Liquidity costs: a new numerical methodology and an empirical study

Christophe Michel, Victor Reutenauer, Denis Talay, Etienne Tanré

Liquidity costs: a new numerical methodology and an empirical study[*]

Christophe Michel[1], Victor Reutenauer[2],

Denis Talay[3] and Etienne Tanré[3]

December 22, 2015

Keywords: Interest rates derivatives, Optimization, Stochastic Algorithms.

Abstract

We consider rate swaps which pay a fixed rate against a floating rate in presence of bid-ask spread costs. Even for simple models of bid-ask spread costs, there is no explicit strategy optimizing an expected function of the hedging error. We here propose an efficient algorithm based on the stochastic gradient method to compute an approximate optimal strategy without solving a stochastic control problem. We validate our algorithm by numerical experiments. We also develop several variants of the algorithm and discuss their performances in terms of the numerical parameters and the liquidity cost.

1 Introduction

Classical models in financial mathematics usually assume that markets are perfectly liquid. In particular, each trader can buy or sell the amount of assets he/she needs at the same price (the "market price"), and the trader's decisions do not

---

[1]CA-CIB, `christophe.michel@ca-cib.com`

[2]Fotonower (former CA-CIB), `victor@fotonower.com`

[3]Inria, `Denis.Talay@inria.fr`, `Etienne.Tanre@inria.fr`

affect the price of the asset. In practice, the assumption of perfect liquidity is never satisfied but the error due to illiquidity is generally negligible with respect to other sources of error such as model error or calibration error, etc.

However, the perfect liquidity assumption does not hold true in practice for interest rate derivatives market: the liquidity costs to hedge interest rate derivatives are highly time varying. Even though there exist maturities for which zero-coupon bonds are liquid, bonds at intermediate maturities may be extremely illiquid (notice that the underlying interest rate is not directly exchangeable). Therefore, hedging such derivatives absolutely needs to take liquidity risk into account. In this context, defining and computing efficient approximate perfect hedging strategies is a complex problem. The main purpose of this paper is to show that stochastic optimization methods are powerful tools to treat it without solving a necessarily high dimensional stochastic control problem, under the constraints that practitioners need to trade at prescribed dates and that relevant strategies depend on a finite number of parameters. More precisely, we construct and analyze an efficient original numerical method which provides practical strategies facing liquidity costs and minimizing hedging errors.

The outline of the paper is as follows. Section 2 introduces the model. In Section 3, we present our numerical method and analyze it from a theoretical point of view within the framework of a Gaussian yield curve model. Section 4 is devoted to a numerical validation in the idealistic perfect liquidity context. In Section 5, we develop an empirical study of the efficiency of our algorithm in the presence of liquidity costs.

2 Our settings: swaps with liquidity cost

2.1 A short reminder on swaps and swaptions hedging without liquidity cost

One of the most common swaps on the interests rate market is as follows. The counterparts exchange two coupons: the first one is generated by a bond (with a constant fixed interest rate) and the second one is generated by a floating rate (e.g. a LIBOR).

Definition 1. In a perfectly liquid market, the price at time t of a zero-coupon bond paying 1 at time T is denoted by $B(t, T)$. The linear forward rate $L(T_F, T_B, T_E)$ is the fair rate decided at time T_F determining the amount at time T_E obtained

by investing 1 at time T_B. The following relation is satisfied:

$$L\left(T_F, T_B, T_E\right) = \frac{1}{T_E - T_B}\left(\frac{B\left(T_F, T_B\right)}{B\left(T_F, T_E\right)} - 1\right). \tag{1}$$

A swap contract specifies:

- an agreement date t

- a time line $(t \leq) T_0 < \cdots < T_N$

- a fixed interest rate r

- a floating interest rate

- the payoff at each time T_i $(1 \leq i \leq N)$, that is,

$$P(i) := (T_i - T_{i-1})\left(r - L\left(T_{i-1}, T_{i-1}, T_i\right)\right). \tag{2}$$

From (1), we deduce the equivalent expression

$$P(i) = r\left(T_i - T_{i-1}\right) - \frac{1}{B(T_{i-1}, T_i)} + 1. \tag{3}$$

In the sequel, we consider that the fixed rate r is chosen at the money (thus the swap at time t has zero value), and that the swap fixed coupons are received by the trader.

In the idealistic framework of a market without liquidity cost, the trader buys or sells quantities of zero-coupon bonds at the same price (i.e. the *market price*), and there exists a discrete time perfect hedging strategy which is independent of any model of interest rates. In view of (3), the replication of the payoff $P(i)$ at time T_i can be split into three parts:

- the fixed part $r(T_i - T_{i-1})$ is replicated statically at time t by selling $r(T_i - T_{i-1})$ zero-coupon bonds with maturity T_i.

- the floating part $1/B(T_{i-1}, T_i)$ is replicated dynamically at time T_{i-1} by buying $1/B(T_{i-1}, T_i)$ zero-coupon bonds with maturity T_i. The price of this transaction is equal to 1.

- the last (fixed) part 1 is used at time T_i to buy $1/B(T_i, T_{i+1})$ zero-coupon bonds with maturity T_{i+1}.

It is easy to see that this strategy is self-financing at times $T_1, \cdots, T_{N-1}$. To make it self-financing at time t also, at this date one buys 1 zero-coupon bond with maturity T_0 and sells 1 zero-coupon bond with maturity T_N.

To summarize, in the idealistic framework, we do not need to consider $(\mathcal{F}_\theta, \theta \geq 0)$ adapted hedging strategies, where $(\mathcal{F}_\theta, \theta \geq 0)$ is the filtration generated by the observations (short rates, derivative prices, etc.) in continuous time, and we may restrict the admissible strategies to be adapted to the filtration generated by the observations at times $\{t, T_0, T_1, \cdots, T_N\}$.

2.2 Hypotheses on markets with liquidity costs

We now consider markets with liquidity costs and need to precise our liquidity cost model. In all the sequel T_{-1} denotes t.

Hypothesis 2.
We assume that, for all $-1 \leq j < i \leq N$, the number $\pi(j, i)$ of zero-coupon bonds with maturity T_i bought or sold at time T_j is measurable with respect to the filtration generated by $(R_t, R_{T_0}, \cdots, R_{T_j})$. That means that the admissible strategies do not depend on the evolution of the rate R_θ between two tenor dates T_m and T_{m+1}.

Denote by $\Psi(T, U, \pi)$ the buy or sell price for π zero coupon bonds. In perfectly liquid markets, $\Psi(T, U, \pi)$ is the linear function $B(T, U)\pi$, where $B(T, U)$ is defined in Definition 1. In the presence of liquidity costs, $\Psi(T, U, \pi)$ becomes a non-linear function of π.

Hypothesis 3.
For all T and U, the price $\Psi(T, U, \pi)$ is a $C^1(\mathbb{R})$, increasing, convex one-to-one map of π from $\mathbb{R}$ to $\mathbb{R}$, and $\Psi(T, U, 0) = 0$.

Under the preceding hypothesis, the function Ψ is positive when $\pi > 0$ and negative when $\pi < 0$.

In the context of the swap we set

$$\Psi_{i,j}(\pi) := \Psi(T_i, T_j, \pi) \tag{4}$$

and we only consider self-financing strategies, that is, satisfying

$$\forall\, 0 \leq j \leq N - 1, \quad \sum_{-1 \leq k < j} \pi(k, j) + P(j) = \sum_{j < i \leq N} \Psi_{j,i}(\pi(j, i)). \tag{5}$$

4

2.3   Optimization objective

In the presence of liquidity risk, the market is not complete any more and the practitioners need to build a strategy which minimizes a given function S (e.g. a risk measure) of the hedging error. Such strategies are usually obtained by solving stochastic control problems. These problems require high complexity numerical algorithms which are too slow to be used in practice. We here propose an efficient and original numerical method to compute approximate optimal strategies. As the perfect hedging leads to a null portfolio at time T_N, we have to solve the optimization problem

$$\inf_{\boldsymbol{\pi} \in \Pi} \mathbb{E}\left[S(W^{\boldsymbol{\pi}})\right], \tag{6}$$

where $W^{\boldsymbol{\pi}}$ is the terminal wealth (at time T_N) given the strategy $\boldsymbol{\pi}$ in the set Π of admissible strategies.

3   Hedging error minimization method in a Gaussian framework

The methodology we introduce in this section is based on the two following key observations:

(1) We consider strategies and portfolios with finite second moment, and thus optimize within $L^2(\mu)$ for some probability measure μ. The Gram-Schmidt procedure provides countable orthogonal bases $\mathcal{B}$ of the separable Hilbert space $L^2(\mu)$. Our set Π of admissible strategies is obtained by truncating of a given basis, which reduces the a priori infinite dimensional optimization problem (6) to a finite dimensional parametric optimization problem of the type $\inf_{\theta \in \Theta} \mathbb{E}\Psi(\theta, X)$, where Θ is a subset of $\mathbb{R}^p$, X is a given random variable, Ψ is a convex function of θ.

(2) The Robbins-Monro algorithm and its Chen extension are stochastic alternatives to Newton's method to numerically solve such optimization problems. These algorithms do not require to compute $\frac{d}{d\theta}\mathbb{E}\Psi(\theta, X)$. They are based on sequences of the type

$$\theta_{\gamma+1} = \theta_\gamma - \rho_{\gamma+1}\frac{\partial}{\partial \theta}\Psi(\theta_\gamma, X_{\gamma+1}), \tag{7}$$

where $(\rho_\gamma, \gamma \geq 1)$ is a decreasing sequence and $(X_\gamma, \gamma \geq 1)$ is an i.i.d. sequence of random variables distributed as X.

We here consider the case of swap in the context of a Gaussian yield curve. This assumption is restrictive from a mathematical point of view but is satisfied by widely used interest rate models such as Vasicek model, Gaussian affine models or HJM (Heath, Jarrow and Morton) models with deterministic volatilities (see e.g. Gibson *et al.* (2010) and Musiela and Rutkowski (2005)). In Babbs and Nowman (1999) it is shown that using a three dimensional Gaussian model is sufficient to fit the term structure of interest rate products.

3.1 Step 1: finite dimensional projections of the admissible controls space

Consider a Gaussian short rate model $(R_\theta, \theta \geq 0)$: we either suppose that the dynamics of the short rate model is given, or that it is deduced from a forward rate model such as in the HJM approach for term structures: see e.g. Eq (6.9) in Gibson *et al.* (2010) under the additionnal assumption that the forward rate volatilities are deterministic. In all cases, the resulting bond price model is log-normal.

In view of Hypothesis 2, each control $\pi(j, i)$ belongs to the Gaussian space generated by $(R_t, R_{T_0}, \cdots, R_{T_j})$ or, equivalently, to a space generated by $\ell(j)+1$ standard independent Gaussian random variables $G^{(0)}, G^{(1)}, \cdots, G^{(\ell(j))}$ (with $\ell(j) = j$ for one-factor models, $\ell(j) = 2j + 1$ for two-factor models, etc.) and R_t. An explicit L^2 orthonormal basis of the space generated by $G^{(0)}, G^{(1)}, \cdots, G^{(\ell(j))}$ is

$$\left(\prod_{m=0}^{\ell(j)} \frac{H_{n_m}(G^{(m)})}{\sqrt{n_m!}} \right)_{(n_0, \cdots, n_{\ell(j)}) \in \mathbb{N}^{\ell(j)+1}}, \tag{8}$$

where $(H_n, n \geq 0)$ are the Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}$$

(see e.g. Malliavin, 1995, p.236).

Thus, the quantities of zero-coupon bonds bought by the trader can be written as

$$\sum_{\boldsymbol{n}_{(j)} \in \mathbb{N}^{\ell(j)+1}} \alpha^{\boldsymbol{n}_{(j)}}(j, i) \prod_{m=0}^{\ell(j)} \frac{H_{n_m}(G^{(m)})}{\sqrt{n_m!}}, \quad \boldsymbol{n}_{(j)} = (n_0, \cdots, n_{\ell(j)}), \tag{9}$$

where the infinite sum has an L^2 limit sense. A strategy can now be defined as a sequence of real numbers $\alpha^{\boldsymbol{n}_{(j)}}(j, i)$ for all $-1 \leq j < i \leq N - 1$ and $\boldsymbol{n}_{(j)} \in \mathbb{N}^{\ell(j)+1}$.

In order to be in a position to solve a finite dimensional optimization problem, we truncate the sequence $(\alpha^{\boldsymbol{n}_{(j)}}(j,i))$. Then a strategy is defined by a finite number of real parameters $\{\alpha^{\boldsymbol{n}_{(j)}}(j,i), -1 \leq j < i \leq N-1\}$ where, for all j, $\boldsymbol{n}_{(j)}$ belongs to a finite subset $\Lambda^{(j)}$ of $\mathbb{N}^{\ell(j)+1}$. The truncated quantities of zero-coupon bonds bought by the trader write

$$\pi(j,i) = \sum_{\boldsymbol{n}_{(j)} \in \Lambda^{(j)}} \alpha^{\boldsymbol{n}_{(j)}}(j,i) \prod_{m=0}^{\ell(j)} \frac{H_{n_m}(G^{(m)})}{\sqrt{n_m!}}. \tag{10}$$

We discuss the efficiency of this truncation and its convergence in Sections 3.5 and 4.1.

To simplify, we denote by $\boldsymbol{\alpha} = (\alpha^{\boldsymbol{n}_{(j)}}(j,i))_{i,j,\boldsymbol{n}_{(j)}}$ the parameters to optimize in $\mathbb{R}^p$ (where the dimension p is known for each truncation $(\Lambda^{(j)}, j = -1, 0, \cdots, N-1)$), by $\boldsymbol{\pi}(\boldsymbol{\alpha})$ (or, when no confusion is possible, simply $\boldsymbol{\pi}$) the hedging strategy corresponding to a vector $\boldsymbol{\alpha} \in \mathbb{R}^p$, see (10). Given the strategy $\boldsymbol{\pi} = \boldsymbol{\pi}(\boldsymbol{\alpha})$, the terminal wealth $W^{(\boldsymbol{\alpha})}$ (at time T_N) satisfies

$$W^{(\boldsymbol{\alpha})} = \sum_{-1 \leq j < N} \pi(\boldsymbol{\alpha})(j,N) + P(N) = \sum_{-1 \leq j < N} \pi(j,N) + P(N). \tag{11}$$

The problem (6) is now formulated as: find $\boldsymbol{\alpha}^*$ in $\mathbb{R}^p$ such that

$$\mathbb{E}\left[S\left(W^{(\boldsymbol{\alpha}^*)}\right)\right] = \inf_{\boldsymbol{\alpha} \in \mathbb{R}^p} \mathbb{E}\left[S\left(W^{(\boldsymbol{\alpha})}\right)\right]. \tag{12}$$

3.2 Step 2: stochastic optimization

Using the self-financing equation (5) one can express $W^{(\boldsymbol{\alpha})}$ as a function of $\boldsymbol{\alpha}$ and $(G^{(0)}, \cdots, G^{(\ell(N))})$. Therefore one needs to minimize the expectation of a deterministic function of the parameter $\boldsymbol{\alpha}$ in $\mathbb{R}^p$ and the random vector $(G^{(0)}, \cdots, G^{(\ell(N))})$. Such problems can be solved numerically by classical stochastic optimization algorithms, such as those introduced in the pioneering work of Robbins and Monro (1951) and its extensions (e.g. Chen and Zhu, 1986). We refer the interested reader to the classical references Duflo (1997); Chen (2002); Kushner and Yin (2003).

In our context (12), the Robbins-Monro algorithm (7) works as follows. Start with an arbitrary initial condition $\boldsymbol{\alpha}_0$ in $\mathbb{R}^p$. At step $\gamma + 1$, given the current approximation $\boldsymbol{\alpha}_\gamma$ of the optimal value $\boldsymbol{\alpha}^*$, simulate independent Gaussian random variables $(G^{(0)}_{\gamma+1}, \cdots, G^{(\ell(N))}_{\gamma+1})$ and compute the terminal wealth $W^{(\boldsymbol{\alpha}_\gamma)}$. Then, update the parameter $\boldsymbol{\alpha}$ by the induction formula

$$\boldsymbol{\alpha}_{\gamma+1} = \boldsymbol{\alpha}_\gamma - \rho_{\gamma+1} \nabla_{\boldsymbol{\alpha}}\left[S(W^{(\boldsymbol{\alpha}_\gamma)})\right], \tag{13}$$

where (ρ_γ) is a deterministic decreasing sequence. In addition, one can use an improvement of this algorithm due to Chen and Zhu (1986). Let $(K_l, l \geq 0)$ be an increasing sequence of compact sets such that

$$K_l \subset \mathrm{Int}(K_{l+1}) \quad \text{and} \quad \lim_l K_l = \mathbb{R}^p, \tag{14}$$

where $\mathrm{Int}(K_{l+1})$ denotes the interior of the set K_{l+1}. The initial condition $\boldsymbol{\alpha}_0$ is assumed to be in K_0 and we set $l(0) = 0$. At each step γ in (13), if $\boldsymbol{\alpha}_{\gamma+1} \in K_{l(\gamma)}$, we set $l(\gamma+1) = l(\gamma)$ and go to step $\gamma+1$. Otherwise, that is if $\boldsymbol{\alpha}_{\gamma+1} \notin K_{l(\gamma)}$, we set $\boldsymbol{\alpha}_{\gamma+1} = \boldsymbol{\alpha}_0$ and $l(\gamma+1) = l(\gamma)+1$. This modification avoids that the stochastic algorithm may blow up during the first steps and, from a theoretical point of view, allows to prove its convergence under weaker assumptions than required for the standard Robbins-Monro method.

3.3 Summary of the method

Our setting

- The interest rate model satisfies: for all $0 \leq j < N$, there exist an integer $\ell(j)$ and a function Φ_j such that

$$(R_t, R_{T_0}, \cdots, R_{T_j}) \overset{\mathcal{L}}{=} \Phi_j(R_t, G^{(0)}, \cdots, G^{(\ell(j))}).$$

- For all $0 \leq j < N$, a finite truncation set $\Lambda^{(j)} \subset \mathbb{N}^{\ell(j)+1}$ is given and $\Lambda^{(-1)} = \{0\}$.

- A strategy $\boldsymbol{\pi} = \boldsymbol{\pi}(\boldsymbol{\alpha})$ is defined by $\boldsymbol{\alpha} = (\alpha^{\boldsymbol{n}_{(j)}}(j,i), -1 \leq j < i \leq N-1, \boldsymbol{n}_{(j)} \in \Lambda^{(j)})$. More precisely, the number of zero-coupon bonds with maturity T_i bought or sold at time T_j is

$$\pi(j,i) = \sum_{\boldsymbol{n}_{(j)} \in \Lambda^{(j)}} \alpha^{\boldsymbol{n}_{(j)}}(j,i) \prod_{m=0}^{\ell(j)} \frac{H_{n_m}(G^{(m)})}{\sqrt{n_m!}}, \quad \text{for } i \leq N-1, \tag{15}$$

and $\pi(j,N)$ is deduced from the self-financing equation

$$\sum_{i<j} \pi(i,j) + P(j) = \sum_{i>j} \Psi_{j,i}(\pi(j,i)) \tag{16}$$

(one possibly needs to use a classical iterative procedure to solve this equation numerically).

- One is given an increasing sequence of compact sets $(K_l, l \geq 0)$ satisfying (14) and a sequence of parameters $(\rho_\gamma, \gamma \geq 1)$ decreasing to 0.

Our stochastic optimization algorithm

Assume that the parameter $\boldsymbol{\alpha}_\gamma = (\alpha_\gamma^{\boldsymbol{n}_{(j)}}(j,i), -1 \leq j < i \leq N-1, \boldsymbol{n}_{(j)} \in \Lambda^{(j)})$ and l_γ are given at step γ. At step $\gamma + 1$:

1. Simulate a Gaussian vector $(G_{\gamma+1}^{(0)}, \cdots, G_{\gamma+1}^{(\ell(N))})$.

2. Deduce the quantities of zero-coupon bonds from (15) and (16):

$$\pi_{\gamma+1}(j,i) = \sum_{\boldsymbol{n}_{(j)} \in \Lambda^{(j)}} \alpha_\gamma^{\boldsymbol{n}_{(j)}}(j,i) \prod_{m=0}^{\ell(j)} \frac{H_{n_m}(G_{\gamma+1}^{(m)})}{\sqrt{n_m!}}, \quad \text{for } i \leq N-1,$$

   and get $\pi_{\gamma+1}(j,N)$ from the self-financing equation

$$\sum_{i<j} \pi_{\gamma+1}(j,i) + P(j) = \sum_{i>j} \Psi_{j,i}(\pi_{\gamma+1}(j,i)).$$

3. Compute the terminal wealth

$$W_{\gamma+1}^{(\boldsymbol{\alpha}_\gamma)} = \sum_{j<N} \pi_{\gamma+1}(j,N) + P(N).$$

4. Update the parameters

$$\boldsymbol{\alpha}_{\gamma+1} = \boldsymbol{\alpha}_\gamma - \rho_{\gamma+1} \nabla_{\boldsymbol{\alpha}} \left[S(W_{\gamma+1}^{(\boldsymbol{\alpha}_\gamma)}) \right].$$

5. If $\boldsymbol{\alpha}_{\gamma+1} \notin K_{l_\gamma}$, set $\boldsymbol{\alpha}_{\gamma+1} = \boldsymbol{\alpha}_0$ and $l(\gamma+1) = l(\gamma) + 1$.

6. Go to 1 (or, in practice, stop after Γ steps).

3.4 Error analysis

In this subsection we study the convergence (Theorem 4) and convergence rate (Theorem 6) of the stochastic algorithm used in Step 2, when the total number of steps Γ tends to infinity. We introduce some notation. Recall (13) and write

$$\boldsymbol{\alpha}_{\gamma+1} = \boldsymbol{\alpha}_\gamma - \rho_{\gamma+1}\mathbb{E}\left[\nabla_{\boldsymbol{\alpha}} S(W_{\gamma+1}^{(\boldsymbol{\alpha}_\gamma)})\right] - \rho_{\gamma+1}\delta M_{\gamma+1} + \rho_{\gamma+1}p_{\gamma+1}. \tag{17}$$

Here, $\delta M_{\gamma+1}$ is given by

$$\delta M_{\gamma+1} = \nabla_{\boldsymbol{\alpha}} S(W_{\gamma+1}^{(\boldsymbol{\alpha}_\gamma)}) - \mathbb{E}\left[\nabla_{\boldsymbol{\alpha}} S(W_{\gamma+1}^{(\boldsymbol{\alpha}_\gamma)})\right]. \tag{18}$$

The last term $\rho_{\gamma+1}p_{\gamma+1}$ in (17) represents the *reinitialization* of the algorithm if $\boldsymbol{\alpha}_{\gamma+1} \notin K_{l(\gamma)}$ i.e. $p_{\gamma+1}$ is fixed such that $\boldsymbol{\alpha}_{\gamma+1} = \boldsymbol{\alpha}_0$.

Let us now recall the convergence theorem obtained by Lelong (2008, Theorem 1) in our setting.

Theorem 4. *Assume*

(A1) The function $\boldsymbol{\alpha} \mapsto \mathbb{E}[S(W^{(\boldsymbol{\alpha})})]$ is strictly concave or convex,

(A2) $\sum_{\gamma} \rho_{\gamma} = \infty$, $\sum_{\gamma} \rho_{\gamma}^2 < \infty$,

(A3) The function $\boldsymbol{\alpha} \mapsto \mathbb{E}[\|\nabla_{\boldsymbol{\alpha}} S(W^{(\boldsymbol{\alpha})})\|^2]$ is bounded on compact sets.

Then the sequence $(\boldsymbol{\alpha}_{\gamma}, \gamma \geq 1)$ converges a.s. to the unique
optimal parameter $\boldsymbol{\alpha}^*$ such that

$$\inf_{\boldsymbol{\alpha} \in \mathbb{R}^p} \mathbb{E}\left[S\left(W^{(\boldsymbol{\alpha})}\right)\right] = \mathbb{E}\left[S\left(W^{(\boldsymbol{\alpha}^*)}\right)\right].$$

Hypothesis 3 and (15) imply that (A3) is satisfied. Before giving examples of situations where (A1) is fulfilled, let us check that $W^{(\boldsymbol{\alpha})}$ is a concave function of $\boldsymbol{\alpha}$.

Proposition 5. *The terminal wealth $W^{(\boldsymbol{\alpha})}$ is a concave function of the parameter $\boldsymbol{\alpha}$.*

Proof. Recall that the terminal wealth is given by (11). The payoff of the swap $P(N)$ does not depend on $\boldsymbol{\alpha}$. We only have to deal with the quantities $\pi(i, N)$ of the zero-coupon bonds with maturity T_N bought at time T_i. They satisfy the self-financing equation (5) and thus

$$\pi(i, N) = \Psi_{i,N}^{(-1)} \left(\sum_{j<i} \pi(j, i) - \sum_{i<j<N} \Psi_{i,j} \left(\pi(i, j) \right) + P(i) \right), \qquad (19)$$

where $\Psi_{i,N}^{(-1)}$ is the inverse of the price function $\Psi_{i,N}$ (see (4)). Moreover, quantities $\pi(i, j)$, $i < j \leq N - 1$ are linear in $\boldsymbol{\alpha}$ (see (9)).

Recall that $\Psi_{i,j}$ is convex, thus $-\Psi_{i,j}$ is concave and the argument in (19) is a concave function of $\boldsymbol{\alpha}$. Finally $\Psi_{i,N}^{(-1)}$ is an increasing concave function, from which $\pi(i, N)$ is a concave function of $\boldsymbol{\alpha}$. $\qquad \square$

The preceding observation shows that (A1) is satisfied when S is a utility function (and thus increasing and concave) and satisfies $S(0) = 0$. Notice that the optimization problem (12) then penalizes the losses and promotes the gains. In Sections 4 and 5 we will see another situation where Theorem 4 applies.

Given suitable functions S, Theorem 4 guarantees the convergence of our algorithm towards the optimal parameters. The following theorem provides the rate of convergence (Lelong, 2013).

Theorem 6. *Let*

$$\rho_\gamma := \frac{v_1}{(v_2 + \gamma)^\beta}, \tag{20}$$

for some positive v_1, v_2 and $\beta \in (1/2, 1)$. Denote by Δ_γ the normalized centered error

$$\Delta_\gamma = \frac{\boldsymbol{\alpha}_\gamma - \boldsymbol{\alpha}^*}{\sqrt{\rho_\gamma}}.$$

Assume

(A1) *The function $\boldsymbol{\alpha} \mapsto \mathbb{E}[S(W^{(\boldsymbol{\alpha})})]$ is concave or convex.*

(A4) *For any $q > 0$, the series*

$$\sum_\gamma \rho_{\gamma+1} \delta M_{\gamma+1} \mathbb{1}_{\{|\boldsymbol{\alpha}_\gamma - \boldsymbol{\alpha}^*| \leq q\}}$$

converges almost surely.

(A5) *There exist two real numbers $A_1 > 0$ and $A_2 > 0$ such that*

$$\sup_\gamma \mathbb{E}\left[|\delta M_\gamma|^{2+A_1} \mathbb{1}_{\{|\boldsymbol{\alpha}_\gamma - \boldsymbol{\alpha}^*| \leq A_2\}} \right] < \infty.$$

(A6) *There exists a symmetric positive definite matrix Σ such that*

$$\mathbb{E}\left[\delta M_\gamma \delta M_\gamma^{\mathsf{t}} \big| \mathcal{F}_{\gamma-1} \right] \mathbb{1}_{\{|\boldsymbol{\alpha}_{\gamma-1} - \boldsymbol{\alpha}^*| \leq A_2\}} \xrightarrow[\gamma \to \infty]{\mathbb{P}} \Sigma.$$

(A7) *There exists $\mu > 0$ such that $\forall n \geq 0$, $d(\boldsymbol{\alpha}^*, \partial K_n) \geq \mu$,*

where ∂K_n denotes the boundary of K_n. Then, the sequence $(\Delta_\gamma, \gamma \geq 1)$ converges in distribution to a normal random variable with mean 0 and covariance matrix linearly depending on Σ.

Remark 7. As explained in detail in Lelong (2013, Sec. 2.4), the assumptions of Theorem 6 are satisfied as soon as

- There exists $A_3 > 0$ such that

$$\forall C > 0, \quad \mathbb{E}\left[\sup_{|\boldsymbol{\alpha}| \leq C} \left| S\left(W^{(\boldsymbol{\alpha})}\right) \right|^{2+A_3}\right] < \infty.$$

- The function $\boldsymbol{\alpha} \mapsto \mathbb{E}[S(W^{(\boldsymbol{\alpha})})]$ is strictly concave or convex.

The first condition is usually easily satisfied, e.g. when S has a polynomial growth at infinity because the moments of the wealth process are typically finite. The both properties are e.g. fulfilled by the example studied in Sections 4 and 5 (see the discussions at the beginning of these sections).

3.5 Performance of the optimal truncated strategy without liquidity cost

The numerical error on the optimal wealth decreases when the $\Lambda^{(j)}$'s tend to $\mathbb{N}^{\ell(j)+1}$. In this subsection, we provide a theoretical estimate on the error resulting from the truncation in (9) in the idealistic context of no liquidity cost and general Gaussian affine models (Dai and Singleton, 2000).

In Dai and Singleton (2000), general Gaussian affine models are introduced for which, for any two times $s < t$, there exist standard independent Gaussian random variables $G^{(0)}, \cdots, G^{(M)}$ and real numbers $\mu, \lambda_0, \cdots, \lambda_M$ such that the prices of zero-coupon bonds have the form

$$B(s,t) = \exp\left(-\mu - \lambda_0 G^{(0)} - \cdots - \lambda_M G^{(M)}\right).$$

A control of the error of truncation is given in the following proposition.

Proposition 8. *In the above context, if the truncation set defined in* (10) *is* $\Lambda^{(j)} := \{n_0 + \cdots + n_{\ell(j)} \leq d\}$, *then*

$$\mathbb{E}\left(W^{(\boldsymbol{\alpha}^*)}\right)^2 \leqslant C_0 \frac{C_1^{d+1}}{(d+1)!}, \tag{21}$$

where C_0 and C_1 are some positive constants.

The proposition is a straightforward consequence of (11) and the next lemma applied to $X = \boldsymbol{\pi}(\boldsymbol{\alpha}^*)$. This lemma also allows one to precise the values of C_0 and C_1.

Lemma 9. *Consider the random variable*

$$X := \exp\left(\mu + \sum_{m=0}^{M} \lambda_m G^{(m)}\right),$$

where $\mu, \lambda_0, \cdots, \lambda_M$ are real numbers, and $G^{(0)}, \cdots, G^{(M)}$ are independent standard Gaussian random variables. Consider the projection X^d of X on the subspace of $L^2(G^{(0)}, \cdots, G^{(M)})$ generated by

$$\left(\prod_{m=0}^{M} \frac{H_{n_m}(G^{(m)})}{\sqrt{n_m!}}, \quad n_0 + \cdots + n_M \leq d\right).$$

We have

$$\|X - X^d\|_2 \leqslant \exp\left(\mu + \lambda_0^2 + \cdots \lambda_M^2\right) \frac{(\lambda_0^2 + \cdots + \lambda_M^2)^{\frac{d+1}{2}}}{\sqrt{(d+1)!}}. \tag{22}$$

We postpone the proof of this lemma to the Appendix.

4 Numerical validation of the optimization procedure: an example without liquidity cost

In this section we study the accuracy of our algorithm in the no liquidity cost case where a perfect replication strategy is known (see Section 2.1). We minimize the quadratic risk measure of the hedging error.

The bond market model is the Vasicek model which is the simplest Gaussian model:

$$dR_\theta = A(r_\infty - R_\theta)d\theta + \sigma dB_\theta, \tag{23}$$

where A is the mean reverting rate, r_∞ is the mean of the equilibrium measure, σ is the volatility and $(B_\theta, \theta \geq 0)$ is a one-dimensional Brownian motion.

Notice that

$$\forall u < v, \quad R_v = r_\infty + (R_u - r_\infty)e^{-A(v-u)} + \sigma e^{-A(v-u)} \int_u^v e^{A(\theta-u)}dB_\theta. \tag{24}$$

Therefore, there exists an *i.i.d.* sequence $(G^{(0)}, G^{(1)}, \cdots, G^{(N)})$ of $\mathcal{N}(0,1)$ Gaussian random variables such that

$$\forall k = 0, \cdots, N, \quad R_{T_k} = \Phi(T_k - T_{k-1}, R_{T_{k-1}}, G^{(k)}), \tag{25}$$

13

where

$$\Phi(\eta, r, G) = r_\infty + e^{-A\eta}(r - r_\infty) + G\sigma\sqrt{\frac{1 - e^{-2A\eta}}{2A}}.$$

In our numerical experiments, we have chosen the following typical values of the parameters $A = 10\%$, $r_\infty = 5\%$, $\sigma = 5\%$. With this choice of parameters, the mean yearly interest zero-coupon rates with maturity less than 10 years take values between 3% and 5%.

Our numerical study concerns the minimization of the quadratic mean hedging error which corresponds to the choice $S(x) = x^2$ in (12). This choice penalizes gains and losses in a symmetric way and aims to construct a strategy as close as possible to the exact replication strategy.

In the no liquidity cost case, the terminal wealth $W^{(\boldsymbol{\alpha})}$ is a linear function of the parameter $\boldsymbol{\alpha}$ and therefore assumption (A1) of Theorem 4 is obviously satisfied.

Given a degree of truncation d, the set $\Lambda_d^{(j)}$ is chosen as

$$\Lambda_d^{(j)} := \{\boldsymbol{n}_{(j)} = (n_0, \cdots, n_j) \in \mathbb{N}^{j+1}, n_0 + \cdots + n_j \leq d\}. \tag{26}$$

We have to optimize the real-valued parameters $\alpha^{\boldsymbol{n}_{(j)}}(j, i)$ for $j < i$ and $\boldsymbol{n}_{(j)} \in \Lambda_d^{(j)}$. The quantities of zero-coupon bonds to exchange are given by (10). The choice of the sequence $(\rho_\gamma, \gamma \geq 1)$ in (13) is crucial. Choose ρ_γ as in (20). We discuss the sensitivity of the method to the parameters v_1, v_2, β in Section 4.2.2. We also discuss the sensitivity of the results to the number Γ of steps.

In all the sequel, we use the following notation.

Notation For all vector $\boldsymbol{\alpha} = (\alpha^{\boldsymbol{n}_{(j)}}(j, i), -1 \leq j < i \leq N - 1)$, we set

$$v(\boldsymbol{\alpha}) := \mathbb{E}\left[W^{(\boldsymbol{\alpha})}\right]^2, \tag{27}$$

where the expectation is computed only with respect to the Gaussian distribution $(G^{(0)}, \cdots, G^{(\ell(N))})$.

4.1 Empirical study of the truncation errors (Step 1)

In this subsection we develop an empirical validation of the projection step presented in Section 3.1

We observe that the quadratic mean hedging error decreases very fast to 0 when the degree of truncation increases. For a notional equal to 1, the error is of the order of one basis point (a hundredth of percent) for a degree $d = 3$ and a small number of dates N, and for $d = 4$ and for larger values of N.

Figure 1 below shows $v(\boldsymbol{\alpha}^{*,d})$, with the optimal parameter $\boldsymbol{\alpha}^{*,d}$ corresponding to the truncation set (26). We have used the explicitly known finite dimensional projections of the optimal strategies without liquidity cost to obtain $\boldsymbol{\alpha}^{*,d}$, and a Monte Carlo procedure to compute v. Table 1 shows some values used to plot Figure 1.
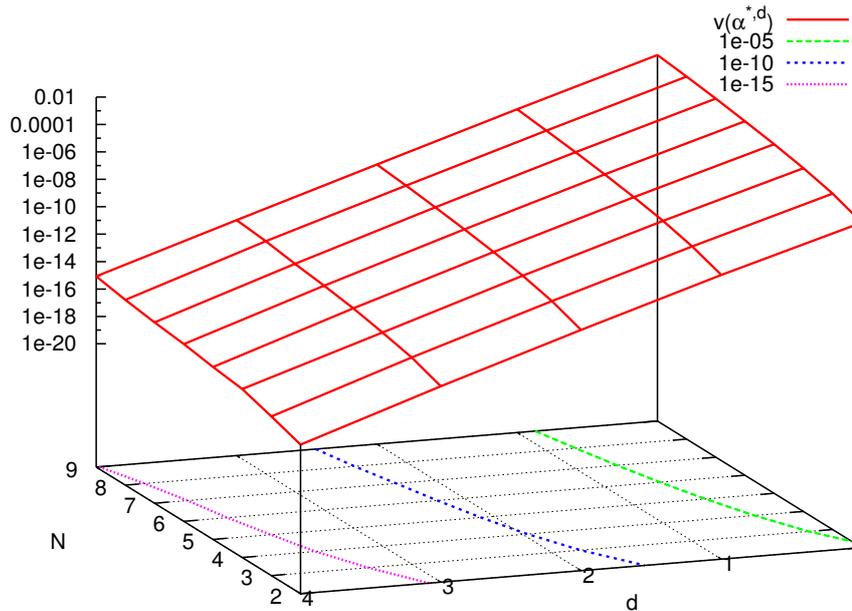

Figure 1

4.2 Empirical study of the optimization step (Step 2)

The stochastic algorithm converges almost surely to the optimal coefficient $\boldsymbol{\alpha}^*$. In this part, we empirically study the convergence rate in terms of the number of steps Γ and the choice of the sequence (ρ_γ).

15

degree d	N=2	N=3
0	5.2 E-6	3.0 E-5
1	5.4 E-9	3.1 E-8
2	3.7 E-12	2.0 E-11
3	1.9 E-15	1.9 E-14
4	2.2 E-18	3.9 E-15

Table 1: $v(\boldsymbol{\alpha}^{*,d})$

4.2.1 A typical evolution of (α_γ)

In this subsection, we consider a swap with two payment dates $(N = 2)$. We consider the truncation set $\Lambda^{(0)} = \{0, 1\}$. The objective is to approximate $\boldsymbol{\alpha}^* = (\alpha^{0,*}(-1, 0), \alpha^{0,*}(-1, 1), \alpha^{0,*}(0, 1), \alpha^{1,*}(0, 1))$. In Figure 2, the four parameters $\boldsymbol{\alpha} = (\alpha^0(-1, 0), \alpha^0(-1, 1), \alpha^0(0, 1), \alpha^1(0, 1))$ evolve according to (13) where the sequence $(\rho_\gamma, \gamma \geq 1)$ is defined by (20) with $v_1 = 10^7$, $v_2 = 1$ and $\beta = 1$.

As expected, the sequence $(\boldsymbol{\alpha}_\gamma)$ converges to $\boldsymbol{\alpha}^*$. However, the evolution is quite *slow* although we have empirically chosen the parameters v_1, v_2 and β in a favorable way.

In Figure 3, we plot (in violet) $v(\alpha^0(-1, 0), \alpha^0(-1, 1), \alpha^{0,*}(0, 1), \alpha^{1,*}(0, 1))$ as a function of $\alpha^0(-1, 0)$ and $\alpha^0(-1, 1)$. We also plot in green the path $(v(\boldsymbol{\alpha}_\gamma), 0 \leq \gamma \leq \Gamma)$. The figure shows that after $\Gamma = 10000$ steps the hedging error is small though the optimal parameters have not been approximated accurately (notice that the violet surface is flat).
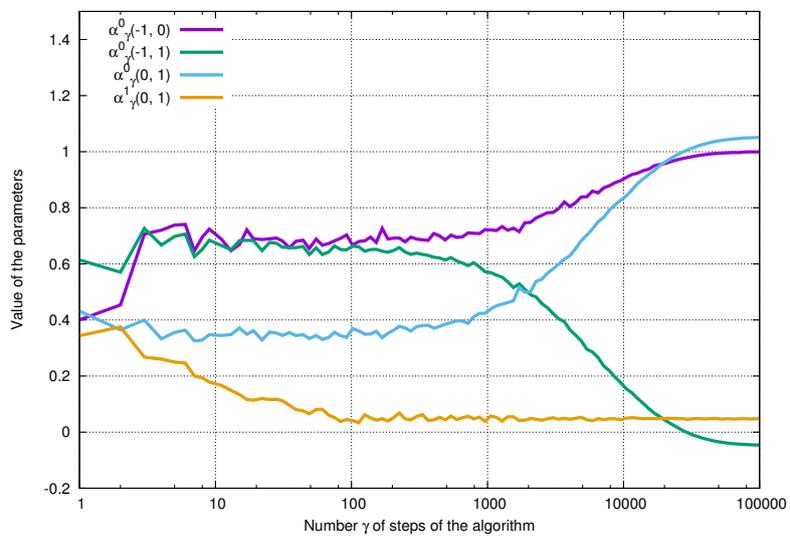
Figure 2: Evolution of the parameters $\alpha_\gamma^0(-1, 0)$, $\alpha_\gamma^0(-1, 1)$, $\alpha_\gamma^0(0, 1)$, $\alpha_\gamma^1(0, 1)$ in terms of γ.

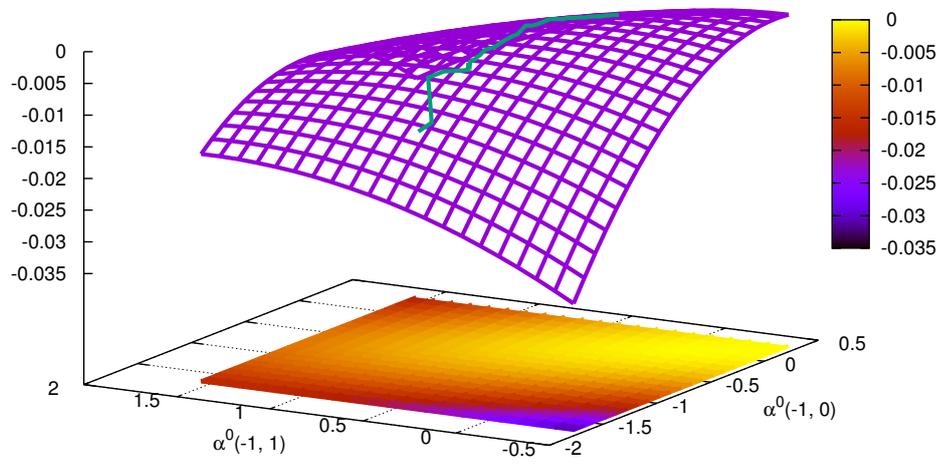Figure 3: Violet surface: $-v(\bullet, \bullet, \alpha^{0,*}(0,1), \alpha^{1,*}(0,1))$. Green curve: evolution of $(\alpha^0\gamma(-1,0), \alpha^0_\gamma(-1,1), -v(\alpha^0_\gamma(-1,0), \alpha^0_\gamma(-1,1), \alpha^{0,*}(0,1), \alpha^{1,*}(0,1))$ in terms of $\gamma = 1, \cdots, 10000$ ($v_1 = 10$, $v_2 = 1$ and $\beta = 0.6$).

4.2.2 Sensitivity to the choice of the sequence (ρ_γ)

Theorem 4 states the convergence of the optimization method for all sequence (ρ_γ) satisfying (A2). We here study the sensitivity of the results to the parameters v_1, v_2, β of sequences of type (20) and to the total number of steps Γ.

Tables 2 and 3 show the expected value function obtained after $\Gamma = 10E4$, $10E5$ and $10E6$ steps. The expected value function is estimated by means of a classical Monte Carlo procedure. Table 4 shows the same results with a sequence $\rho_\gamma = v_1$ which does not satisfy condition (A2).

v_1 / Γ	1	10	100	1000	10 000	20 000	10E5	10E6
10E4	7.2 E-4	2.6 E-6	1.1 E-6	8.9 E-7	5.1 E-4	8.3 E-4	1.0 E-1	1.3 E-1
10E5	1.4 E-5	9.3 E-7	8.9 E-7	4.7 E-7	2.3 E-8	1.7 E-2	9.1 E+3	9.5 E+5
10E6	3.8 E-6	9.2 E-7	7.7 E-7	1.5 E-7	6.6 E-12	5.4 E-12	1.0 E-5	15.8

Table 2: $v(\boldsymbol{\alpha}_\Gamma)$ ($v_2 = 1000$, $\beta = \mathbf{0.6}$)

v_1 / Γ	1	10	100	1000	10 000	13 000	2E4
10E4	8.8 E-3	1.0 E-3	7.8 E-6	9.8 E-7	8.8 E-7	1.7 E-6	9.7 E-7
10E5	6.6 E-3	1.1 E-4	9.3 E-7	9.0 E-7	6.8 E-7	5.5 E-7	4.8 E-7
10E6	4.5 E-3	1.8 E-5	9.3 E-7	8.7 E-7	5.1 E-7	3.6 E-7	2.9 E-7

v_1 / Γ	1E5	5E5	1E6	2E6	3E6	4E6	5E6
10E4	1.3 E-3	3.3 E-2	7.3 E-1	4.5 E+1	1.1 E-2	6.5 E+2	6.9 E+1
10E5	6.6 E-7	1.2 E-5	4.1 E-4	2.7 E-1	8.1 E-3	1.7 E-2	1.4 E+1
10E6	6.8 E-8	1.1 E-10	5.3 E-12	7.4 E-12	7.0 E-12	2.6 E-5	1.5 E-6

Table 3: $v(\boldsymbol{\alpha}_\Gamma)$ ($v_2 = 1000$, $\beta = \mathbf{0.9}$)

v_1 / Γ	1	2	4	6	8	10	12	20
10E4	7.6 E-7	6.9 E-7	2.5 E-7	2.9 E-7	1.4 E-6	1.9 E-7	7.6 E-7	6.1 E-6
10E5	6.9 E-7	6.8 E-7	4.1 E-7	1.3 E-7	2.9 E-7	3.2 E-7	3.9 E-6	3.2 E-4
10E6	3.0 E-8	1.0 E-9	5.1 E-12	4.3 E-12	5.1 E-12	4.2 E-12	5.9 E-12	7.8 E-6

Table 4: $v(\boldsymbol{\alpha}_\Gamma)$ for a constant sequence $\rho_\gamma = v_1$

We observe that the efficiency of the algorithm depends on the choice of the parameters v_1, v_2, β and is really sensitive to it when the total number of steps Γ is small.

When Γ becomes large (e.g $\Gamma = 10E6$), then the algorithm may seem to diverge if β is chosen carelessly. In fact, as the sequence $(\rho_\gamma, \gamma \geq 1)$ satisfies hypothesis (A2) of Theorem 4, the algorithm converges to the optimal parameters but it is far from $\boldsymbol{\alpha}^*$ after $10E6$ steps. However, for each value β, some v_1 reduces the mean square hedging error to 5 E-12.

5 An empirical study of the bid-ask spread costs impact

We here present numerical results corresponding to two piecewise linear liquidity cost functions Ψ:

$$\Psi^1_\lambda(T, U, \pi) = (1 + \lambda \operatorname{sign}(\pi))B(T,U)\pi \tag{28}$$

$$\Psi^2_{\lambda,C}(T, U, \pi) = \begin{cases} B(T,U)\pi \text{ for } |\pi| \leq C \\ B(T,U)(C + (1+\lambda)(\pi - C) \text{ for } \pi > C \\ B(T,U)(-C + (1-\lambda)(\pi + C) \text{ for } \pi < -C. \end{cases} \tag{29}$$

Despite the fact that we know there is no perfect hedging strategy in this context, we suppose the holder receives a null cash at time t (which is the price of the swap in a no liquidity cost market).

We now shortly check the convergence of the algorithm. Given piecewise linear cost functions Ψ, it is easy to prove that the terminal wealth $W^{(\boldsymbol{\alpha})}$ is piecewise linear in $\boldsymbol{\alpha}$ (see the proof of proposition 5). Therefore, assumption (A1) of Theorem 4. However, Theorem 4 does not apply to our context since Ψ^1_λ and Ψ^2_λ are piecewise linear and therefore are not continuously differentiable everywhere. Replace Ψ^1_λ and Ψ^2_λ by smooth approximations obtained by convolutions with kernels of the type $1/\sqrt{2\pi\varepsilon}\exp(-x^2/(2\varepsilon))$, ε small. Let $\boldsymbol{\alpha}^*_\varepsilon$ be the unique optimal parameter corresponding to the new cost functions (existence and uniqueness of $\boldsymbol{\alpha}^*_\varepsilon$ are provided by Theorem 4). In view of Rockafellar and Wets (1998, Th 7.33) $(\boldsymbol{\alpha}^*_\varepsilon)$ tends to $\boldsymbol{\alpha}^*$ when ε tends to 0.

The preceding consideration to prove convergence is more theoretical than practical: in practice, the numerical results do not differ when ε is small or ε is null.

Finaly, notice that the piecewise linearity in $\boldsymbol{\alpha}$ of $W^{(\boldsymbol{\alpha})}$ implies that conditions in Remark 7 are satisfied, so that Theorem 6 precises the convergence rate.

5.1 Taking liquidity costs into account is really necessary

Consider two different strategies: (i) the strategy corresponding to the optimal parameters $\boldsymbol{\alpha}_0$ in the idealistic model without liquidity costs and (ii) the null strategy $\boldsymbol{\delta}_0$ defined as

$$\boldsymbol{\delta}_0^{\boldsymbol{n}_{(j)}}(j,i) = 0, \quad \text{for all } -1 \le j < i \le N-1 \text{ and } \boldsymbol{n}_{(j)} \in \Lambda^{(j)}.$$

To satisfy the self-financing assumption (5), at time T_j the payoff $P(j)$ of the swap (2) is used to buy zero-coupon bonds with maturity T_N.

Figure 4 shows $-v(\boldsymbol{\alpha}_0)$ and $-v(\boldsymbol{\delta}_0)$ in terms of the parameter λ where the cost function Ψ is as in (28). The mean square hedging error dramatically increases when, in the presence of liquidity costs, the trader uses the strategy which is optimal in the no liquidity cost context. When the liquidity cost λ is larger than 4%, it is even worse to use this strategy than to use the $\boldsymbol{\delta}_0$ strategy!
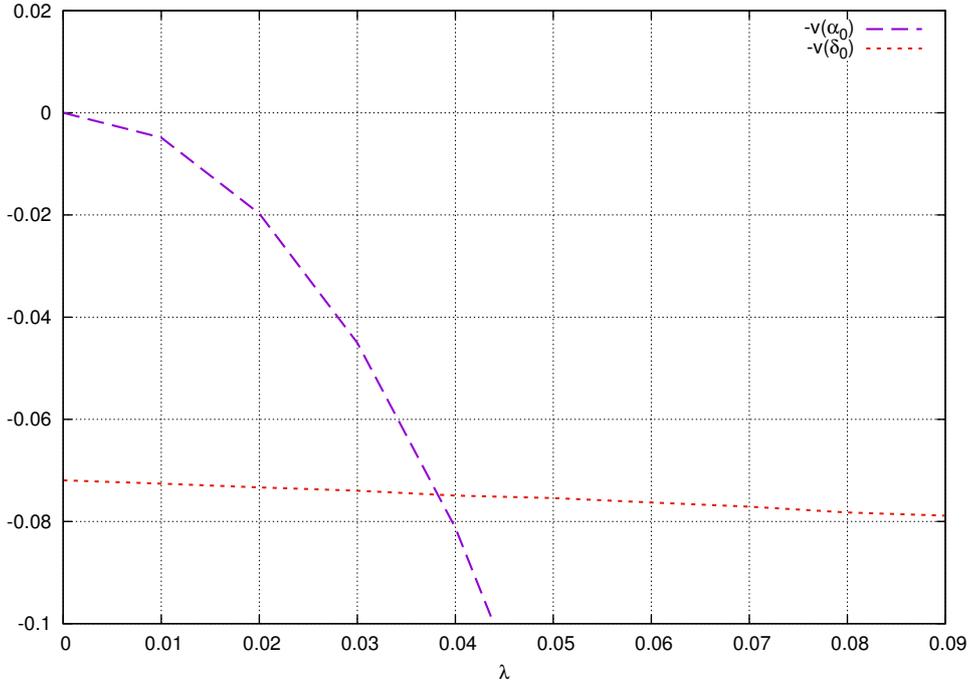

Figure 4: (short dashes) $-v(\boldsymbol{\delta}_0)$ (long dashes) $-v(\boldsymbol{\alpha}_0)$ in terms of the liquidity cost λ

21

5.2 Probability distribution of the hedging error in the case (28)

In this section, the liquidity cost function Ψ is chosen as in (28).

After Γ steps of the stochastic optimization procedure with a sample ω of the Gaussian vector $((G_\gamma^{(0)}, \cdots, G_\gamma^{(N)}), \gamma = 1, \cdots, \Gamma)$, one obtains a random approximation $\boldsymbol{\alpha}_\Gamma(\omega)$ of the optimal parameter $\boldsymbol{\alpha}^*$.

Figure 5 shows the probability distribution of the random variable $v(\boldsymbol{\alpha}_\Gamma(\omega))$ for $\Gamma = 10000$ and Table 5 shows its mean and standard deviation for different values of λ.

λ	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
Mean	2.5E-32	0.0031	0.012	0.024	0.038	0.049	0.058	0.065	0.11	0.12
Std dev.	6.7E-33	4.2E-5	2.8E-4	1.7E-3	0.016	0.017	0.046	0.081	1.1	1.3

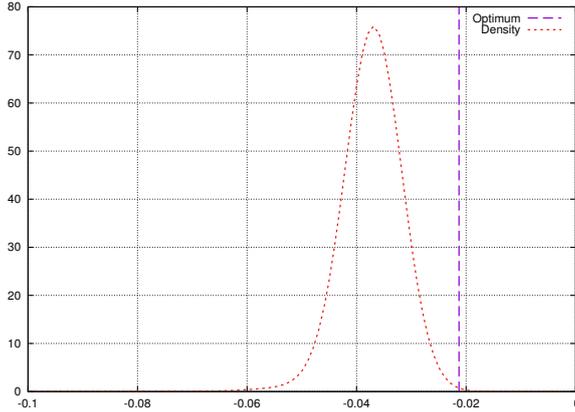Table 5: Empirical mean and standard deviation of $v(\boldsymbol{\alpha}_{10000})$.

Figure 5: Empirical density of $-v(\boldsymbol{\alpha}_\Gamma(\omega))$ ($v_1 = v_2 = \beta = 1$, $\Gamma = 10000$, $\lambda = 0.04$). The vertical line corresponds to the $v(\boldsymbol{\alpha}^*)$.

5.3 Hedging error in the case (29)

In this section, the liquidity cost function Ψ is chosen as in (29).

Figure 6 shows $-v(\boldsymbol{\alpha}_\Gamma)$ for $\Gamma = 0$ (green) and $\Gamma = 10E6$ (violet). The initial parameter $\boldsymbol{\alpha}_0$ is the optimal one for a market without liquidity cost.

22

We observe that the main part of the loss is saved thanks to the optimization procedure. In Figure 7, we zoom on the surface resulting from the optimization procedure. Notice that the cost function (28) is equal to the cost function (29) in the particular case $C = 0$. So, Figure 7 allows to compare the value functions corresponding to these two cost functions.
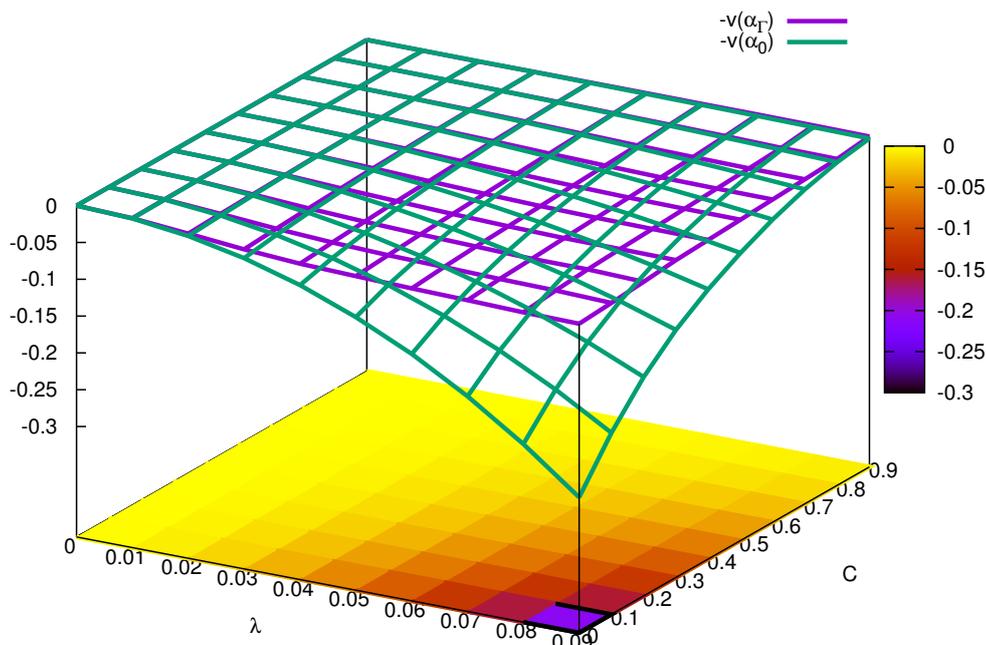

Figure 6: Violet surface: $-v(\boldsymbol{\alpha}_\Gamma)$ ($\Gamma = 10E6$, $v_1 = v_2 = 100$, $\beta = 0.6$) for a liquidity cost function (29) in terms of the value λ of the liquidity cost and the size C of the compact. Green surface:$-v(\boldsymbol{\alpha}_0)$.

5.4 Influence of the initial value $\boldsymbol{\alpha}_0$ of the optimization procedure

In Figure 8, we draw two functions of the liquidity cost λ: $-v(\boldsymbol{\alpha}_\Gamma)$ and $-v(\boldsymbol{\delta}_\Gamma)$, where $(\boldsymbol{\alpha}_\Gamma)$ and $(\boldsymbol{\delta}_\Gamma)$ are the parameters obtained after Γ steps of the optimization procedure but with different initial values $\boldsymbol{\alpha}_0$ and $\boldsymbol{\delta}_0$ as described in Subsec. 5.1. The performance of the strategies obtained after $\Gamma = 10E6$ steps are quite similar.

23

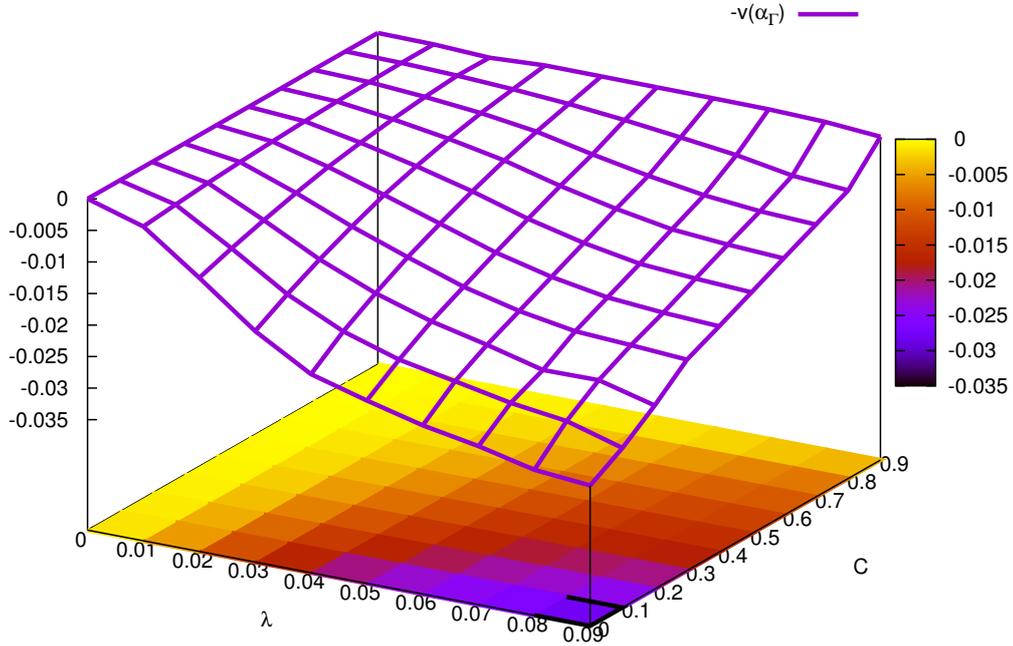Figure 7: A zoom on the violet surface of Figure 6

It means that the sensitivity to the arbitrary initial parameter $\boldsymbol{\alpha}_0$ is not observable any more after $\Gamma = 10E6$ steps.

5.5 Reducing the set of admissible strategies

Recall Hypothesis 2. So far, our admissible strategies at time T_j depend on all the past and present rates $R_t, \cdots, R_{T_j}$. Thus the number of parameters $\alpha^{\boldsymbol{n}(j)}(j, i)$ to optimize is at least of the order of magnitude of the binomial coefficient $\binom{N}{d}$, where N is the number of dates and the degree of truncation d is defined as in (26). This order of magnitude is a drastically increasing function of N. This crucial drawback leads us to try to simplify the complexity of the control problem (6) by reducing the size of the set of the admissible strategies Π. Observe that the optimal strategy under the perfect liquidity assumption has the property that $\pi^*(j, i)$ only depends on R_{T_j}. This observation suggests to face large numbers of dates by reducing the set of controls to controls depending only on a small number of recent interest rates $R_{T_j}, R_{T_{j-1}}, \cdots, R_{T_{j-q}}$.
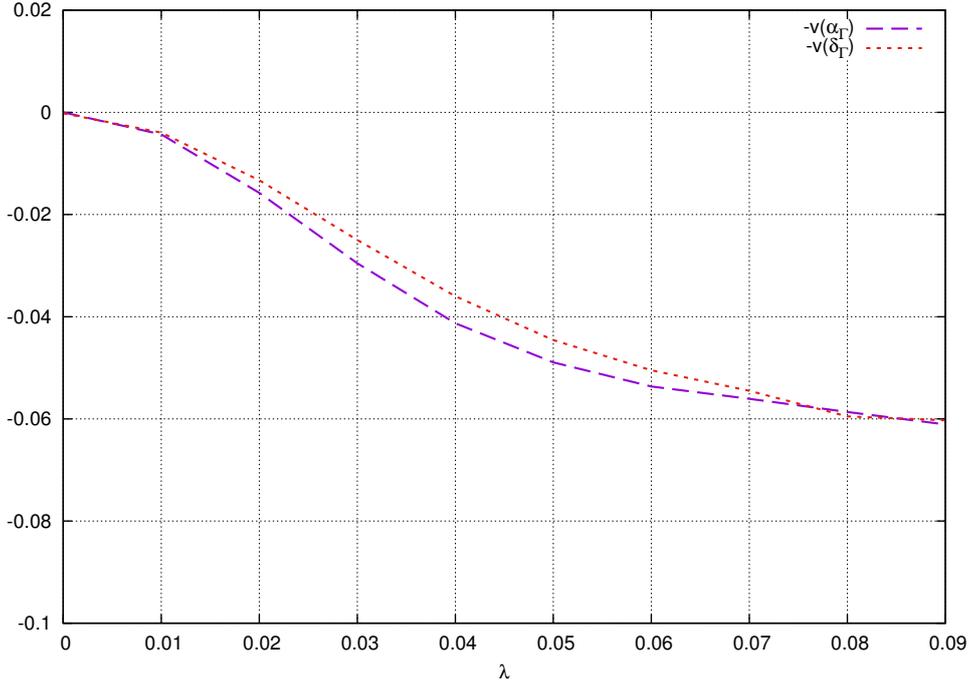
24

Figure 8: (short dashes) $-v(\boldsymbol{\delta}_\Gamma)$ (long dashes) $-v(\boldsymbol{\alpha}_\Gamma)$, in terms of the liquidity cost λ ($\Gamma = 10E6$, $v_1 = 0.1$, $v_2 = 100$, $\beta = 0.6$, $d = 3$).

Figures 9 and 10 illustrate that the optimal control problem (6) with admissible strategies defined as in hypothesis 2 may be used as benchmarks to solve control problems. Consider swaps with $N = 5$ and $N = 10$ dates of payment. In each one of these two cases, we study the effect of choosing $q = 0$ (that is at time T_j, admissible strategies only depend on R_{T_j}), $q = 1$ (admissible strategies depend on R_{T_j} and $R_{T_{j-1}}$), $q = 2$.

Figure 9 shows the performance of the corresponding strategies obtained after $\Gamma = 10E6$ steps of the optimization algorithm for a swap with $N = 5$ dates of payments.

Figure 10 shows similar quantities for a swap with $N = 10$ dates of payment. We observe that the numerical computation of the optimal strategy is quite unstable when q is too big, which reflects the difficulty to solve a high dimensional optimization problem. Therefore, one necessarily must choose q small in order to get accurate approximations of optimal strategies belonging to reduced sets of admissible strategies.
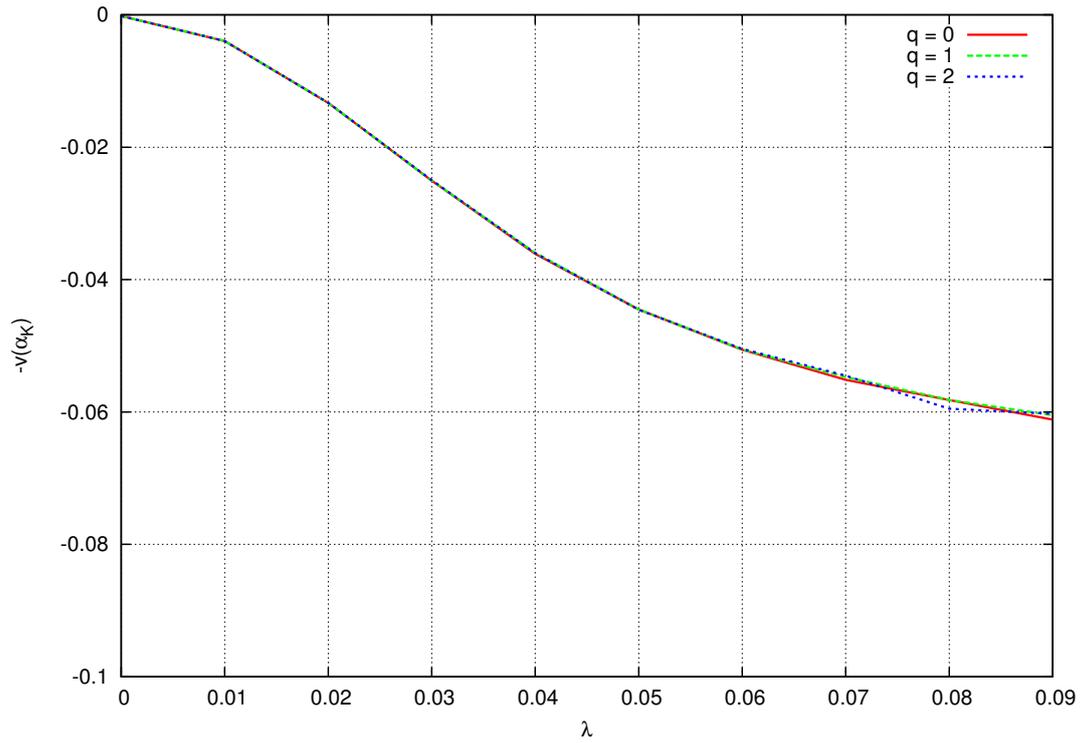
Figure 9: $-v(\boldsymbol{\alpha}_\Gamma)$ in terms of λ for strategies depending only on recent rates $R_{T_j}, \cdots, R_{T_{j-q}}$ ($q = 0, 1, 2$, $\Gamma = 10E6$, $v_1 = 0.1$, $v_2 = 100$, $\beta = 0.6$, $d = 3$) for a swap with $N = 5$ dates of payment.
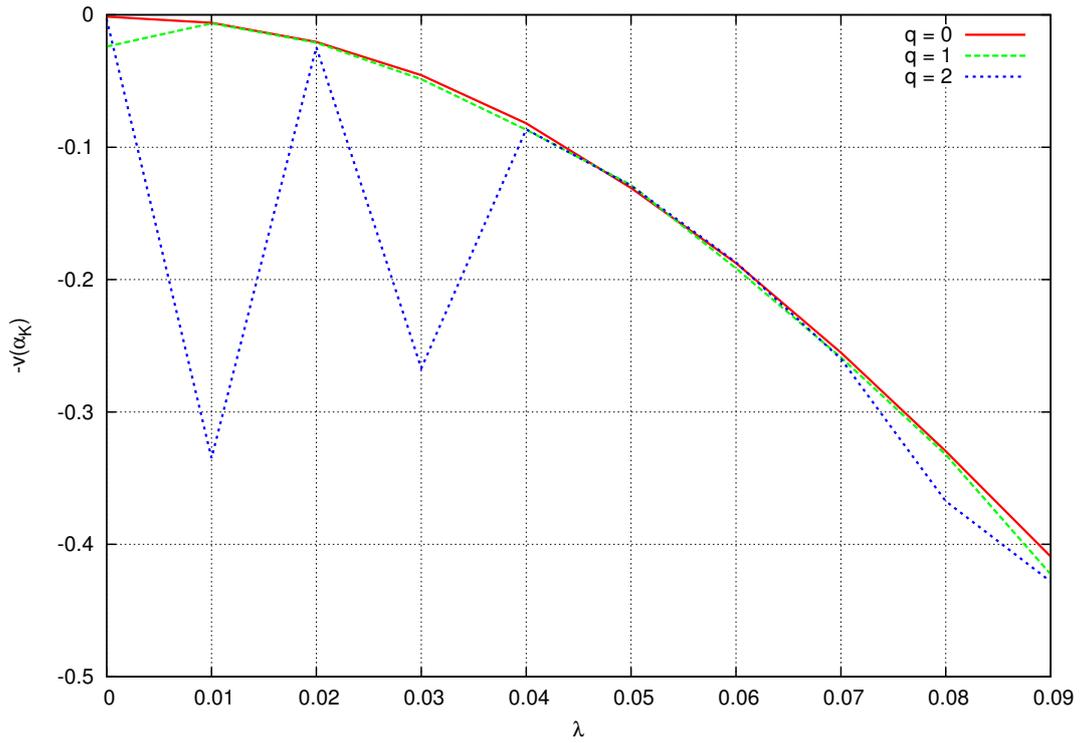
Figure 10: $-v(\boldsymbol{\alpha}_\Gamma)$ in terms of λ for strategies depending only on recent rates $R_{T_j}, \cdots, R_{T_{j-q}}$ ($q = 0, 1, 2$, $\Gamma = 10E6$, $v_1 = 0.1$, $v_2 = 100$, $\beta = 0.6$, $d = 3$) for a swap with $N = 10$ dates of payment.

6   Conclusion

Stochastic control problems generally have no explicit solutions and are difficult to solve numerically. In this paper, we have proposed an efficient algorithm to approximate optimal allocation strategies to hedge interest rate derivatives subject to liquidity costs.

As discussed above, our methodology is constructive and efficient in a Gaussian paradigm. We project the admissible allocation strategies to the space generated by the first Hermite polynomials and use a classical stochastic algorithm to optimally choose the coefficients of the projection in order to optimize an expected function of the terminal hedging error.

We have illustrated this general approach by studying swaps in the presence of liquidity costs. We have discussed the performances of the numerical method in terms of all its algorithmic components.

We emphasize that our methodology can be applied to many control problems, e.g., the computation of indifference prices, when the model under consideration belongs to a Gaussian space.

Acknowledgment

References

Babbs, S. H. and Nowman, K. B. (1999) Kalman Filtering of Generalized Vasicek Term Structure Models, *Journal of Financial and Quantitative Analysis*, 34(1), pp. 115–130.

Chen, H. F. (2002) *Stochastic Approximation and its Applications*, Nonconvex Optimization and its Applications Vol. 64 (Kluwer Academic Publishers, Dordrecht).

Chen, H. F. and Zhu, Y. (1986) Stochastic approximation procedures with randomly varying truncations, *Sci. Sinica Ser. A*, 29(9), pp. 914–926.

Dai, Q. and Singleton, K. J. (2000) Analysis of Affine Term Structure Models, *The Journal of Finance*, 55(5), pp. 1943–1978.

Duflo, M. (1997) *Random Iterative Models*, Applications of Mathematics (New York) Vol. 34 (Springer-Verlag, Berlin).

Gibson, R., Lhabitant, F. S. and Talay, D. (2010) *Modeling the Term Structure of Interest Rates: A Review of the Literature*, Foundations and Trends in Finance Vol. 5(1-2) (Now Publisher).

Kushner, H. J. and Yin, G. G. (2003) *Stochastic Approximation and Recursive Algorithms and Applications*, Second, Applications of Mathematics (New York) Vol. 35 (Springer-Verlag, New York) Stochastic Modelling and Applied Probability.

Lelong, J. (2008) Almost sure convergence for randomly truncated stochastic algorithms under verifiable conditions, *Statist. Probab. Lett.*, 78(16), pp. 2632–2636.

Lelong, J. (2013) Asymptotic normality of randomly truncated stochastic algorithms, *ESAIM Probab. Stat.*, 17, pp. 105–119.

Malliavin, P. (1995) *Integration and Probability*, Graduate Texts in Mathematics Vol. 157 (New York: Springer-Verlag).

Musiela, M. and Rutkowski, M. (2005) *Martingale methods in financial modelling*, Second, Stochastic Modelling and Applied Probability Vol. 36 (Springer-Verlag, Berlin).

Robbins, H. and Monro, S. (1951) A stochastic approximation method, *Ann. Math. Statistics*, 22, pp. 400–407.

Rockafellar, R. T. and Wets, R. J. B. (1998) *Variational analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] Vol. 317 (Springer-Verlag, Berlin).

7 Appendix

Proof of Lemma 9

Proof. Let us first prove the result for $M = 0$, that is $X = \exp\left(\mu + \lambda G\right).$

$$X = \sum_{n=0}^{+\infty} \alpha(n) \frac{H_n(G)}{\sqrt{n!}}$$

where

$$\alpha(n) = \mathbb{E}\left[X \frac{H_n(G)}{\sqrt{n!}}\right] = \exp(\mu)\mathbb{E}\left[\exp(\lambda G)\frac{H_n(G)}{\sqrt{n!}}\right].$$

We then use the identity: $e^{\lambda x - \lambda^2/2} = \sum_{j \geqslant 0} \frac{\lambda^j}{j!} H_j(x)$ and obtain

$$\alpha(n) = \exp(\mu + \lambda^2/2)\mathbb{E}\left[\sum_j \frac{\lambda^j}{j!} H_j(G)\frac{H_n(G)}{\sqrt{n!}}\right].$$

As $(H_j(G)/\sqrt{j!}, j \geq 0)$ is an orthonormal basis of $L^2(G)$ we have

$$\alpha(n) = \exp(\mu + \lambda^2/2)\frac{\lambda^n}{\sqrt{n!}}.$$

Let X^d be the projection of X on the subspace generated by the $H_0, \cdots, H_d$:

$$X^d = \sum_{n=0}^d \alpha(n)\frac{H_n(G)}{\sqrt{n!}} = e^{\mu+\lambda^2/2}\sum_{n=0}^d \frac{\lambda^n}{\sqrt{n!}}\frac{H_n(G)}{\sqrt{n!}}$$

$$X - X^d = e^{\mu+\lambda^2/2}\sum_{n=d+1}^\infty \frac{\lambda^n}{\sqrt{n!}}\frac{H_n(G)}{\sqrt{n!}}.$$

The truncation error is

$$\|X - X^d\|_2^2 = e^{2\mu+\lambda^2}\sum_{n=d+1}^\infty \frac{\lambda^{2n}}{n!}$$

$$\leq e^{2\mu+\lambda^2}\frac{\lambda^{2(d+1)}}{(d+1)!}\sum_{n=0}^\infty \frac{\lambda^{2n}}{n!}$$

$$\leq e^{2\mu+2\lambda^2}\frac{\lambda^{2(d+1)}}{(d+1)!}.$$

The desired result thus holds true for $M = 0$. Let M be a positive integer. We have

$$\|X - X^d\|_2^2 = e^{2\mu+\lambda_0^2+\cdots+\lambda_M^2}\sum_{n_0+\cdots+n_M>d}\prod_{m=0}^M \frac{\lambda_m^{2n_m}}{(n_m)!}$$

Recall the classical identity

$$\frac{(a_0 + \cdots + a_L)^N}{N!} = \sum_{n_0+\cdots+n_L=N}\prod_{m=0}^L \frac{a_m^{n_m}}{(n_m)!}. \tag{30}$$

Thus,

$$\sum_{n_0+\cdots+n_M>d}\prod_{m=0}^{M}\frac{\lambda_m^{2n_m}}{(n_m)!} = \sum_{N=d+1}^{\infty}\sum_{n_0+\cdots+n_M=N}\prod_{m=0}^{M}\frac{\lambda_m^{2n_m}}{(n_m)!}$$

$$= \sum_{N=d+1}^{\infty}\frac{\left(\lambda_0^2+\cdots+\lambda_M^2\right)^N}{N!}$$

$$\leq \frac{\left(\lambda_0^2+\cdots+\lambda_M^2\right)^{d+1}}{(d+1)!}\exp\left(\lambda_0^2+\cdots+\lambda_M^2\right).$$

This ends the proof for all positive M. $\qquad\square$