In the following, $(\Omega, \mathcal{A}, \mathbb{P})$ is a probability space, and E is a separable reflexive Banach space. The norm on E is denoted $|\cdot|$.

The principle of the quantization of a random variable X taking its values in E is to estimate X by a random variable Y taking a finite number of values in E. The discrete random variable Y is a quantizer of X.

The resulting error of this discretization is the L^p-norm of $X-Y$. One wants to minimize this induced error. This gives the following minimization problem:

$$\min\Big\{\|X-Y\|_p, Y:\Omega\to E \text{ measurable, } \operatorname{card}(Y(\Omega))\leq N\Big\}. \tag{1}$$

Definition (Nearest neighbour projection)**.** *Let us consider the settled point set $\Gamma=\{y_1,\cdots,y_N\}\subset E$. A nearest neighbour projection on Γ is a Borel application $\mathrm{Proj}_\Gamma : E\to\Gamma$ such that $\forall x\in E\ |x-\mathrm{Proj}_\Gamma(x)| = \min\limits_{1\leq i\leq N}|x-y_i|$.*

Definition (Voronoi partition)**.** *$N\in\mathbb{N}^*$, $\Gamma=\{y_1,\cdots,y_N\}\subset E$ and $C=\{C_1,\cdots,C_N\}$ a partition of E. C is a Voronoi partition associated with Γ if $\forall i\in[\![1,N]\!],\ C_i\subset\{\xi\in E, |\xi-y_i| = \min\limits_{j\in[\![1,N]\!]}|\xi-y_j|\}$.*

Proposition 1.1. *Let X be an E-valued L^p random variable, and Y taking its values in the settled point set $\Gamma=\{y_1,\cdots,y_N\}\subset E$ where $N\in\mathbb{N}$. Let $\widehat{X}^\Gamma$ the random variable defined by $\widehat{X}^\Gamma := \mathrm{Proj}_\Gamma(X)$ where Proj_Γ is a nearest neighbour projection on Γ, called a Voronoi Γ-quantizer of X.*

Then we clearly have $|X-\widehat{X}^\Gamma|\leq|X-Y|$ a.s.. Hence $\|X-\widehat{X}^\Gamma\|_p\leq\|X-Y\|_p$.

If $C=\{C_1,\cdots,C_N\}$ is a Voronoi partition associated with $\Gamma=\{y_1,\cdots,y_N\}$, it is clear that $\forall i\in[\![1,N]\!], y_i\in C_i$. C_i is called Voronoi slab associated with y_i in C and y_i is the centre of the slab C_i.
One denotes $C_i=\mathrm{slab}_C(y_i)$, and for every $a\in\Gamma$, $W(a|\Gamma)$ is the closed subset of E defined by

$$W(a|\Gamma)=\Big\{y\in E\Big||y-a| = \min_{b\in\Gamma}|y-b|\Big\}.$$

It is clear that $\sum\limits_{1=1}^{N} y_i\mathbf{1}_{C_i}$ is a nearest neighbour projection on Γ.

If we want to approximate a random variable X by Y valued in a finite subset Γ, the best option is to choose for Y a measurable nearest neighbour projection of X on Γ.

As a consequence of the previous remark, solving the minimization problem (1) comes to solving the simpler minimization problem

$$\min\Big\{\|X-\mathrm{Proj}_\Gamma(X)\|_p,\ \Gamma\subset E, \mathrm{card}(\Gamma)\leq N\Big\}. \tag{2}$$

The quantity $\|X-\mathrm{Proj}_\Gamma(X)\|_p$ is called the L^p-quantization error. When this minimum is reached, one refers to optimal quantization.

The problem of the existence of a minimum have been investigated for decades on its numerical and theoretical aspects in the finite dimensional case [10].

- N being settled, the L^p-distortion is Lipschitz continuous and reaches a minimum. A Ntuple that achieves the minimum has pairwise distinct components, as soon as $\mathrm{card}(\mathrm{supp}(\mathbb{P}_X))\geq N$. This result stands in the general abstract case of a random value in a reflexive Banach space.

- If $\mathrm{card}(X(\Omega))$ is infinite, this minimum strictly decreases to 0 as N goes to infinity. The rate of convergence is ruled by Zador theorem in the finite dimensional case.

Theorem 1.2 (Zador)**.** *Assume that for some $\varepsilon>0$, $\mathbb{E}[|X|^{p+\varepsilon}]<+\infty$. If $\Gamma\subset\mathbb{R}^d$ is finite, let us denote $\widehat{X}^\Gamma=\mathrm{Proj}_\Gamma(X)$, a measurable nearest neighbour projection. Then*

$$\lim_{n\to+\infty}\Big(N^{\frac{p}{d}}\min_{\mathrm{card}(\Gamma)\leq N}\|X-\widehat{X}^\Gamma\|_p^p\Big)=J_{p,d}\Big(\int_{\mathbb{R}^d}g^{\frac{d}{d+p}}(\xi)\lambda_d(d\xi)\Big)^{1+\frac{p}{d}}$$

where $\mathbb{P}_X(d\xi)=g(\xi)\lambda_d(d\xi)+\mu(d\xi)$, $\mu\perp\lambda_d$ (λ_d is the Lebesgue measure on $\mathbb{R}^d$.) The constant $J_{p,d}$ corresponds to the case of the uniform distribution on $[0,1]^d$.

This mainly says us that $\min\Big\{\|X-\widehat{X}\|_p,\mathrm{card}(\Gamma)\leq N\Big\}\sim C_{\mathbb{P}_X,p,d}N^{-\frac{1}{d}}$.

In figure 1, the Voronoi partition of a random N-quantizer and an L^2-optimized N quantizer of the $\mathcal{N}(0,I_2)$ distribution are given.

1.2 Stationarity and centroidal Voronoi tessellations

We now stand in the case where E is a separable Hilbert space $(H,\langle .,.\rangle_H)$, and in the quadratic case ($p=2$ and $X\in L^2(H)$).

Notation Consider X a $L^2(H)$ random variable. We introduce the following notations:

- $\mathcal{C}_n(X)$ is the set of L^2-optimal quantizers of X of level n.

- $e_n(X)$ is the minimal quadratic distortion that can be achieved when approximating X by a quantizer of level n.

Definition (Stationarity)**.** *A quantizer Y of X is stationary (or self-consistent) if*

$$Y=\mathbb{E}[X|Y]. \tag{3}$$

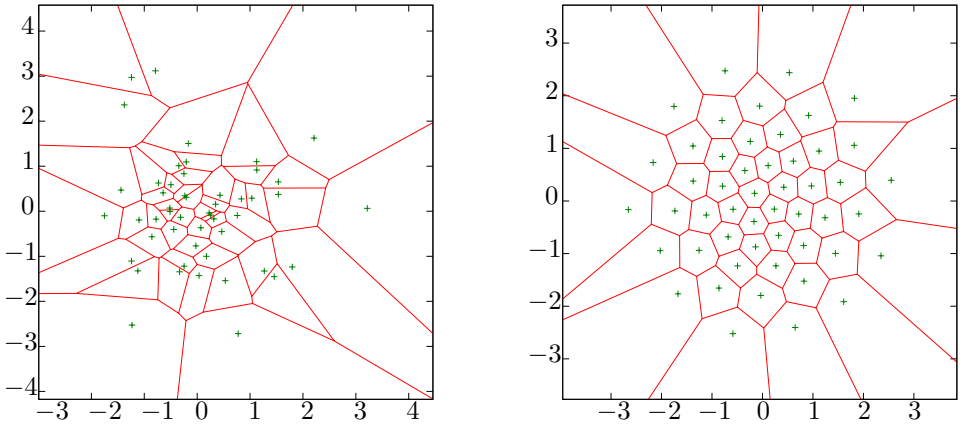

Figure 1: Voronoi partition of a random quantizer and a L^2-optimized N-quantizer of the $\mathcal{N}(0, I_2)$ distribution in $\mathbb{R}^2$. ($N = 20$).

Proposition 1.3 (Stationarity of L^2-optimal quantizer)**.** *A (quadratic) optimal quantizer is stationary.*

The stationarity is a particularity of the quadratic case. In other L^p cases, a similar property involving the notion of p-centre occurs. A proof of is available in [11].

A consequence is, if $Y = \mathrm{Proj}_\Gamma(X)$ is an L^2-optimal quantizer, and $C = \{C_1, \cdots, C_n\}$ is the associated Voronoi partition, one has $\forall y \in \Gamma, y = \mathbb{E}[X|X \in \mathrm{slab}_C(y)]$.

Proposition 1.4. *Let X be an H-valued L^2 random variable. Let us denote D_N^X the squared quadratic quantization error associated with a codebook of size N with respect to X.*

$$\begin{array}{rlcl} D_N^X: & H^N & \to & \mathbb{R}_+ \\ & (x_1, \cdots, x_N) & \to & \mathbb{E}\Big[\min\limits_{1\le i\le N} |X - x_i|_H^2\Big], \end{array}$$

The distortion function D_N^X is $|.|_H$-differentiable at N-quantizers $x \in H^N$ with pairwise distinct components and

$$\nabla D_N^X(x) = 2\Big(\int_{C_i(x)} (x_i - \xi)\mathbb{P}_X(d\xi)\Big)_{1\le i\le N} = 2\Big(\mathbb{E}(\hat{X}^{\Gamma(x)} - X)\mathbf{1}_{\{\hat{X}^{\Gamma(x)} = x_i\}}\Big)_{1\le i\le N}. \tag{4}$$

Hence any Voronoi quantizer associated with a critical point of D_N^X is a stationary quantizer.

Definition (Centroidal projection)**.** *Let $C = \{C_1, \cdots, C_N\}$ be a Borel partition of H. Let us define for $1 \le i \le N$, $G_i = \begin{cases} \mathbb{E}[X|X \in C_i] \text{ if } \mathbb{P}[X \in C_i] \neq 0, \\ 0 \text{ in the other case.} \end{cases}$ the centroids associated with X and the partition C.*

We call centroidal projection associated with the partition C and the random variable X and denote $\mathrm{Proj}_{C,X}$ the application

$$\begin{array}{rlcl} \mathrm{Proj}_{C,X}: & H & \to & \{x_1, \cdots, x_N\} \subset H \\ & x & \to & \sum\limits_{i=1}^N G_i \mathbf{1}_{C_i}(x). \end{array}$$

Lemma 1.5 (Huyghens, Variance Decomposition)**.** *Let $X \in L^2(\mathbb{P})$ be a H-valued L^2 random vector, $N \in \mathbb{N}^*$ and $C = (C_i)_{1 \le i \le N}$ a Borel partition of H. Consider* $\mathrm{Proj}_{C,X} = \sum_{i=1}^{N} G_i \mathbf{1}_{C_i}$ *the associated centroidal projection.*

Then one has, $\mathrm{Var}(X) = \underbrace{\mathbb{E}[|X - \mathrm{Proj}_{C,X}(X)|^2]}_{:=(1)} + \underbrace{\mathbb{E}[|\mathrm{Proj}_{C,X}(X) - \mathbb{E}[X]|^2}_{:=(2)}$.

The variance of the probability distribution X decomposes itself into the ***intraclass inertia*** (1) *plus the* ***interclass inertia*** (2).

Proof of lemma:

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[|X - \mathrm{Proj}_{C,X}(X) + \mathrm{Proj}_{C,X}(X) - \mathbb{E}[X]|^2] \\ &= \underbrace{\mathbb{E}[|X - \mathrm{Proj}_{C,X}(X)|^2]}_{=(1)} + \underbrace{\mathbb{E}[|\mathrm{Proj}_{C,X}(X) - \mathbb{E}[X]|^2]}_{=(2)} \\ &\quad + \underbrace{2\mathbb{E}[\langle X - \mathrm{Proj}_{C,X}(X), \mathrm{Proj}_{C,X}(X) - \mathbb{E}[X]\rangle]}_{:=(3)}.\end{aligned}$$

Now $(3) = 0$ since $\mathrm{Proj}_{C,X}(X) = \mathbb{E}[X|\mathrm{Proj}_{C,X}(X)]$. □

1.3 Optimal quantization and principal component analysis

1.3.1 Covariance operator of a random variable

A point of interest of the stationarity of L^2 optimal quantizers is that stationary quantizers necessarily lies in the reproducing kernel Hilbert space (or Cameron-Martin space) of the covariance operator of X (if $\mathbb{E}[X] = 0$).

Definition. *Let X be a centred H valued L^2 random variable.*

The covariance operator $C_X : H \to H$ of X is defined by $C_X y = \mathbb{E}[\langle y, X\rangle X]$.

1. C_X is a symmetric positive trace class operator.
2. In the finite dimensional case, the matrix of C_X in the canonical basis is the covariance matrix of X.

Definition. *The reproducing kernel Hilbert space K_X is a subspace of H defined by*

$$\begin{aligned}K_X &:= \left\{\mathbb{E}[ZX] | Z \in \overline{\{\langle y, X\rangle | y \in H\}}^{L^2(\mathbb{P})}\right\} \\ &= \left\{\mathbb{E}[g(X)X] | g \in \overline{\{\langle y, \cdot\rangle | y \in H\}}^{L^2(\mathbb{P}_X)}\right\}.\end{aligned}$$

The set K_X is equipped with the inner product

$$\langle k_1, k_2\rangle_X := \mathbb{E}[Z_1 Z_2] \text{ if } k_1 = \mathbb{E}[Z_1 X] \text{ and } k_2 = \mathbb{E}[Z_2 X].$$

so that $(K_X, \langle\cdot\rangle)$ is a Hilbert space, isometric with $\overline{\{\langle y, X\rangle | y \in H\}}^{L^2(\mathbb{P})}$.

Hence, K_X is spanned as a Hilbert space by $\{C_X(y) | y \in H\}$. Taking any $g \in L^2(P)$ does not enlarge K_X, so that

$$K_X = \{\mathbb{E}[g(X)X] | g \in L^2(\mathbb{P})_X\}. \tag{5}$$

Furthermore, for every $y, z \in H$, thanks to the Fubini theorem, one has the so called reproducing property: $\langle k, C_X y\rangle_X = \langle k, y\rangle,\ k \in K_X, y \in H$.

Proposition 1.6. *If $\mathbb{E}[X] = 0$ and $\Gamma \subset H$ is a N-stationary codebook for X, then $\Gamma \subset K_X$.*

Proof: Consider $a \in \Gamma$. $a = \mathbb{E}[X|X \in W(a|\Gamma)] = \mathbb{E}[g(X)X]$, where $g = \mathbf{1}_{W(a|\Gamma)}/\mathbb{P}_X[W(a|\Gamma)] \in L^2(\mathbb{P})$. The assertion follows from equation (5).

1.3.2 The case of a bi-measurable stochastic processes

Consider the Hilbert space $H = L^2([0,T], dt)$ and a bi-measurable centred $L^2(\mathbb{P})$-process $X = (X_t)_{t\in[0,T]}$ with paths in $L^2([0,T], dt)$ *a.s.* and covariance function $\Gamma_X(s,t) := \mathbb{E}[X_s X_t]$ satisfying $\int_{[0,T]} \Gamma_X(s,t)ds < +\infty$. Then X can be seen as a H-valued random vector with $\mathbb{E}[\|X\|^2] < \infty$.

$$C_X y = \int_{[0,T]} y(s)\Gamma_X(s,\cdot)ds, \;\; y \in L^2([0,T], dt). \tag{6}$$

In this setting, proposition 1.6 indicates that the components of a stationary quantizer have certain smoothness properties. Proofs are available in [13, 11].

1.3.3 Reduction of dimension

The aim is now the reduction of the quantization problem to finite dimensional subspaces of H. For any finite dimensional subspace U of H, let Π_U denote the orthogonal projection from H onto U.

Proposition 1.7. *Let U be a finite dimensional linear subspace of H. Then*

$$\begin{aligned} e_n(\Pi_U(X))^2 \le e_n(X)^2 \quad &\le \inf\left\{\mathbb{E}[\min_{a\in\Gamma} \|X-a\|^2] \middle| \Gamma \subset U, 1 \le \operatorname{card}\Gamma \le n\right\} \\ &= \mathbb{E}\|X - \Pi_U(X)\|^2 + e_n(\Pi_U(X))^2. \end{aligned}$$

In other words, the quadratic quantization error with respect to $\Gamma \subset U$ consists of the projection error and the quantization error of the projected random vector.

Notation: Let us denote $d_n(X) = \min\{\dim \operatorname{span}(\Gamma), \Gamma \in \mathcal{C}_n(X)\}$ the quantization dimensiion of the level n of the quantization problem for X.

It follows from proposition 1.7 that $e_n^2(X) = \min\{\mathbb{E}[\|X - \Pi_V(X)\|^2] + e_n^2(\Pi_V(X)) | V \subset H \text{ linear subspace } \dim V \ge d_n(X)\}$.

1.3.4 The case of Gaussian measures

In this section, let X be a centred H-valued random vector with Gaussian distribution $\mathbb{P}_X$. Since we wish to investigate the infinite dimensional situation, we assume throughout that $\dim K_X = \infty$. ($\operatorname{supp}(\mathbb{P}_X) = \overline{K_X}$.)

In this case, the proposition 1.6 can be considerably improved.

Theorem 1.8. *Le $\Gamma \subset H$ be a n-stationary codebook for X and let $U = \operatorname{span}(\Gamma)$. Then $\Pi_U(X)$ and $X - \Pi_U(X)$ are independent so that $C_X(U) = U$. In particular, $\Gamma \subset C_X(H) \subset K_X$.*

This theorem shows that linear subspaces U of H spanned by n-stationary codebooks correspond to principal components of X, in other words, are spanned by eigenvectors of C_X.

In order to deal with n-optimal codebooks, let $\lambda_1^X \geq \lambda_2^X \geq \cdots > 0$ be the ordered non-zero eigenvalues of C_X (written as many times as its multiplicity) and note that $\mathbb{E}[\|X\|^2] = \sum_{j=1}^{\infty} \lambda_j^X$.

Theorem 1.9. *Let Γ be an optimal codebook for X, $U = \mathrm{span}(\Gamma)$ and $m = \dim U$. Then $C_X(U) = U$ and*

$$\mathbb{E}\|X - \Pi_U(X)\|^2 = \sum_{j \geq m+1} \lambda_j^X.$$

Observe that

$$\sum_{j \geq m+1} \lambda_j^X = \inf\{\mathbb{E}\|X - \Pi_V(X)\|^2 | V \subset H \text{ linear subspace}, \dim V = m\}.$$

This theorem shows that m-dimensional subspaces of H spanned by n-optimal codebooks are spanned by eigenvectors of C_X that belong to the m largest eigenvalues. Thus these subspaces correspond to the first m principal components of X.

We now deduce the final representation of $e_n(X)$.

$$e_n(X)^2 = \sum_{j \geq m+1} \lambda_j^X + e_n\Big(\bigotimes_{j=1}^{m} \mathcal{N}(0, \lambda_j^X)\Big)^2 \quad \text{for } m \geq d_n(X), \tag{7}$$

$$e_n(X)^2 < \sum_{j \geq m+1} \lambda_j^X + e_n\Big(\bigotimes_{j=1}^{m} \mathcal{N}(0, \lambda_j^X)\Big)^2 \quad \text{for } 1 \leq m < d_n(X). \tag{8}$$

This two equations (7) and (8) show that for the quantization of a Gaussian process, as soon as we know its Karhunen-Loève basis $(e_n)_{n \in \mathbb{N}^*}$ and its eigenvalues $(\lambda_n^X)_{n \in \mathbb{N}^*}$, the problem of optimal L^2-quantization comes to the problem of the quantization of a Gaussian vector of dimension d_n.

In the following, one will see that the Karhunen-Loève basis of many common Gaussian processes are known (Brownian motion, Brownian bridge and the Ornstein-Uhlenbeck process).

1.3.5 Product quantization

Let $(e_n)_{n \in \mathbb{N}^*}$ be a Hilbertian basis of H and $I \subset \mathbb{N}^*$ is a finite subset of $\mathbb{N}^*$. For every $k \in I$, consider a N_k-tuple $\Gamma^{N_k} = \{x_1^{N_k}, \cdots, x_{N_k}^{N_k}\} \subset \mathbb{R}$.

An easy way to construct a quantizer is to define the codebook Γ by the set of the points x such that for every $k \in I, \langle x, e_k \rangle \in \Gamma_k$ and for every $k \in \mathbb{N}^* - I, \langle x, e_k \rangle = \mathbb{E}[\langle X, e_k \rangle]$.

To quantize X, one will consider $Y = \mathrm{Proj}_\Gamma(X)$, where Proj_Γ is a nearest neighbour projection on Γ. The Voronoi cells associated with such a codebook are hyper-parallelepipeds.

Proposition 1.10 (Case of independent marginals)**.** *With the same notations, if one assumes that the marginals of X, $(\langle X, e_1 \rangle, \langle X, e_2 \rangle, \cdots)$ are independent, then one can choose for each $k \in I$ the values $\Gamma^k = \{x_1^{N_k}, \cdots, x_{N_k}^{N_k}\}$ such that $Y^k = \mathrm{Proj}_{\Gamma^k}(\langle X, e_k \rangle)$ is a stationary quantizer of $\langle X, e_k \rangle$. Then $Y = \mathrm{Proj}_\Gamma(X)$ is a stationary quantizer of X.*

This method gives us a way to get a stationary quantizer with an easy projection rule.

A drawback of product quantization is that one needs to restrict themselves to the case of independent marginals in order to have stationarity.

1.4 Numerical optimal quantization

Various numerical algorithm have been developped to get numerically an optimal N-grid with a minimal qudratic quantization error in the finite-dimensional setting. A review of these methods is available in [16]. Let us mention the Lloyd's algorithm for the quadratic case, which is the natural probabilistic counterpart of a classification algorithm due to Forgy [7].

Another algorithm is a stochastic gradient method which is suggested by the fact that the L^2-quantization distortion function is differentiable at any N-tuple having pairwise distinct components and a $\mathbb{P}_X$-negligible Voronoi tessellation boundary and has an integral representation. The algorithm is deeply investigated in [17].

Equation (4) shows that any Voronoi quantizer associated with a critical point of D_N^X is a stationary quantizer. In the case of one dimensional distributions, as the Gaussian distribution, the Hessian of the distortion is known and can be represented by a tridiagonal matrix. Hence, it is easy to invert and a Newton method can be implemented. It is completely detailed in [17] in the Gaussian case. It is still the fastest way to compute L^2-optimal quantizers of one-dimensional Gaussian variables.

1.5 Application to numerical integration

Let X be an integrable E-valued random vector. The quantization method for numerical integration consists in approximating the distribution $\mathbb{P}_X$ by the distribution of a quantizer Y, valued on $\Gamma = \{y_1, \cdots, y_N\} \subset E$. $\mathbb{P}_Y = \sum_{i=1}^{N} p_i \delta_{y_i}$, where $p_i = \mathbb{P}[Y = y_i]$ and δ_{y_i} is the Dirac mass at y_i.

Then, one approximate the quantity $\mathbb{E}[\phi(X)] = \int_E \phi(\xi)\mathbb{P}_X(d\xi)$ by $\mathbb{E}[\phi(Y)] = \int_E \phi(\xi)\mathbb{P}_Y(d\xi) = \sum_{i=1}^{N} p_i \phi(x_i)$, where ϕ is a Lipschitz continuous function on E.

Remark. *If one supposes that Y is X-measurable, there is a measurable partition of E, $C = \{C_1, \cdots, C_N\}$ such that $Y = \sum_{i=1}^{N} y_i \mathbf{1}_{C_i}(X)$, hence $p_i = \mathbb{P}(Y = y_i) = \mathbb{P}_X(C_i)$.*

Error control for numerical integration:

In [17], Pagès and Printems have established error control inequalities when doing this approximation.

• **The Lipschitz case**

When ϕ is Lipschitz continuous, the resulting error is controlled by:

$$|\mathbb{E}[\phi(X)] - \mathbb{E}[\phi[Y]]| \quad \leq [\phi]_{Lip} \|X - Y\|_1 \quad \leq [\phi]_{Lip} \|X - Y\|_p \quad (p \geq 1).$$

• **The Lipschitz derivative case**

If ϕ is continuously differentiable with a Lipschitz continuous differential $D\phi$. Taylor's formula gives $|\phi(X) - (\phi(Y) + D\phi(Y).(X-Y))| \leq [D_\phi]_{Lip}|X-Y|^2$,

$$\text{so that } \Big|\mathbb{E}[\phi(X)] - \mathbb{E}[\phi(Y)] - \mathbb{E}[D\phi(Y).(X-Y)]\Big| \leq [D\phi]_{Lip}\|X-Y\|_2^2$$

Now, the stationnarity (3) of Y ensures

$$\mathbb{E}[D\phi(Y).(X-Y)] = \mathbb{E}[D\phi(Y).\mathbb{E}[(X-Y)|Y]] = 0,$$

hence $|\mathbb{E}[\phi(X)] - \mathbb{E}[\phi[Y]]| \leq [D\phi]_{Lip}\|X-Y\|_2^2 = O(N^{-2/d})$.

In other words, stationarity of a quantizer makes us reach an order in the convergence rate of approximation of the expectation.

• **The convex case**

When ϕ is a convex function and Y is a stationary quantizer, Jensen inequality yields $\mathbb{E}\big[\phi(Y)\big] = \mathbb{E}\big[\phi\big(\mathbb{E}[X|Y]\big)\big] \leq \mathbb{E}[\phi(X)]$. Hence $\mathbb{E}[\phi(Y)]$ is always a lower bound for $\mathbb{E}[\phi(X)]$.

1.6 Quantization of Gaussian processes

1.6.1 Quantization

From now on, as in section 1.3.2, we will assume that X is a bi-measurable Gaussian process defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying $\mathbb{E}[|X|^2_{L^2_T}] = \int_0^T \mathbb{E}[X_s^2]ds < \infty$, so that it can be viewed as an L^2_T-valued random vector.

We have seen in section 1.3.4 that in this context, as soon as one knows the Karhunen-Loève system $(e_n^X, \lambda_n^X)_{n\in\mathbb{N}^*}$ of the covariance operator of X, the problem of the L^2-optimal quantization of the process X comes to the quantization of the Gaussian vector $\bigotimes_{j=1}^m \mathcal{N}(0, \lambda_j^X)$. ($(\lambda_n^X)_{n\in\mathbb{N}^*}$ is sorted in the decreasing order.)

Once we have computed an optimal quantizer of the finite dimensional Gaussian vector $\bigotimes_{j=1}^m \mathcal{N}(0, \lambda_j^X)$ given in equation (7), we have an optimal quantizer of the underlying Gaussian process at our disposal. The companion parameters of the functional quantizer are easily deduced from the ones of the quantizer of $\bigotimes_{j=1}^m \mathcal{N}(0, \lambda_j^X)$ we use.

All this is true for any Gaussian process X, except that one needs to know explicitly its Karhunen-Loève Basis to compute the optimal quantizer. Common Gaussian processes have explicit Karhunen-Loève expansions, like the Brownian motion and the Brownian bridge. The Ornstein-Uhlenbeck process admits a semi-closed form for its Karhunen-Loève expansion. (The formula is derived for standard parameters in the stationary case in [12].) In section 7, (annex about the Ornstein-Uhlenbeck Karhunen-Loève basis), the computation of Karhunen-Loève decomposition of the Ornstein-Uhlenbeck process is detailed. The general case where $r_0 \sim \mathcal{N}(m_0, \sigma_0^2)$ is handled. As far as we know, the K-L expansion of the fractional Brownian motion is not known.

Later in the paper, numerical illustrations will be given for the following cases:

1. The Brownian motion $(W_t)_{t\in[0,T]}$,

$$e_n^W(t) := \sqrt{\frac{2}{T}}\sin\Big(\pi(n-1/2)\frac{t}{T}\Big), \qquad \lambda_n^W := \Big(\frac{T}{\pi(n-1/2)}\Big)^2, \quad n\geq 1. \tag{9}$$

2. The Brownian bridge on $[0,T]$,

$$e_n^B(t) := \sqrt{\frac{2}{T}}\sin\Big(\pi n\frac{t}{T}\Big), \qquad \lambda_n^B := \Big(\frac{T}{\pi n}\Big)^2, \quad n\geq 1. \tag{10}$$

3. The Ornstein-Uhlenbeck process on $[0,T]$, starting from 0, (See section 7.)

$$e_n^{OU}(t) := \Big(\frac{1}{\sqrt{\frac{T}{2}-\frac{\sin(2\omega_{\lambda_n}T)}{4\omega_{\lambda_n}}}}\Big)\sin(\omega_{\lambda_n}t), \qquad \lambda_n^{OU} := \frac{\sigma^2}{\omega_{\lambda_n}^2+\theta^2}, \quad n\geq 1, \tag{11}$$

where ω_{λ_n} are the (sorted) strictly positive solutions of the equation

$$\theta\sin(\omega_{\lambda_n}T)+\omega_{\lambda_n}\cos(\omega_{\lambda_n}T)=0.$$

4. The stationary Ornstein-Uhlenbeck process on $[0,T]$. (See section 7.)

On figure 2, one can see a N-optimal L^2-quantizer of the standard Brownian motion.

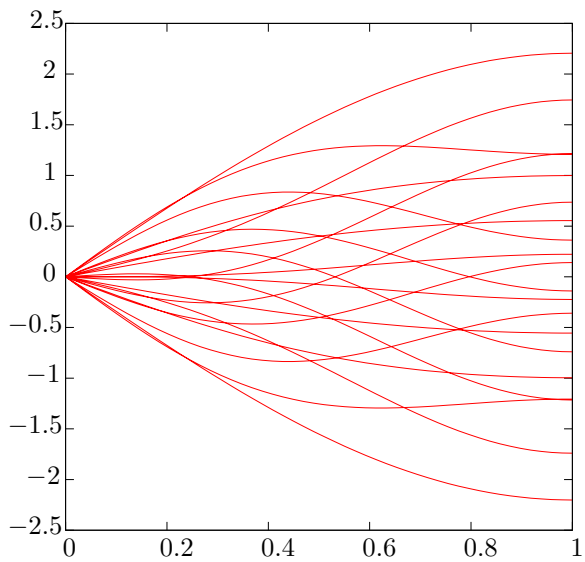

Figure 2: Optimal quantizer of a standard Brownian motion on $[0,1]$.

1.6.2 Product quantization

Thanks to equations (7) and (8), product quantization of the finite dimensional Gaussian vector $\xi\sim\bigotimes_{j=1}^{m}\mathcal{N}(0,\lambda_j^X)$ yields a stationary quantizer of the process X. In this context, let us give the following notations:

The quantizer of X is $\widehat{X}_t = \sum_{n\geq 1} \sqrt{\lambda_n^X}\widehat{\xi}_n e_n^X(t)$, where $\widehat{\xi}_n = \mathrm{Proj}_{\Gamma^{(N_n)}}(\xi_n)$ is an optimal N_n-quantization of ξ_n and $N_1 \times \cdots \times N_n \leq N,\ N_1, \cdots, N_n \geq 1$. (Hence for large enough n, $N_n = 1$ so that $\widehat{\xi}_n = 0$.)

The paths of an $N_1 \times \cdots \times N_n$-quantizer χ and a multi-index $\underline{i} = \{x_1, \cdots, x_n, \cdots\}$ that produces this quantization are of the form

$$\chi_{\underline{i}}(t) = \sum_{n\geq 1} \sqrt{\lambda_n^X} x_{i_n}^{N_n} e_n^X(t). \tag{12}$$

A quantizer χ defined by equation (12) is called a K-L product quantizer. Furthermore, one denote by $\mathcal{O}_{pq}(X, N)$ the set of the K-L product quantizers of size at most N of X.

In the case of a product quantization, the counterpart of equation (7) is

$$\begin{aligned}\mathbb{E}[\min_{\underline{i}} |X - \chi_{\underline{i}}|^2] &= \sum_{n=1}^{d_x} \lambda_n^X \mathbb{E}\Big[\min_{1\leq i_n\leq N_n} |\xi_n - x_{i_n}^{(N_n)}|\Big] + \sum_{n\geq d_x+1} \lambda_n^X \\ &= \sum_{n=1}^{d_x} \lambda_n^X \mathbb{E}\Big[\min_{1\leq i_n\leq N_n} |\xi_n - x_{i_n}^{(N_n)}|\Big] + \mathbb{E}[|X|^2_{L^2_T}] - \sum_{n=1}^{d_x} \lambda_n^X.\end{aligned} \tag{13}$$

1.6.3 Product decomposition blind optimization

As a consequence, the lowest quadratic quantization error induced by a K-L-product quantizer having at most N codebooks is obtained as a solution of the minimization problem

$$\min\Big\{e(\chi), \chi \in \mathcal{O}_{pq}(X, N)\Big\}, \tag{14}$$

that is, thanks to equation (13)

$$\min\Big\{\sum_{n=1}^{d} \lambda_n^X \min_{\mathbb{R}^{N_n}} \|\xi - \widehat{\xi}^{(N_n)}\|_2^2 + \sum_{n\geq d+1} \lambda_n^X, N_1 \times \cdots \times N_n \leq N, d \geq 1\Big\}. \tag{15}$$

A solution of (14) is called an optimal K-L product quantizer.

The blind optimization procedure consists of computing the criteria for every possible decomposition $N_1 \times \cdots \times N_n \leq N$. For a given Gaussian process X, results can be kept off-line for a future use. Optimal decompositions for a wide range of values of N for both Brownian bridge and Brownian motion are available on the web site `www.quantize.maths-fi.com` [15] for download. The blind optimization procedure is more thoroughly described in [14].

Some values of the optimal decomposition for the standard Brownian motion are given in figure 3.

Proceeding in this article, we will be confronted with other similar optimization problems (with an other criteria than the quadratic distortion). The blind optimization procedure will be the way to compute optimal product decomposition databases.

In figure 4, one can see examples of optimal product quantizers of the Brownian motion and the Brownian bridge on $[0, 1]$. In figure 4, one can see optimal product quantizers of the centred Ornstein-Uhlenbeck process starting from $r_0 = 0$ and a stationary Ornstein-Uhlenbeck on $[0, 3]$.

N	N_{rec}	Quantization Error	N_{rec} decomposition
1	1	0.70711	1
10	10	0.31376	5 - 2
100	96	0.22644	12 - 4 - 2
1000	966	0.18760	23 - 7 - 3 - 2
10000	9984	$0,16259$	26 - 8 - 4 - 3 - 2 - 2

Figure 3: Record of optimal product decomposition values of the standard Brownian motion.

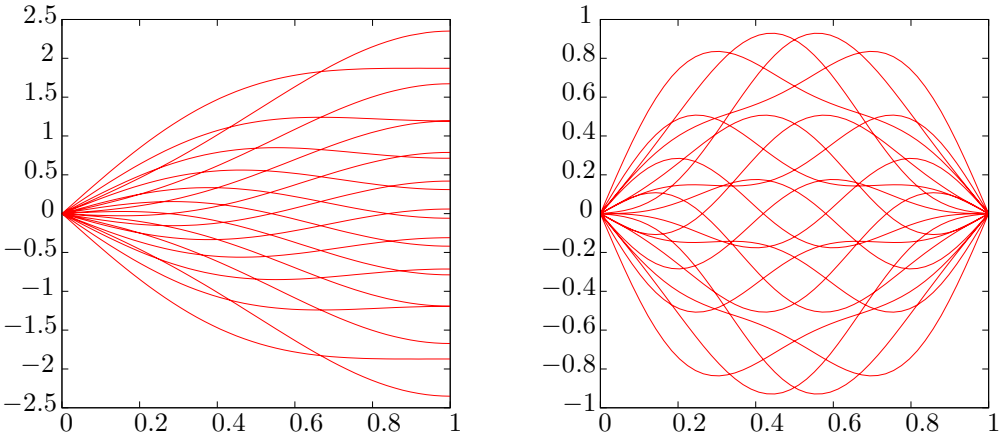

Figure 4: Optimal product quantizer of a standard Brownian motion (left) and a standard Brownian bridge (right) on $[0, 1]$.

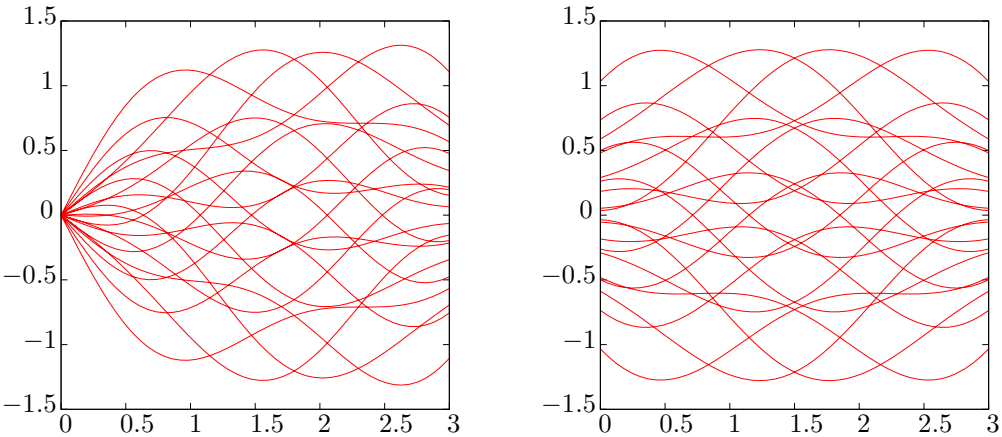

Figure 5: Optimal product quantizer of a centred Ornstein-Uhlenbeck process, starting from $r_0 = 0$ (left) and stationary (right) given by $dr_t = -r_t dt + dW_t$, on $[0, 3]$.

1.6.4 Rate of decay for the quantization error

In [13], H. Luschgy and G. Pagès have established a precise link between the rate problem and Shannon-Kolmogorov's entropy of X. This allowed them to

compute the exact rate of convergence of the minimal L^2-quantization error under rather general conditions on the eigenvalues of the covariance operator. Typical rates are $O(\log(n)^{-a}),\ a > 0$. This conditions are fulfilled by a large class of processes as the Ornstein-Uhlenbeck process and the Brownian motion.

2 Quantization as a control variate: a first attempt to quantization based variance reduction

This method has been initially proposed in [14].

2.1 Quantization as a control variate variable

Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to E$ be a square integrable random vector, $N \in \mathbb{N}^*$ and $\Gamma = \{y_1, \cdots, y_N\}$ a N-codebook. We suppose that we have access to a Γ-valued quantizer $Y = \mathrm{Proj}(X) = \sum_{i=1}^{N} y_i \mathbf{1}_{C_i}(x)$ where $C = \{C_1, \cdots, C_N\}$ is a partition of E. At this step, we do not suppose that Proj is a nearest neighbour projection on Γ.

Let $F : E \to E$ be a Lipschitz continuous function such that $F(X) \in L^2(\mathbb{P})$. In order to compute $\mathbb{E}[F(X)]$, one writes:

$$\begin{aligned}\mathbb{E}[F(X)] &= \mathbb{E}\big[F(\mathrm{Proj}(X))\big] + \mathbb{E}\big[F(X) - F(\mathrm{Proj}(X))\big] \\ &= \underbrace{\mathbb{E}\big[F(\mathrm{Proj}(X))\big]}_{(a)} + \underbrace{\frac{1}{M}\sum_{m=1}^{M} F(X^{(m)}) - F(\mathrm{Proj}(X^{(m)}))}_{(b)} + R_{N,M},\end{aligned} \tag{16}$$

where $X^{(m)}, 1 \le m \le M$ are M independent copies of X, and $R_{N,M}$ is a remainder term defined by equation (16).

Here, term (a) can be computed by quantization and term (b) can be computed by a Monte-Carlo simulation. Now

$$\begin{aligned}\|R_{N,M}\|_2 = \ & \frac{\sigma(F(X)-F(\mathrm{Proj}(X)))}{\sqrt{M}} \le \frac{\|F(X)-F(\mathrm{Proj}(X))\|_2}{\sqrt{M}} \\ & \le [F]_{Lip}\frac{\|X-\mathrm{Proj}(X)\|_2}{\sqrt{M}}.\end{aligned}$$

Furthermore, $\sqrt{M}R_{N,M} \to^{\mathcal{L}} \mathcal{N}\Big(0, \mathrm{Var}\Big(F(X) - F(\mathrm{Proj}(X))\Big)\Big)$.

Consequently, in the d-dimensional case, if F is simply a Lipschitz function and if $(Y_n)_{n\in\mathbb{N}} = (\mathrm{Proj}^N(X))_{N\in\mathbb{N}}$ is a rate optimal sequence of quantizers of X,

$$\|F(X) - F(\mathrm{Proj}^N(X))\|_2 \le [F]_{Lip}\frac{C_X}{N^{1/d}}$$

and

$$\|R_{N,M}\|_2 \le [F]_{Lip}\frac{C_X}{M^{1/2}N^{1/d}}.$$

In a similar way, in the case of the Brownian motion, if $(\widehat{W}^N)_{N\ge1}$ is a rate optimal sequence of product quantization of the Brownian motion, if

F is simply a Lipschitz functional, then

$$\|F(W) - F(\widehat{W}^N)\|_2 \leq [F]_{Lip} \frac{C_W}{\log(N)^{1/2}}$$

and

$$\||R_{N,M}|_{L^2_T}\|_2 \leq [F]_{Lip} \frac{C_W}{M \log(N)^{1/2}}.$$

2.2 Practical implementation: the problem of fast nearest neighbour search

• **The complexity of the projection:** Concerning practical implementation, one notices in equation 16 that for every step of the Monte-Carlo method, one has to compute the projection $\text{Proj}(X^{(m)})$. This is the critical part of the algorithm when dealing with optimal quantization. Hence, the efficiency of the quantization as a control variate variable is conditioned by the efficiency of the projection procedure. When dealing with Voronoi quantization, this projection is simply the nearest neighbour projection.

In a more general setting, the problem of nearest neighbor projection, known as the post office problem has been widely investigated in the area of computational geometry. It is encountered for many applications, as pattern recognition and information retrieval.

The problem has been solved near optimally for the case of low dimensions. Algorithms differ on their practical efficiency on real data sets. For large dimensions, most solutions require storage space exponential in the dimension, or require a bigger query time than the obvious brute force algorithm. In fact for dimension $d > \log N$, a brute force algorithm is usually the best choice. This effect is known as the curse of dimensionality. Still, even in low dimension, fast nearest neighbour search is a critical part of the algorithm.

Concerning vector quantization, the speed of the projection can also be increased by relaxing the hypothesis within the projection on the quantizer is a nearest neighbour projection. It can be done by designing other kind of partitions of the state space. See [5] for more details.

• **The functional case:** One other drawback of the method, when dealing with the functional case is that one does not simulate the whole trajectory of the stochastic process but only its marginals at discretes dates. Hence it is not possible to compute its projection. This problem finds its solution in the simulation scheme for Gaussian processes derived in 4.2 for the functional stratification.

3 Application of quantization to stratification

3.1 Introduction to stratification

The base idea of stratification is to localize the Monte-Carlo method on the element of a measurable partition of the state space of a L^2 random vector $X : (\Omega, \mathcal{A}) \to (E, \varepsilon)$.

- Let $(A_i)_{i\in I}$ be a finite ε-measurable partition of E. The sets A_i are called *strata*. Assume that the weights $p_i = \mathbb{P}(X \in A_i)$ are known for $i \in I$ and strictly positive.
- Let us define the collection of independent random variables $(X_i)_{i\in I}$ with distribution $\mathcal{L}(X|X \in A_i)$.

Remark:

- One assumes that one can write $X_i = \phi_i(U)$ where U is uniformly distributed on $[0,1]^{r_i}$ and $\phi : [0,1]^{r_i} \to \mathbb{R}$ is an easily computable function. (One has $r_i \in \mathbb{N} \cup \{+\infty\}$, the case $r_i = +\infty$ occurs for example in the case of the acceptance-rejection method.) This condition simply means that the random variables $X_i \sim \mathcal{L}(X|X \in A_i)$ are easy to simulate on a computer.
- It is a major constraint for practical implementation of stratification methods. This simulability condition usually has a strong impact on the possible design of the strata. In the following, one will come back several times on this condition.

Let $F : (E, \varepsilon) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mathbb{E}[F^2(X)] < +\infty$.

$$\begin{aligned}\mathbb{E}[F(X)] &= \sum_{i\in I} \mathbb{E}[\mathbf{1}_{\{X_i\in A_i\}} F(X)] = \sum_{i\in I} p_i \mathbb{E}[F(X)|X \in A_i] \\ &= \sum_{i\in I} p_i \mathbb{E}[F(X_i)].\end{aligned}$$

The stratification concept comes into play now. Let M be the global budget allocated to the computation of $\mathbb{E}[F(X)]$ and $M_i = q_i M$ the budget allocated to compute $\mathbb{E}[F(X_i)]$ in each stratus. One assumes that $\sum_{i\in I} q_i = 1$. This leads to define the (unbiased) estimator of $\mathbb{E}[F(X)]$:

$$\overline{F(X)}^I_M := \sum_{i\in I} p_i \frac{1}{M_i} \sum_{k=1}^{M_i} F(X_i^k), \tag{17}$$

where $(X_i^k)_{1\le k\le M_i}$ is a $\mathcal{L}(X|X \in A_i)$-distributed random sample.

Proposition 3.1. *With the same notations:*

$$\mathrm{Var}\left(\overline{F(X)}^I_M\right) = \frac{1}{M} \sum_{i\in I} \frac{p_i^2}{q_i} \sigma^2_{F,i}, \tag{18}$$

where $\sigma^2_{F,i} = \mathrm{Var}(F(X)|X \in A_i) = \mathrm{Var}(F(X_i))\ \forall i \in I$.

Proof: Let us denote $Z_i = \frac{1}{M_i} \sum_{k=1}^{M_i} F(X_i^k)$. For $i \neq j$, Z_i is independent from Z_j.

One has $\overline{F(X)}^I_M = \sum_{i\in I} p_i Z_i$. Hence, by independence,

$$\mathrm{Var}(\overline{F(X)}^I_M) = \sum_{i\in I} p_i^2 \,\mathrm{Var}(Z_i) = \sum_{i\in I} p_i^2 \tfrac{1}{M_i} \,\mathrm{Var}(F(X_i)) = \tfrac{1}{M} \sum_{i\in I} \tfrac{p_i^2}{q_i} \sigma^2_{F,i}.$$

□

Optimizing the simulation allocation to each stratus amounts to solving the following minimization problem:

$$\min_{(q_i)\in\mathcal{P}_I} \sum_{i\in I} \frac{p_i^2}{q_i}\sigma_{F,i}^2 \text{ where } \mathcal{P}_I = \Big\{(q_i)_{i\in I}\in\mathbb{R}_+^I \Big| \sum_{i\in I} q_i = 1\Big\}. \tag{19}$$

3.1.1 Sub-optimal choice

The first natural choice is to set $q_i = p_i, i \in I$. The two motivations for this choice are the facts that the weights p_i are known and because it always reduces the variance.

$$\begin{aligned}\sum_{i\in I}\tfrac{p_i^2}{q_i}\sigma_{F,i}^2 &= \sum_{i\in I} p_i\sigma_{F,i}^2 = \sum_{i\in I}\mathbb{E}\Big[\Big(F(X)-\mathbb{E}[F(X)|X\in A_i]\Big)^2\mathbf{1}_{A_i}(X)\Big]\\ &= \|F(X)-\mathbb{E}[F(X)|\sigma(\{X\in A_i\}, i\in I)]\|_2^2.\end{aligned}$$

Consequently $\sum_{i\in I}\frac{p_i^2}{q_i}\sigma_{F,i}^2 \le \|F(X)-\mathbb{E}[F(X)]\|_2^2 = \mathrm{Var}(F(X))$.

Remark. *When dealing with empirical distributions, in the opinion pool world, this corresponds to the so-called quota method.*

3.1.2 Optimal choice

The optimal choice is the solution of the constrained minimization problem 19.

The Schwartz inequality yields

$$\sum_{i=1}^N p_i\sigma_{F,i} = \sum_{i=1}^N \frac{p_i\sigma F,i}{\sqrt{q_i}}\sqrt{q_i} \le \Big(\sum_{i=1}^N \frac{p_i^2\sigma F,i^2}{q_i}\Big)^{1/2}\underbrace{\Big(\sum_{i=1}^N q_i\Big)}_{=1}^{1/2}.$$

As a consequence, the solution of the minimization problem corresponds to the equality case into the Schwartz inequality.

Hence the solution of the minimization problem is given by $q_i^* = \frac{p_i\sigma_{F,i}}{\sum_{j=1}^N p_j\sigma_{F,j}}$, and the corresponding minimal variance is given by $\Big(\sum_{i=1}^N p_i\sigma_{F,i}\Big)^2$.

At this point, the problem is that one does not know the local inertia $\sigma_{F,i}^2$. Still, using the fact that L^p norms are decreasing with p, one sees that

$$\sigma_{F,i} \ge \mathbb{E}\Big[\big|F(X)-\mathbb{E}[F(X)|\{X\in A_i\}]\big|\Big|\{X\in A_i\}\Big],$$

so that

$$\Big(\sum_{i=1}^N p_i\sigma_{F,i}\Big)^2 \ge \Big\|F(X)-\mathbb{E}\big[F(X)|\sigma(\{X\in A_i\}, i\in I)\big]\Big\|_1^2.$$

In [6], Etoré and Jourdain proposed an algorithm for adaptively modify the proportion of further drawings in each stratum, that converge to the optimal allocation. This can be used in a general framework.

In section 3.2.1, one will see that the problem of designing good strata, in term of Variance reduction is linked with the problem of optimal quantization. Moreover, the case of quantization based strata have two other advantages:

- The weights p_i are already known, what saves us for evaluation their values during the Monte-Carlo evaluation, when dealing with the "sub-optimal" choice.
- What concerns the optimal choice for the allocation parameters q_i, one shows that weights can be chosen such that stratification has a uniform efficiency among the class of Lipschitz continuous functionals. This weights have an explicit expression in the case of quantization based stratification.

3.2 Stratification and quantization

3.2.1 Quantization based stratification

The main drawback of using quantization as a control variate variable is that it requires repeated computations of projections on the quantizer. (That is nearest neighbour search in the case of a Voronoi quantizer.) The point when dealing with stratification is that one does not have to use a projection procedure.

The critical point now is the cost of the simulation of conditional distribution $\mathcal{L}(X|X \in A_i)$.

Theorem 3.2 brings together previous results about stratification and highlights the relationships with the notions of local inertia and intraclass inertia. The other point is that it shows that stratification has a uniform efficiency among the class of Lipschitz continuous functionals.

Theorem 3.2 (Universal stratification)**.** *Let $A = (A_i)_{i\in I}$ be a partition (stratification) of E. (Keep in mind the notation $\mathrm{Proj}_{A,Z}$ for the centroidal projection of the distribution Z on a partition A, defined in definition 1.2).*

1. *For every $i \in I$, consider the local inertia of the random vector X,*

$$\sigma_i^2 = \mathbb{E}\Big[|X - \mathbb{E}[X|X \in A_i]|^2 \Big| X \in A_i\Big].$$

Then, for every Lipschitz continuous function $F : E \to \mathbb{R}$,

$$\forall i \in I, \sigma_{F,i} \leq [F]_{Lip}\sigma_i \text{ so that } \sup_{[F]_{Lip}\leq 1} \sigma_{F,i} \leq \sigma_i. \tag{20}$$

2. *In the case of the sub-optimal choice (see section 3.1.1),*

$$\begin{aligned}\sup_{[F]_{Lip}\leq 1}\Big(\sum_{i\in I} p_i\sigma_{F,i}^2\Big) \leq \sum_{i\in I} p_i\sigma_i^2 &= \Big\|X - \mathbb{E}[X|\sigma(\{X \in A_i\}, i \in I)]\Big\|_2^2 \\ &= \Big\|X - \mathrm{Proj}_{A,X}(X)\Big\|_2^2.\end{aligned} \tag{21}$$

3. *In the case of the optimal choice (see section 3.1.2),*

$$\sup_{[F]_{Lip}\leq 1}\Big(\sum_{i\in I} p_i\sigma_{F,i}^2\Big) \leq \Big(\sum_{i\in I} p_i\sigma_i\Big)^2, \tag{22}$$

and

$$\begin{aligned}\Big(\sum_{i\in I} p_i\sigma_i\Big)^2 \geq &\Big\|X - \mathbb{E}[X|\sigma(\{X \in A_i\}, i \in I)]\Big\|_1^2, \\ &= \Big\|X - \mathrm{Proj}_{A,X}(X)\Big\|_1^2.\end{aligned}$$

4. *If one considers vector valued Lipschitz continuous functions $F : E \to E$, then inequalities (20), (21) and (22) hold as equalities.*

Proof: One has

$$\begin{aligned}\sigma_{F,i}^2 &= \mathrm{Var}(F(X)|X \in A_i)\\ &= \mathbb{E}\Big[\big|F(X) - \mathbb{E}[F(X)|X \in A_i]\big|^2\Big|X \in A_i\Big]\\ &\le \mathbb{E}\Big[\big|F(X) - F(\mathbb{E}[X|X \in A_i])\big|^2\Big|X \in A_i\Big].\end{aligned}$$

Now using that F is Lipschitz, it follows that

$$\sigma_{F,i}^2 \le [F]_{Lip}^2 \frac{1}{p_i}\mathbb{E}\Big[\big|X - \mathbb{E}[X|X \in A_i]\big|^2 \mathbf{1}_{\{X\in A_i\}}\Big] = [F^2]_{Lip}\sigma_i^2.$$

Items 2 and 3 easily follow from item 1. Claim 4 is obvious by considering $F = Id_E$. □

The general case:

The idea is now to use the partition $\{A_1, \cdots, A_N\}$ and the N-codebook $\Gamma = \{y_1, \cdots, y_N\}$ associated with our projection

$$\mathrm{Proj}(x) = \sum_{i=1}^N y_i \mathbf{1}_{A_i}(x).$$

In the case of a Voronoi quantization, this amounts to setting $I = \{1, \cdots, N\}$ and $A_i = \mathrm{slab}_A(x_i)$.

Then for every $i \in \{1, \cdots, N\}$, there exists a Borel function $\phi(x_i, .) : [0,1]^q \to E$ such that $\phi(x_i, U) \sim \mathcal{L}(X|X \in C_i) = \frac{\mathbf{1}_{C_i}\mathbb{P}_X(d\xi)}{\mathbb{P}[X\in C_i]}$, where $U \sim \mathcal{U}([0,1]^q)$.

Remark. *Note that the dimension q is arbitrary: one may always assume that $q = 1$ by the fundamental theorem of simulation, but in order to obtain some closed forms for $\phi(x_i, .)$, we are led to consider situations where $q \ge 2$ or even infinite when considering a Von Neumann acceptance-rejection method.*

Now let (ξ, U) be a couple of independent random variables such that ξ has the distribution of $Y = \mathrm{Proj}(X)$ and $U \sim \mathcal{U}([0,1]^q)$. Then one checks that $\phi(\xi, U)$ has the same distribution as X, so that one may assume without loss of generality that $X = \phi(\mathrm{Proj}(X), U)$ and which in turn implies that $\xi = \mathrm{Proj}(X)$ i.e.

$$X = \phi(\mathrm{Proj}(X), U),\ U \sim \mathcal{U}([0,1]^q), U, \mathrm{Proj}(X) \text{ independent.}$$

In terms of implementation as mentioned above, one needs a simple form for the function ϕ (in term of computational complexity) which induces some stringent constraints on the choice of the strata.

3.3 Closed forms for hyper-rectangles strata in the independent Gaussian case.

Consider a random variable $X \sim \mathcal{N}(0, I_d)$, $d \ge 1$.

Let $(e_1, \cdots, e_d)$ be an orthonormal basis of $E = \mathbb{R}^d$. We set $N_1, \cdots, N_d \ge 1$ the number of strata in each direction. So we consider for $1 \le i \le d$, $-\infty = x_0^i \le x_1^i \le \cdots \le x_{N_i}^i = +\infty$.

The strata are

$$A_{\underline{i}} = \prod_{l=1}^{d} \Big\{ x \in \mathbb{R}^d \text{ such that } \langle e_l, x \rangle \in [y^l_{i_l-1}, y^l_{i_l}] \Big\}, \qquad \underline{i} \in \prod_{l=1}^{d} \{1, \cdots, N_l\}.$$

Then for every multi-index $\underline{i} \in \prod_{l=1}^{d} \{1, \cdots, N_l\}$,

$$\mathcal{L}(X|X \in A_i) = \bigotimes_{l=1}^{d} \mathcal{L}(Z|Z \in [x^l_{i_l-1}, x^l_{i_l}]),$$

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

Then $p_{\underline{i}} = \mathbb{P}(A_i) = \prod_{k=1}^{d} \big(\mathcal{N}(x_{i_k}) - \mathcal{N}(x_{i_{k-1}})\big)$, and for $-\infty \leq a \leq b \leq \infty$,

$$\mathcal{L}(Z|Z \in [a,b]) = \mathcal{N}^{-1}\Big(\big(\mathcal{N}(b) - \mathcal{N}(a)\big)U + \mathcal{N}(a)\Big), \qquad U \sim \mathcal{U}([0,1]). \tag{23}$$

4 Functional stratification of a Gaussian process

In the functional case, the state space of the random values are functional spaces. What is usually done is to simulate a scheme to approximate marginals of the underlying process.

In this section, we assume that X is a $\mathbb{R}$-valued Gaussian process on $[0,T]$. We are interested in the value of $V := \mathbb{E}[F(X_{t_1}, \cdots, X_{t_n})]$ where $0 \leq t_1 < \cdots, < t_n \leq T$ are n dates of interest for the underlying process.

(For example, X can be a standard Brownian motion on $[0,T]$, and V the risk-neutral expectation of a path-dependent payoff of a diffusion based on X.)

What is done in this section can be easily generalized to multi-dimensional processes in the case where their coordinates are independent. (For example, when dealing with multi-factor Brownian diffusions, it does not matter how the Brownian motions are being correlated afterward.) Still we restrict ourselves to the one dimensional setting for clarity.

Let us assume that $\chi \in \mathcal{O}_{pq}(X,N)$ is a K-L quantizer of X. The codebook associated with this product quantizer is the set of the paths of the form

$$\chi_{\underline{i}}(t) = \sum_{n \geq 1} \sqrt{\lambda^X_n} x_i^{(N_n)} e^X_n, \qquad \underline{i} = \{x_1, \cdots, x_n, \cdots\}.$$

with the same notation as in section 1.6.2.

We now need to be able to simulate the conditional distribution

$$\mathcal{L}(X|X \in A_{\underline{i}})$$

where $A_{\underline{i}}$ is the slab associated with $\chi_{\underline{i}}$ in the codebook.

To simulate according to the distribution $\mathcal{L}(X|X \in A_{\underline{i}})$, one will :

- First, simulate the first K-L coordinates and X, using (23).
- Then simulate the conditional distribution of the marginals of the Gaussian process, its first coordinates being settled.

4.1 Simulation of marginals of the Gaussian process, its first K-L coordinates being settled.

In this setting, the aim is to simulate the conditional distribution

$$\mathcal{L}\Big(X_{t_1},\cdots,X_{t_n}\Big|\int_0^T X_s e_1^X ds,\int_0^T X_s e_2^X(s)ds,\cdots,\int_0^T X_s e_d^X(s)ds\Big) \tag{24}$$

where $(X_t)_{t\in}[0,T]$ is a L^2 $\mathbb{R}$-valued Gaussian process, and $(e_k^X,\lambda_k^X)_{k\in\mathbb{N}^*}$ is the Karhunen-Loève system associated with the process X.

As X is a Gaussian process, the random vector

$$\left(X_{t_1},\cdots,X_{t_n},\int_0^T X_s e_1^X(s)ds,\int_0^T X_s e_2^X(s)ds,\cdots,\int_0^T X_s e_d^X(s)ds\right)$$

is a Gaussian vector. Hence, the conditional distribution (24) is given by the transition kernel $\nu(y,A)=\mathcal{N}(Af_1(y)),\mathrm{cov}(X-Af_1(Y)))$, where $Af_1:\mathbb{R}^d\to\mathbb{R}^n$ is an affine function corresponding to the linear regression of X on Y. $Af_1(Y):=\mathbb{E}[X|Y]$.

One has $\Big(\Big(\mathrm{cov}(Y_i,Y_j)\Big)\Big)_{1\le i,j\le d}=\Big(\Big(\lambda_i^X\delta_{ij}\Big)\Big)_{1\le i,j\le d}$, and $\mathrm{cov}(X_{t_i},Y_k)=\mathbb{E}[X_{t_i}\int_0^T X_s e_k^X(s)ds]=\lambda_k^X e_k^X(t_i)$, for $1\le i\le n$ and $1\le k\le d$.

The linear part R_1 of Af_1 is then given by $R_1=\Big(\Big(e_j^X(t_i)\Big)\Big)_{1\le i\le n,1\le j\le n}$.

As a consequence, the conditional expectation writes

$$Af_1(Y)=\mathbb{E}[X]+\begin{pmatrix} e_1^X(t_1) & \cdots & e_d^X(t_1)\\ \vdots & \ddots & \vdots\\ e_1^X(t_n) & \cdots & e_d^X(t_n)\end{pmatrix}\begin{pmatrix} Y_1\\ \vdots\\ Y_d\end{pmatrix}.$$

$$\begin{aligned} K_1:=\mathrm{cov}(X-Af_1(Y)) &= \mathbb{E}\Big[(X-R_1Y)(X-R_1Y)\Big]\\ &= \mathrm{cov}(X)-2\,\mathrm{cov}(X,R_1Y)+\mathrm{cov}(R_1Y),\end{aligned}$$

with $\mathbb{E}[Y_iY_j]=\lambda_i\delta_i^j,\ \forall(i,j)\in\{1,\cdots,d\}$, and

$$\begin{aligned}\mathrm{cov}(R_1Y)_{kl} &= \sum_{i=1}^d \lambda_i e_i^X(t_k)e_i^X(t_l)=\sum_{i=1}^d \lambda_i R_{1ki}e_i(t_l)\\ &= \sum_{i=1}^d R_{1ki}\mathbb{E}[X_lY_i]=\mathbb{E}[XY]_{kl}.\end{aligned}$$

Hence, K_1 is given by $K_{1lk}=\mathrm{cov}(X_l,X_k)-\sum_{i=1}^d \lambda_i e_i^X(t_l)e_i^X(t_k)$.

Now, one is able to simulate according to this probability distribution, using a square root of this matrix.

The easiest way of doing this in the definite positive case is to compute the Cholesky factorization of the matrix K_1, but in this case, the simulation of a simple path requires a $n\times n$ matrix multiplication, whose complexity is quadratic. This solution is not satisfactory for our purpose.

4.2 Faster simulation of conditional paths - Bayesian simulation

The natural method to simulate $\mathcal{L}(X|Y)$ requires for each path a multiplication by a square root of K_1 whose cost is $O(n^2)$. This cost is to high.

- Yet, in the context of this paper, d is the quantization dimension of the process. It is close to $\log(N)$ if N is the number of strata, and n, the number of time steps in the simulation attempts to be very large in comparison to d.
- Moreover, we make the assumption that the simulation of $(X_{t_1}, \cdots, X_{t_n})$ is $O(n)$. (It is the case for the Brownian motion, the Ornstein-Uhlenbeck process and the Brownian bridge.)
- The idea here is that the conditional distribution $\mathcal{L}(X, Y)$ is determined through the Bayes lemma, by the conditional distribution $\mathcal{L}(Y, X)$ and the two marginal distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$. The simulation cost of $\mathcal{L}(Y, X)$, $\mathcal{L}(X)$, $\mathcal{L}(Y)$ are respectively $O(d^2), O(n), O(d^2)$.

One knows that $X = \mathbb{E}[X|Y] + Z$ where $Z \sim \mathcal{N}(Af_1(Y), \mathrm{cov}(X - \mathbb{E}[X|Y]))$ is independent of Y.

Hense one is able to simulate according to $\mathcal{L}(X|Y = y)$ if one can simulate the distribution of Z, by writing

$$\mathcal{L}(X|Y = y) = E[X|Y = y] + \underbrace{\mathcal{L}(Z|Y = y)}_{=\mathcal{L}(Z)}.$$

To simulate Z, one simulates the distribution of X and the conditional distribution $\mathcal{L}(Z|X)$. This yield the following steps :

- Simulate X. (cost of $O(n)$).
- Simulate $\mathcal{L}(Z|X) \sim \mathcal{L}(X - \mathbb{E}[X|Y]|X) \sim X - \mathcal{L}(\mathbb{E}[X|Y]|X)$
  - $\mathcal{L}(\mathbb{E}[X|Y]|X) = Af_1\mathcal{L}(Y|X)$.

    $$\mathcal{L}(Y|X) \sim \mathcal{N}(\mathbb{E}[Y|X], \mathrm{cov}(Y - \mathbb{E}[Y|X]))$$

    Let us denote Af_2 the affine function corresponding to the regression of Y on X and R_2 its linear part.

    $$\begin{aligned}\mathrm{cov}(Y - \mathbb{E}[Y|X]) &= \mathrm{cov}(Y) + \mathrm{cov}(\mathbb{E}[Y|X]) - 2\,\mathrm{cov}(Y, \mathbb{E}[Y|X]) \\ &= \mathrm{cov}(Y) - R_2\,\mathrm{cov}(X)^t R_2.\end{aligned}$$

  - This yields $Z = Af_1(G)$ where $G \sim \mathcal{N}(Af_2(X), \mathrm{cov}(Y) - R_2\,\mathrm{cov}(X)^t R_2)$.
  - This simulation complexity is $O(d \times n)$.
- The random variable $T = Af_1(y) + Z$ satisfies $T \sim \mathcal{L}(X|Y = y)$.

In the general case, the matrix R_2 needed by the method can be computed by performing a least square regression.

Still, in the case of the Brownian motion, the matrix R_2 is explicitly computed. Indeed, one has $\mathrm{cov}(X) = ((\min(t_i, t_j)))_{0\leq i,j\leq n}$. If $t_0 < t_1 < \cdots < t_n$, it is invertible and $R_2 = \mathrm{cov}(Y, X)\,\mathrm{cov}(X)^{-1}$. If $t_j = \frac{jT}{n}$, $0 \leq j \leq n = jh$, this yields $R_2 = ((\alpha_{ij}))_{1\leq i\leq d, 1\leq j\leq n}$, with

- for $j \notin \{1, n\}$, $\alpha_{ij} = \lambda_i^W \frac{he_i^W(t_j) - he_i^W(t_{j-1}) - he_i^W(t_{j+1})}{h^2}$,
- $\alpha_{i0} = \lambda_i^W \left(e_i^{W\prime}(t_0) - \frac{e_i^W(t_1) - e_i^W(t_0)}{h} \right)$,
- $\alpha_{in} = \lambda_i^W \left(\frac{e_i^W(t_n) - e_i^W(t_{n-1})}{h} - e_i^{W\prime}(t_n) \right)$.

The proof is available in section 6. The case of a non uniform subdivision is also handled.

Now, we have a very fast and easy way to simulate the conditional distribution (24) at our disposal.

In figures 6 and 7, we plot a few paths of the conditional distribution of various Gaussian processes knowing that they belong to the L^2 Voronoi cell of one of the curve. The appearance of the drawing suggests to consider the method as a "guided Monte-Carlo simulation".

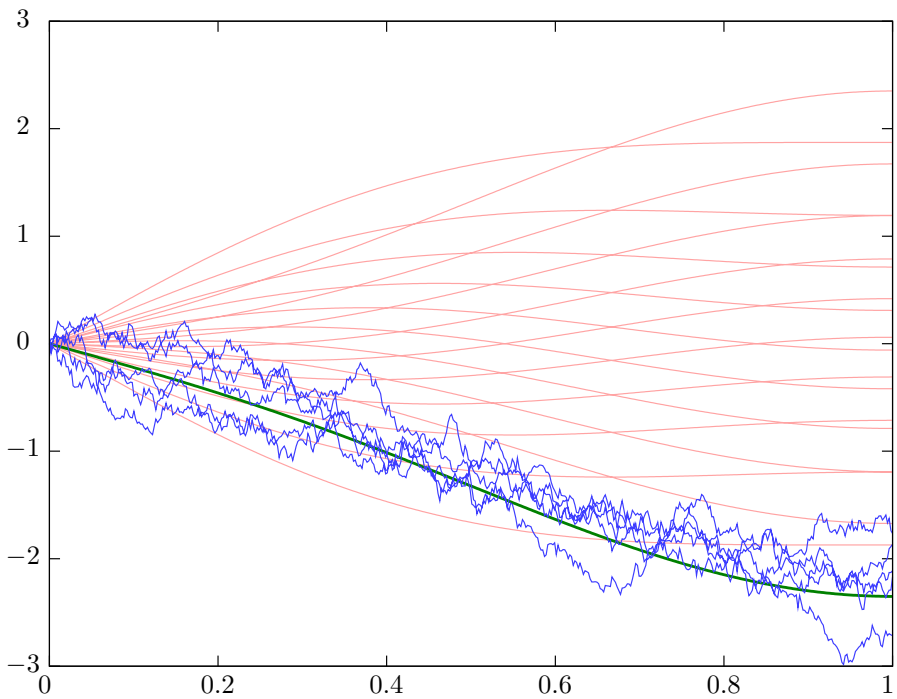

Figure 6: Plot of a few paths of the conditional distribution of the Brownian motion, knowing that its path belong to the L^2 Voronoi cell of the highlighted curve in the quantizer.

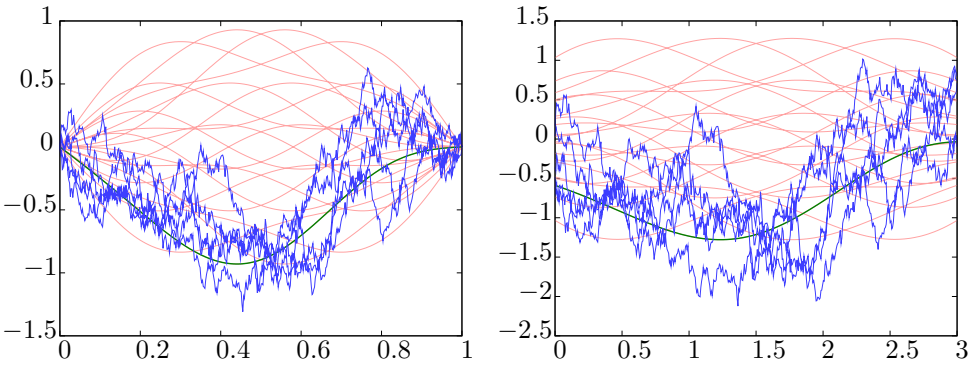

Figure 7: Plot of a few paths of the conditional distribution of the Brownian bridge (left) and the stationary Ornstein-Uhlenbeck process (right), knowing that its path belong to the L^2 Voronoi cell of the highlighted curve in the quantizer.

4.3 Blind optimization procedures

4.3.1 Universal stratification criteria

We have seen in section 3.2.1 that the quantity $d(\chi) = \Big(\sum_{\chi_{\underline{i}} \in \Gamma} p_{\underline{i}} \sigma_{\underline{i}}\Big)^2$ is an upperbound of the Variance of the estimator, given in equation (17) in the case where the functional is 1-Lipschitz continuous. Hence one may want to minimize this criteria instead of the L^2-quantization error.

This yields the minimization problem

$$\min \Big\{ d(\chi), \chi \in \mathcal{O}_{pq}(X, N) \Big\} \tag{25}$$

instead of the minimization problem (14).

The same kind of blind optimization procedure as in section 1.6.3 can be performed. Some values of the optimal decomposition for the standard Brownian motion are given in figure 8.

N	N_{rec}	$d(\chi)$	N_{rec} decomposition
1	1	0.5	1
10	10	$9.75689 \cdot 10^{-2}$	5 - 2
100	96	$5.10548 \cdot 10^{-2}$	12 - 4 - 2
1000	966	$3.51289 \cdot 10^{-2}$	23 - 7 - 3 - 2
10000	9984	$2.63721 \cdot 10^{-2}$	26 - 8 - 4 - 3 - 2 - 2

Figure 8: Record of optimal product decomposition record values of the standard Brownian motion with respect to the criteria (25).

Optimal product decompositions for both Brownian bridge and Brownian motion and for a wide range of values of N are available on the web site `www.quantize.maths-fi.com` [15] for download. One can notice that the obtained optimal decompositions for this criterium are the same than in table 3. Comparing all the decompositions obtained for a quantizer size smaller than

11000, one notices that in the case of the Brownian motion, the optimal decompositions for both criteria are "almost" always the same. The only values when decompositions differ are the ranges $270 - 271$ and $3328 - 3359$.:

In the case of the Brownian motion, the two criteria do not have a very different values for the two decompositions, hence one can use the same decomposition database for the two applications. Concerning the Brownian bridge, one notice that there are more differing values between the optimal decompositions for the variance and the optimal decomposition for the L^2-distortion.

4.3.2 Composition with the dimensionality

The problem of the simulation complexity

When performing a Monte-Carlo simulation, one makes the number of simulations that is needed to reach a given accuracy, which is given by a confidence interval. Hence, when the number of simulation goes to infinity, as the convergence rule is driven by the Central Limit Theorem, the number of needed simulation M is proportional to the Variance of the simulated random variable.

$$M \propto \big(\mathrm{Var}(X)\big). \tag{26}$$

Now, if κ is the simulation complexity of a path of X, the complexity C of the whole Monte-Carlo simulation is proportional to the product $\kappa \times \mathrm{Var}(X)$.

$$C \propto \big(\kappa \times \mathrm{Var}(X)\big). \tag{27}$$

Hence if one knows an other random variable X' having the same expectation, but an other variance X' and an other simulation cost κ', one will choose to use X' instead of X only if

$$C' = \big(\kappa' \times \mathrm{Var}(X')\big) < \big(\kappa \times \mathrm{Var}(X)\big) = C. \tag{28}$$

As a consequence, the criteria that has to be minimized in any Variance reduction method is the quantity C', and not only the variance.

The case of functional stratification

In our case, one acknowledged in section 4.2 that the simulation complexity of a conditional path is proportional to

$$(C_1 + d) \times n + (C_2 \times d^2 + C_3 \times d),$$

where n is the number of dates, d is the quantization dimension and C_1, C_2 and C_3 are constants independent of d and n. The quantity $C_1 \times n$ corresponds to the cost of the simulation of the first unconditioned path (first step in the algorithm), and is related to the random number generator that is used.

Assuming that the number of dates is much bigger than the quantization dimension (one reaches a quantization dimension of 9 with a quantizer size of 10^4 with the two criteria (14) and (25)), one has to minimize the quantity $(C_1 + d) \times n \times d(\chi)$ in the case of the optimal weights, and $(C_1 + d) \times n \times e(\chi)$ in the case of the standard weights.

To compute the corresponding optimal decomposition, one propose to evaluate the constant C_1 in the case of the user implementation, and perform an other blind optimization procedure to obtained the corresponding optimal decomposition.

In the case of the author's implementation (with a Mersenne Twister pseudo random generator), the constant C_1 is close to 3.0.

5 Application to option pricing

Now, one is able to simulate the conditional law of a Gaussian process, knowing it's Voronoi cell in a product quantizer. One condition is to know an orthonormal Hilbert basis that diagonalizes its covariance operator. The cases of the Brownian motion, the Brownian bridge and the Ornstein-Uhlenbeck process have been handled.

The particular case of the Brownian motion allows to use functional stratification as a generic Variance reduction method for the case of functional of Brownian diffusions. Even in the multi-dimensonal case, no matter how the independent Brownian motions are correlated or used afterwards ; no matter if it is used for diffusing the underlying stock, a stochastic volatility process or an actualization factor. It can be used as a variance reduction method.

Hence, this is a very interesting variance method to be used "in production" when using Monte-Carlo simulations in a industrial way, independently of the path-dependent payoff or the model (as soon as it uses Brownian diffusions or one of the other proposed Gaussian processes). User does not have to set up complicated adjustment when using it.

In the following of this section, the method is used to illustrate its performance on simple one dimensional cases. One begins with the case of a continuous time Up-In Call in the Black and Scholes model, for which a closed formula is known, and used as a Benchmark.

Benchmark with a Up-In-Call pricing in the Black and Scholes model.

Here, one benchmarks the numerical method for a path dependent option in a case where a theoretical value is known : a barrier option in the Black and Scholes Model.

For simplicity's sake, consider a log - normal Black and Scholes diffusion with no drift (no interest rate and no divident).

One has a closed form for the continuous barrier option. A numerical correction proposed by Broadie and Glasserman [4] is done to get the theoretical price to be compared to. The number of Monte-Carlo Simulation is 100000 in every case.

One prices a Up-In-Call with different values of the initial spot S, the strike K, the barrier H, the volatility σ, the maturity T, and the number of fixing dates for the discrete barrier n. In every case, a 95% confidence interval is given. So is the variance of the estimator.

The numerical results are given in table 9 when using the method with 20 stratas and table 10 when using the method with 100 stratas.

6 Annex: Special case of the Brownian motion for R_2 computation.

In this section, we give the closed form of the matrix $R_2 := ((\alpha_{ij}))_{1\leq i\leq d, 1\leq j\leq n} \in M_{d,n}(\mathbb{R})$ corresponds to the affine function Af_2 defined by $\mathbb{E}[Y|X] = Af_2(X)$.

Parameters	Theoritical value	Simple Estimator	Strat. Estimator sub-optimal weights	Strat. Estimator Lip.-optimal weights	Strat. Estimator Optimal weights
$S = 100,\ K = 100$ $H = 125,\ \sigma = 0.3,$ $T = 1.5,\ n = 365$	13.9597	14.0379 $[13.8705, 14.2053]$ Var = 729.2518	13.9281 $[13.8491, 14.0071]$ Var = 162.4650	13.9283 $[13.8519, 14.0047]$ Var = 151.9481	13.9617 $[13.9069, 14.0165]$ Var = 78.1973
$S = 100,\ K = 100$ $H = 200,\ \sigma = 0.3,$ $T = 1,\ n = 365$	1.3665	1.4206 $[1.3442, 1.4969]$ Var = 151.6366	1.3659 $[1.3106, 1.4211]$ Var = 79.5118	1.3510 $[1.3039, 1.3981]$ Var = 57.7425	1.35828 $[1.3460, 1.3706]$ Var = 3.9550

Figure 9: Numerical results for the Up In Call option, with 20 stratas.

Parameters	Theoritical value	Simple Estimator	Strat. Estimator sub-optimal weights	Strat. Estimator Lip.-optimal weights	Strat. Estimator Optimal weights
$S = 100,\ K = 100$ $H = 125,\ \sigma = 0.3,$ $T = 1.5,\ n = 365$	13.9597	14.0379 $[13.8705, 14.2053]$ Var = 729.2518	13.9382 $[13.8720, 14.0043]$ Var = 114.0634	13.9511 $[13.8874, 14.0150]$ Var = 105.8760	13.9450 $[13.9012, 13.9888]$ Var = 50.0103
$S = 100,\ K = 100$ $H = 200,\ \sigma = 0.3,$ $T = 1,\ n = 365$	1.3665	1.4206 $[1.3442, 1.4969]$ Var = 151.6366	1.3296 $[1.2825, 1.3768]$ Var = 57.8899	1.3493 $[1.3093, 1.3893]$ Var = 41.6666	1.3563 $[1.3461, 1.3664]$ Var = 2.6831

Figure 10: Numerical results for the Up In Call option, with 100 stratas.

Consider $t_0 = 0 \leq t_1 \leq \cdots \leq t_n = T$ a subdivision of $[0, T]$.

$$\mathbb{E}\Big[\int_0^T W_s e_i^W(s)ds | W_{t_1}, \cdots, W_{t_n}\Big] = \sum_{j=0}^{n-1} \underbrace{\mathbb{E}\Big[\int_{t_j}^{t_{j+1}} W_s e_i^W(s)ds \Big| W_{t_j}, W_{t_{j+1}}\Big]}_{=f_j^i(W_{t_j}, W_{t_{j+1}})},$$

where f_j^i is an affine function.

If $t_j \neq t_{j+1}$, $f_j^i(x, y) = \mathbb{E}\Big[\int_{t_j}^{t_{j+1}} \Big(x + \frac{s-t_j}{t_{j+1}-t_j}(y-x) + (Y_{s-t_j}^{B,t_{j+1}-t_j})\Big) e_i^W(s)ds\Big]$ (where $Y_{s-t_j}^{B,t_{j+1}-t_j}$ is a standard Brownian bridge)

$$= x \underbrace{\Big(\int_{t_j}^{t_{j+1}} \frac{t_{j+1}-s}{t_{j+1}-t_j} e_i^W(s)ds\Big)}_{:=A_j} + y \underbrace{\Big(\int_{t_j}^{t_{j+1}} \frac{s-t_j}{t_{j+1}-t_j} e_i^W(s)ds\Big)}_{:=B_j} = xA_j^i + yB_j^i.$$

Simple computations lead to:

$$\int_{t_j}^{t_{j+1}} e_i^W(s)ds = \sqrt{\frac{2}{T}} \frac{T}{\pi(i-\frac{1}{2})} \Big(\cos\big(\pi(i-\frac{1}{2})\frac{t_j}{T}\big) - \cos\big(\pi(i-\frac{1}{2})\frac{t_{j+1}}{T}\big)\Big),$$

and

$$\begin{aligned}\int_{t_j}^{t_{j+1}} s e_i^W(s)ds &= \sqrt{\tfrac{2}{T}} \tfrac{T}{\pi(i-\frac{1}{2})} \Big(t_j \cos(\pi(i-\tfrac{1}{2})\tfrac{t_j}{T}) - t_{j+1}\cos(\pi(i-\tfrac{1}{2})\tfrac{t_{j+1}}{T})\Big) \\ &+ \sqrt{\tfrac{2}{T}} \Big(\tfrac{T}{\pi(i-\frac{1}{2})}\Big)^2 \Big(\sin(\pi(i-\tfrac{1}{2})\tfrac{t_{j+1}}{T}) - \sin(\pi(i-\tfrac{1}{2})\tfrac{t_j}{T}))\Big).\end{aligned}$$

Hence $\mathbb{E}\Big[\int_0^T W_s e_i^W(s)ds | W_{t_1}, \cdots, W_{t_n}\Big] = \sum_{j=0}^{n-1} A_j^i W_{t_j} + B_j^i W_{t_{j+1}} = \sum_{i=0}^{n} \alpha_{ij} W_{t_i}$ with, for every $1 \leq j < n$, $\alpha_{ij} = A_j^i + B_{j-1}^i$, $\alpha_{i0} = A_0^i$ and $\alpha_{in} = B_{n-1}^i$.

Finally one gets the following closed forms for $R_2 := ((\alpha_{ij}))_{1 \leq i \leq d, 1 \leq j \leq n}$.

- If $t_{j-1} < t_j < t_{j+1}$,

$$\alpha_{ij} = \lambda_i^W \frac{(t_{j+1}-t_{j-1})e_i^W(t_j) - (t_{j+1}-t_j)e_i^W(t_{j-1}) - (t_j - t_{j-1})e_i^W(t_{j+1})}{(t_{j+1}-t_j)(t_j-t_{j-1})}.$$

  If $t_{j-1} = t_j < t_{j+1}$,

$$\alpha_{ij} = \lambda_i^W \Big(\Big(-\frac{e_i^W(t_{j+1})}{t_{j+1}-t_j} + \frac{e_i^W(t_j)}{t_{j+1}-t_j} \Big) + e_i^{W\prime}(t_j) \Big).$$

  If $t_{j-1} < t_j = t_{j+1}$,

$$\alpha_{ij} = \lambda_i^W \Big(-e_i^{W\prime}(t_j) + \Big(\frac{e_i^W(t_j)}{t_j - t_{j-1}} - \frac{e_i^W(t_{j-1})}{t_j - t_{j-1}} \Big) \Big).$$

  If $t_{j-1} = t_j = t_{j+1}$, $\alpha_{ij} = 0$.

- $\alpha_{i0} = A_0^i = \begin{cases} \lambda_i^W \left(e_i^{W\prime}(t_0) - \frac{e_i^W(t_1) - e_i^W(t_0)}{t_1 - t_0} \right) & \text{if } t_1 \neq t_0, \\ 0 \text{ in the other case.} \end{cases}$

- $\alpha_{in} = B_{n-1}^i = \begin{cases} \lambda_i^W \left(\frac{e_i^W(t_n) - e_i^W(t_{n-1})}{t_n - t_{n-1}} - e_i^{W\prime}(t_n) \right) & \text{if } t_n \neq t_{n-1}, \\ 0 \text{ in the other case.} \end{cases}$

7 Annex: Computation of the Karhunen-Loève decomposition the Ornstein-Uhlenbeck process

7.1 The Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck process is defined by the SDE

$$dr_t = \theta(\mu - r_t)dt + \sigma dW_t, \qquad \text{with } \sigma \geq 0 \text{ and } \theta > 0. \tag{29}$$

The equation is solved by applying Itô's formula to the process $U_t := r_t e^{\theta t}$. One gets

$$r_t = r_0 e^{-\theta t} + \mu(1 - e^{-\theta t}) + \int_0^t \sigma e^{\theta(s-t)} dW_s. \tag{30}$$

If one assumes that r_0 is Gaussian ($r_0 \sim \mathcal{N}(m_0, \sigma_0)$) and is independent from $\bigvee_{s \in \mathbb{R}_+} \sigma\{W_t, t < s\}$, the process $(r_t)_{t>0}$ is Gaussian.

One has

$$\mathbb{E}[r_t] = m_0 e^{-\theta t} + \mu(1 - e^{-\theta t})$$

and

$$\operatorname{cov}(r_s, r_t) = \frac{\sigma^2}{2\theta} e^{-\theta(s+t)} (e^{2\theta \min(s,t)} - 1) + \sigma_0^2 e^{-\theta(s+t)}.$$

One easily sees that $\lim_{t\to\infty} \operatorname{Var}(r_t) = \frac{\sigma^2}{2\theta}$ (the long term variance). If the initial variance σ_0^2 is equal to long term variance $\frac{\sigma^2}{2\theta}$, the process is stationary and the covariance writes $\operatorname{cov}(r_s, r_t) = \frac{\sigma^2}{2\theta} e^{-\theta|s-t|}$.

The total variance of the process on $[0, T]$ is

$$\|r_2\|_2^2 = \int_0^T \operatorname{Var}(r_s) ds = \frac{\sigma^2 T}{2\theta} + \Big(\sigma_0^2 - \frac{\sigma^2}{2\theta} \Big) \Big(\frac{1}{2\theta} - \frac{e^{-2\theta T}}{2\theta} \Big)$$

7.2 The Ornstein-Uhlenbeck covariance operator:

The Ornstein-Uhlenbeck covariance operator is given by

$$T^{OU}f(s) = \int_0^T \frac{\sigma^2}{2\theta} e^{-\theta(s+t)}\Big(e^{2\theta\min(s,t)}-1\Big)f(s)ds + \int_0^T \sigma_0^2 e^{-\theta(s+t)} f(s)ds. \quad (31)$$

Computing the Karhunen-Loève decomposition of the Ornstein-Uhlenbeck process

T^{OU} is a compact Hermitian positive operator on the separable Hilbert space $L^2([0,T])$. Hence there is an orthonormal basis of V consisting of eigenvectors of T^{OU} and each eigenvalue is real and strictly positive. Moreover $\|T^{OU}\|^2 \leq \frac{\sigma^2 T}{2\theta} + \frac{\sigma^2}{4\theta^2}\Big(e^{-2\theta T}-1\Big)$.

One has

$$T^{OU}f(t) = \int_0^t \frac{\sigma^2}{2\theta} e^{\theta(s-t)} f(s)ds + \int_t^T \frac{\sigma^2}{2\theta} e^{\theta(t-s)} f(s)ds + \int_0^T (\sigma_0^2 - \frac{\sigma^2}{2\theta}) e^{-\theta(s+t)} f(s)ds.$$

Proposition 7.1. *If $f \in C([0,1])$, and if $g = T^{OU}f$, then*

$$g'' - \theta^2 g = -\sigma^2 f, \quad (32)$$

with

$$\sigma_0^2 g'(0) = \Big(\sigma^2 - \theta\sigma_0^2\Big)g(0) \qquad \text{and} \qquad g'(T) = -\theta g(T) \quad (33)$$

Proof:

$$g(t) = \int_0^t \frac{\sigma^2}{2\theta} e^{\theta(s-t)} f(s)ds + \int_t^T \frac{\sigma^2}{2\theta} e^{\theta(t-s)} f(s)ds + \int_0^T (\sigma_0^2 - \frac{\sigma^2}{2\theta}) e^{-\theta(s+t)} f(s)ds.$$

$$g'(t) = -\frac{\sigma^2}{2}\int_0^t e^{\theta(s-t)} f(s)ds + \frac{\sigma^2}{2}\int_t^T e^{\theta(t-s)} f(s)ds - (\theta\sigma_0^2 - \frac{\sigma^2}{2})\int_0^T e^{-\theta(s+t)} f(s)ds$$

$$g''(t) = \frac{\sigma^2\theta}{2}\Big[\int_0^t f(s)e^{\theta(s-t)}ds + \int_t^T f(s)e^{\theta(t-s)}ds\Big] + \theta\int_0^T (\theta\sigma_0^2 - \frac{\sigma^2}{2}) e^{-\theta(s+t)} f(s)ds - \sigma^2 f(t).$$

One gets $g''(t) = \theta^2 g(t) - \sigma^2 f(t)$. Moreover, equation (33) comes when identifying expressions with $t = 0$ and $t = T$. □

Proposition 7.2. *Conversely, if $g \in C^2([0,T])$ and if f satisfies equations (32) and (33) then $g = T^{OU}f$.*

Proof: Computing $T^{OU}g''$ yields:

$$T^{OU}g'' = \int_0^t \tfrac{\sigma^2}{2\theta} e^{\theta(s-t)} g''(s)ds + \int_t^T \tfrac{\sigma^2}{2\theta} e^{\theta(t-s)} g''(s)ds + \int_0^T \Big(\sigma_0^2 - \tfrac{\sigma^2}{2\theta}\Big) g''(s)ds$$

An integration by parts twice yields

$$\begin{aligned} T^{OU}g'' &= -\sigma_0^2 g'(0)e^{-\theta t} - \sigma^2 g(t) + \tfrac{\sigma^2}{2} g(0)e^{-\theta t} - \Big(\theta\sigma_0^2 - \tfrac{\sigma^2}{2}\Big) g(0)e^{-\theta t} + \theta^2 T^{OU} g(t) \\ &= -\sigma^2 g(t) + \theta^2 T^{OU} g(t) \quad \text{thanks to equation (33).} \end{aligned}$$

Now, by necessary conditions, $T^{OU}f = \lambda f \Leftrightarrow \sigma^2 g = \lambda(\theta^2 g - g'')$. One obtains

$$\lambda g'' + (\sigma^2 - \lambda\theta^2)g = 0. \tag{34}$$

□

Hence the solution of the ordinary differential equation (34) on $[0, T]$ has the form $g(t) = A\cos(\omega t) + B\sin(\omega t)$, with $\omega = \sqrt{\frac{\sigma^2-\lambda\theta^2}{\lambda}} \Leftrightarrow \lambda = \frac{\sigma^2}{\omega^2+\theta^2}$.

Equation (33) yields $\omega B\sigma_0^2 = (\sigma^2 - \theta\sigma_0^2)A$. Hence, function $g(x)$ writes

$$g(t) = K\Big(\omega\sigma_0^2\cos(\omega t) + (\sigma^2 - \theta\sigma_0^2)\sin(\omega t)\Big).$$

Hence $g'(T) = -\theta g(T)$ yields

$$\omega\sigma^2\cos(\omega T) + \big[-\omega^2\sigma_0^2 + \theta\sigma^2 - \theta^2\sigma_0^2\big]\sin(\omega T) = 0. \tag{35}$$

Conversely, by the same computation, one sees that $\lambda_n \in]0, \|T^{OU}\|_2]$ is an eigenvalue of T^{OU} if and only if equation (35) is fulfilled.

Proposition 7.3. *Finaly, if one knows the sorted increasing sequence (ω_n) of the stricly positive solutions of equation (35), the Karhunen-Loève basis $(\lambda_n^{OU}, e_n^{OU})$ of the Ornstein-Uhlenbeck covariance operator T^{OU} are given by: $\lambda_n^{OU} = \frac{\sigma^2}{\omega_n^2+\theta^2}$, and*

$$e_n^{OU}(t) = K_n\Big(\omega_n\sigma_0^2\cos(\omega_n t) + (\sigma^2 - \theta\sigma_0^2)\sin(\omega_n t)\Big),$$

where K_n is the normalization constant. If $(\sigma, \sigma_0) \neq (0, 0)$, K_n is given by

$$1/K_n^2 = \frac{1}{2w_n}\sigma_0^2(\sigma^2 - \theta\sigma_0^2)[1 - \cos(2\omega_n T)] + \frac{1}{2}\sigma_0^4\omega_n^2\Big(T + \frac{1}{2\omega_n}\sin(2\omega_n T)\Big) + \frac{1}{2}(\sigma^2 - \theta\sigma_0^2)^2\Big(T - \frac{1}{2\omega_n}\sin(2\omega_n T)\Big). \tag{36}$$

Case of a deterministic start point: In this case ($\sigma_0 = 0$), one has

$$e_n^{OU}(t) = \frac{1}{\sqrt{\frac{T}{2} - \frac{\sin(2\omega_n T)}{4\omega_n}}}\sin(\omega_n t).$$

Stationary case: In the stationary case, $\sigma_0^2 = \frac{\sigma^2}{2\theta}$, one has

$$e_n^{OU}(t) = C_n\Big(\omega_n\cos(\omega_n t) + \theta\sin(2\omega_n t)\Big),$$

where C_n is the normalization constant. C_n is given by

$$1/C_n^2 = \frac{\theta}{2}\Big(1 - \cos(2\omega_n T)\Big) + \frac{\omega_n^2}{2}\Big(T + \frac{\sin(2\omega_n T)}{2\omega_n}\Big) + \frac{\theta^2}{2}\Big(T - \frac{\sin(2\omega_n T)}{2\omega_n}\Big).$$

7.3 Numerical computation of the Karhunen-Loève decomposition of the Ornstein-Uhlenbeck process

As we have seen in previous section, everything comes to evaluate numerically the stritly positive solutions of equation (35).

7.3.1 Deterministic start point

In this case, ($\sigma_0 = 0$), one can check that elements of $\{\frac{\pi}{2} + k\frac{\pi}{T} | k \in \mathbb{N}\}$ are not solutions of equation (35), thus the equation comes to

$$\theta \tan(\omega T) = -\omega. \tag{37}$$

The case where $\theta = 0$ comes to the case of the Brownian motion, hence one assumes that $\theta \neq 0$. Solutions of this equation are illustrated in figure 11. One can easily show that a unique solution w_n lies in each interval $\left]\frac{n\pi}{T} - \frac{\pi}{2T}, \frac{n\pi}{T}\right[$, for $n \in \mathbb{N}^*$.

$$\lim_{n\to\infty} \omega_n - \left(\frac{n\pi}{T} - \frac{\pi}{2T}\right) = 0.$$

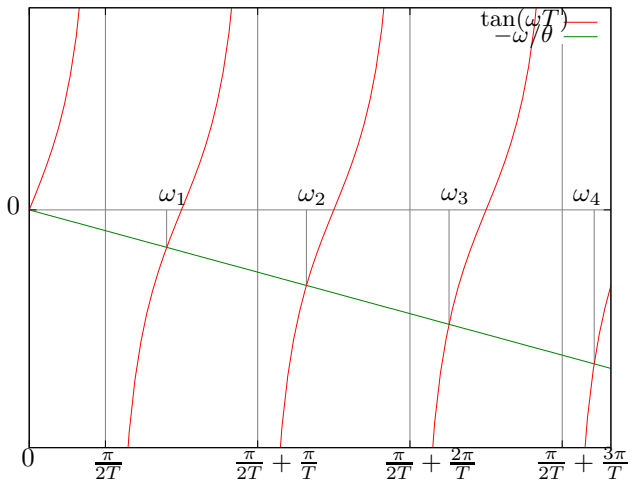

Figure 11: **(Deterministic start point)**. Solutions of equation (37). (Ornstein-Uhlenbeck process starting from a determined point r_0, $\sigma_0 = 0$.) Values for this figure are $T = 3$, $\sigma = 1$ and $\theta = 3$.

7.3.2 Non deterministic start point

Let us assume now that $\sigma_0 \neq 0$ and consider equation (35) again.

The term $-\omega^2\sigma_0^2 + \theta\sigma^2 - \theta^2\sigma_0^2$ never vanishes one $]0, +\infty[$ if $\theta^2\sigma_0^2 - \theta\sigma^2 \geq 0$.

First case: $\theta^2\sigma_0^2 - \theta\sigma^2 \geq 0$.

Here, everything comes to the equation

$$\tan(\omega T) = \frac{\omega\sigma^2}{\omega^2\sigma_0^2 + \theta^2\sigma_0^2 - \theta\sigma^2}. \tag{38}$$

Solutions of this equation are illustrated in figure 12.

One can easily show that $\forall n \in \mathbb{N}^*, \exists! \omega \in \left]\frac{n\pi}{T}, \frac{n\pi}{T} + \frac{\pi}{2T}\right[$, that is solution of equation (35). Moreover a solution lies in $]0, \frac{\pi}{2T}[$ if and only if $(\theta^2\sigma_0^2 - \theta\sigma^2)T - \sigma^2 < 0$.

Second case: $\theta^2\sigma_0^2 - \theta\sigma^2 < 0$.

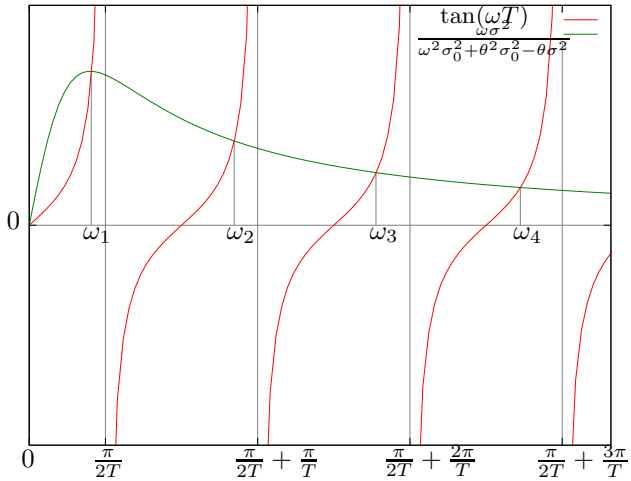

Figure 12: **(Non deterministic start point, $\theta^2\sigma_0^2 - \theta\sigma^2 \geq 0$).** Solutions of equation (37). (Ornstein-Uhlenbeck process starting from $r_0 \sim \mathcal{N}(0, \sigma_0^2)$, $\sigma_0 \neq 0$.) Values for this figure are $T = 3$, $\sigma = 1$, $\theta = 3$ and $\sigma_0^2 = 0.4$.

Here, the term $-\omega^2\sigma_0^2+\theta\sigma^2-\theta^2\sigma_0^2$ vanishes for $\omega = V := \sqrt{\theta\frac{\sigma^2}{\sigma_0^2} - \theta^2}$. If V is not a solution of equation (35), (that is if V does not belong to $\{\frac{\pi}{2T}+k\frac{\pi}{T}|k \in \mathbb{N}\}$), no other value of this set is a solution, and everything comes again to the same equation (38).

Solutions of this equation are illustrated in figure 13 a One can then easily show that $\forall n \in \mathbb{N}^* \cap]V, +\infty[, \exists! \omega \in \left]\frac{n\pi}{T}, \frac{n\pi}{T} + \frac{\pi}{2T}\right[$, that is solution of equation (35). Moreover, in every non empty interval $\left]\frac{k\pi}{T} - \frac{\pi}{2T}, \frac{k\pi}{T} + \frac{\pi}{2T}\right[\cap \left]0, V\right[$, $k \in \mathbb{N}^*$ there is an other solution of the equation.

7.3.3 Pseudo-Code for computing Ornstein-Uhlenbeck eigenvalues

A pseudo-code for the computation of the n-th eigenvalue of the Ornstein-Uhlenbeck covariance operator is given in algorithm 1. We point out the fact that, in this pseudo-code, eigenvalues indexes start at 1 what is consistent with previous notations. In the practical implementation, one prefer to start indexing at 0. In this code, the function **search**$(a, \text{left}, \text{right})$ stands for a root finding method. It fills argument a with the root of equation (35) that is bracketted by $[\text{left}, \text{right}]$.

In the author's implementation, one uses the Brent method as reliable root finding method for equation (35). As Newton-likes method, Brent method can take advantage of a guess of the value of the root as soon as one has a bracketting method. (The idea is to start from a small interval arround the guess that is geometrically expanded, until the limiting range $[\text{left}, \text{right}]$ is reached.)

In section 7.3.4, an accurate numerical guess of ω_n for the deterministic start point is proposed.

Algorithm 1 $OUEigenvalue(\theta, \sigma, \sigma_0, T, n)$

Require: $\theta > 0,\ \sigma \geq 0,\ \sigma_0 \geq 0,\ T \geq 0,\ n \geq 1.$

```
if σ_0 = 0 then
    {There is a unique solution w_n of (35) in the interval ]nπ/T − π/2T, nπ/T[. }
    search(w_n, nπ/T − π/2T, nπ/T).
else
    Here σ_0 > 0.
    if (θ^2 σ_0^2 − θσ^2) ≥ 0 then
        {The vertical asymptote of the right hand of equation 38 lies on the left of 0. }
        if (θ^2 σ_0^2 − θσ^2)T − σ^2 < 0 then
            {There is a unique solution w_n of (35) in the interval ]0, π/2T[. }
            search(w_n, (n−1)π/T, (n−1)π/T + π/2T).
        else
            {The smaller stricly positive solution w_1 of equation (35) lie in the interval ]π/2T, π/T[.}
            search(w_n, nπ/T, nπ/T + π/2T).
        end if
    else
        {The vertical asymptote of the right hand of equation 38 lies on the right of 0. }
        if (n−1)π/T − π/2T > sqrt(θ σ^2/σ_0^2 − θ^2) then
            search(w_n, (n−1)π/T, (n−1)π/T + π/2T).
        else if (n+1)π/T − π/2T < sqrt(θ σ^2/σ_0^2 − θ^2) then
            search(w_n, nπ/T − π/2T, nπ/T).
        else if nπ/T − π/2T < sqrt(θ σ^2/σ_0^2 − θ^2) and (n+1)π/T − π/2T > sqrt(θ σ^2/σ_0^2 − θ^2) then
            search(w_n, nπ/T − π/2T, sqrt(θ σ^2/σ_0^2 − θ^2)).
        else
            search(w_n, sqrt(θ σ^2/σ_0^2 − θ^2), nπ/T − π/2T).
        end if
    end if
end if
λ_n ← σ^2/(ω_n^2 + θ^2).
Return λ_n.
```

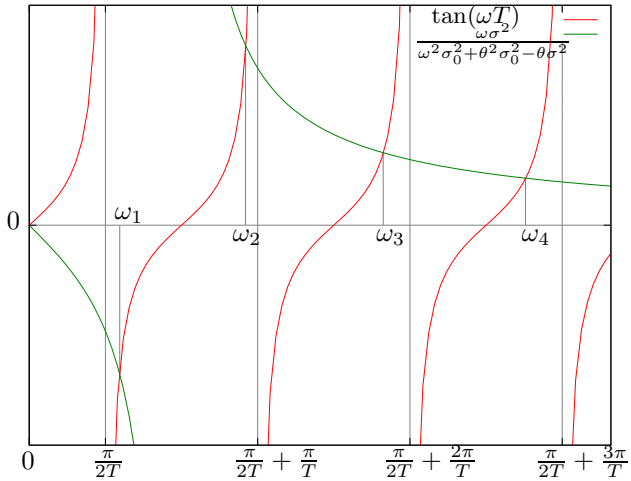

Figure 13: **(Non deterministic start point, $\theta^2\sigma_0^2 - \theta\sigma^2 < 0$).** Solutions of equation (37). (Ornstein-Uhlenbeck process starting from $r_0 \sim \mathcal{N}(0, \sigma_0^2)$, $\sigma_0 \neq 0$.) Values for this figure are $T = 3$, $\sigma = 1$, $\theta = 3$ and $\sigma_0^2 = 0.3$.

7.3.4 A numerical guess for the value of ω_n.

As one has seen, one uses a root finding method for evaluating numerically the value of ω_n. In this section, one proposes a numerical guess for this quantity, that can be used as a starting point in the root finding method.

The starting point is to approximate function tan on interval $]-\frac{\pi}{2}, \frac{\pi}{2}[$ by the rational fraction $\text{tanapprox}(x) := \frac{\frac{4(8-\pi^2)x^3}{\pi^4}+x}{1-\frac{4x^2}{\pi^2}}$, which is a good uniform approximation of tan on $]-\frac{\pi}{2}, \frac{\pi}{2}[$. One has $\|\tan - \text{tanapprox}\|_{\infty,]-\frac{\pi}{2},\frac{\pi}{2}[} = \frac{1}{2}\frac{10-\pi^2}{\pi} \approx 0.02075310$.

Now, in the case of the Ornstein-Uhlenbeck eigenvalue computation, for a deterministic start point, equation 37 can be approximated by

$$\theta\,\text{tanapprox}(\omega_n T + n\pi) = -\omega_n \quad n \geq 1. \tag{39}$$

This comes to a polynomial equation of degree 3 for every $n > 0$ which has a unique solution $w_n^{guess} \in]\frac{n\pi}{T} - \frac{\pi}{2T}, \frac{n\pi}{T}[$.

Using this numerical guess yields a good accuracy for approximating the real value of ω_n.

References

[1] Vlad Bally, Gilles Pagès, and Jacques Printems. A quantization tree method for pricing and hedging multidimensional american options. *Mathematical Finance*, 15(1):119–168, 2005.

[2] Olivier Bardou, Sandrine Bouthemy, and Gilles Pagès. Optimal quantization for the pricing of swing options, May 15 2007. Comment: 27p.

[3] Olivier Bardou, Sandrine Bouthemy, and Gilles Pagès. When are swing options bang-bang and how to use it, May 03 2007. Comment: 28 pages.

[4] Mark Broadie, Paul Glasserman, and Steven Kou. A continuity correction for discrete barrier options. *Mathematical Finance*, 7:325–349, 1997.

[5] Sylvain Corlay. Projection rules for vector quantization. 2009.

[6] Pierre Étoré and Benjamin Jourdain. Adaptive optimal allocation in stratified sampling methods. *Methodology and Computing in Applied Probability*, 2008.

[7] E. W. Forgy. Cluster analysis of multivariate data: efficiency vs interpretability of classifications. *Biometrics*, 21:768 769, 1965.

[8] Allen Gersho and Robert M. Gray. *Vector quantization and signal compression.* Kluwer Academic Publishers, Norwell, MA, USA, 1991.

[9] Paul Glasserman. *Monte Carlo Methods in Financial Engineering.* Springer-Verlag New York, Inc., 2004.

[10] Siegfried Graf and Harald Luschgy. *Foundations of Quantization for Probability Distributions.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2000.

[11] Siegfried Graf, Harald Luschgy, and Gilles Pagès. Optimal quantizers for radon random vectors in a banach space. *J. Approx. Theory*, 144(1):27–53, 2007.

[12] Francis Hirsch and Gilles Lacombes. *Elements d'analyse fonctionnelle ; Cours et exercices avec réponses.* 2009.

[13] H. Luschgy and G. Pagès. Functional quantization of Gaussian processes. *Journal of Functional Analysis*, 196(2):486–531, December 2002.

[14] G. Pages and J. Printems. Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods and Appl.*, 11(11):407–446, 2005.

[15] Gilles Pagès and Jacques Printems. `http://www.quantize.maths-fi.com`, 2005. "Web site devoted to optimal quantization".

[16] Gilles Pagès, Jacques Printems, and Huyen Pham. Optimal quantization methods and applications to numerical problems in finance. 2003.

[17] Gilles Pagès Jacques Printems. Optimal quadratic quantization for numerics: the Gaussian case. *Monte Carlo Methods and Applications*, 9:135–166, 2003.