Optimal discretization of hedging strategies with directional views

Jiatu Cai, Masaaki Fukasawa, Mathieu Rosenbaum, Peter Tankov

Optimal discretization of hedging strategies with directional views

Jiatu Cai[1], Masaaki Fukasawa[2], Mathieu Rosenbaum[3] and Peter Tankov[1,4]

[1] Laboratoire de Probabilités et Modèles Aléatoires,
Université Paris Diderot (Paris 7)

[2] Department of Mathematics,
Osaka University

[3] Laboratoire de Probabilités et Modèles Aléatoires,
Université Pierre et Marie Curie (Paris 6)

[4] International Laboratory of Quantitative Finance,
National Research University Higher School of Economics, Moscow

July 16, 2014

Abstract

We consider the hedging error of a derivative due to discrete trading in the presence of a drift in the dynamics of the underlying asset. We suppose that the trader wishes to find rebalancing times for the hedging portfolio which enable him to keep the discretization error small while taking advantage of market trends. Assuming that the portfolio is readjusted at high frequency, we introduce an asymptotic framework in order to derive optimal discretization strategies. More precisely, we formulate the optimization problem in terms of an asymptotic expectation-error criterion. In this setting, the optimal rebalancing times are given by the hitting times of two barriers whose values can be obtained by solving a linear-quadratic optimal control problem. In specific contexts such as in the Black-Scholes model, explicit expressions for the optimal rebalancing times can be derived.

Key words: Discretization of hedging strategies, delta hedging, hitting times, asymptotic optimality, expectation-error criterion, semi-martingales, limit theorems, linear-quadratic optimal control.

1 Introduction

In order to manage the risks inherent to the derivatives they buy and sell, practitioners use continuous time stochastic models to compute their prices and hedging portfolios. In the simplest cases, notably in that of the so-called delta hedging strategy, the hedging portfolio obtained from the model is a time varying self financed combination of cash and the underlying of the option. We denote the price at time t of the underlying asset by Y_t and assume it to be a one-dimensional semi-martingale. Hence, in such situations, the outputs of the model are the price of the option together with the number of underlying assets to hold in the hedging portfolio at any time t, denoted by X_t (the weight in cash is then deduced from the self financing property). Therefore, assuming zero interest rates, the theoretical value of the model based hedging portfolio at the maturity of the option T is given by

$$\int_0^T X_t dY_t.$$

Typically, the process X_t derived from the model is a continuously varying semi-martingale, requiring continuous trading to be implemented in practice. This is of course physically impossible and would be anyway irrelevant because of the costs induced by microstructure effects. Hence practitioners do not use the strategy X_t, but rather a discretized version of it. This means the hedging portfolio is only rebalanced at some discrete times and thus is held constant between two rebalancing times. Let us denote by $(\tau_j^n)_{j\geq 0}$ an increasing sequence of rebalancing times over $[0, T]$ (the meaning of the parameter n will be explained below). With respect to the target portfolio obtained from the model, the hedging error due to discrete trading Z_T^n is therefore given by

$$Z_T^n = \sum_{j=0}^{+\infty} X_{\tau_j^n}(Y_{\tau_{j+1}^n \wedge T} - Y_{\tau_j^n \wedge T}) - \int_0^T X_t dY_t.$$

Thus, some important questions in practice are:

- What is the order of magnitude of Z_T^n in the case of classical discretization strategies ?
- For a given criterion, how to optimize the rebalancing times ?

The most classical rebalancing scheme is that of equidistant trading dates of the form

$$\tau_j^n = jT/n, \; j = 0, \ldots, n,$$

where n represents the total number of trades on the period $[0, T]$. In this setting, the first question has been addressed in details. There are two popular approaches to quantify the hedging error Z_T^n, both of them being asymptotic, assuming the rebalancing frequency n/T tends to infinity (that is n tends to infinity since T is fixed). A first possibility is to use the L^2 norm, where one typically looks for asymptotic bounds of the form

$$\mathbb{E}[(Z_T^n)^2] \leq cn^{-\theta}, \quad n \to \infty.$$

Many authors have explored various aspects of this problem in this deterministic rebalancing dates framework. For European call and put options in the Black-Scholes model, it is shown in [1] and [17] that the L^2 error has a convergence rate $\theta = 1$. For other options, the convergence rate depends on the regularity of the payoff. For example, it is shown in [8] that for binary options, the convergence rate is $\theta = 1/2$. However, in this context, the convergence rate $\theta = 1$ can be achieved by choosing a suitable non equidistant deterministic rebalancing grid defined by

$$\tau_j^n = T - T\,(1 - j/n)^{1/\beta},$$

with $\beta \in (0,1]$ being the fractional smoothness in the Malliavin sense of the option payoff, see [6]. An asymptotic lower bound for the L^2 error is given in [3, 5] for a general class of rebalancing schemes.

The second way to assess the hedging error is through the weak convergence of the sequence of the suitably rescaled random variables Z_T^n. When X and Y are Itô processes, the case of equidistant rebalancing dates has been investigated in this approach in [1, 9, 13], where the following convergence in law is proved:

$$\sqrt{n} Z_T^n \xrightarrow{\mathcal{L}} \sqrt{\frac{T}{2}} \int_0^T \sigma_t^X \sigma_t^Y dB_t, \tag{1.1}$$

where σ^X and σ^Y are the volatilities of X and Y and B is a Brownian motion independent of the other quantities. The case where X and Y are processes with jumps is treated in [16].

This asymptotic approach has also been recently used in the context where the rebalancing times are random stopping times. Some specific hitting times based schemes derived from a microstructure model are investigated in [12]. In [4], the author works with quite general sampling schemes based on stopping times. More precisely, for a given parameter n driving the asymptotic, one considers an increasing sequence of stopping times

$$0 = \tau_0^n \leq \tau_1^n \leq \ldots \leq \tau_j^n \leq \ldots$$

so that almost surely, $\lim_{j \to \infty} \tau_j^n = T$ (meaning in fact that the stopping times are all equal to T for large enough j) and

$$\sup_j (\tau_{j+1}^n - \tau_j^n)$$

tends to 0 in a suitable sense as n goes to infinity. Under some regularity conditions on the (random) rebalancing times, a general limit theorem for the hedging error is obtained in [4]. It is shown that after suitable renormalization (specified in the next sections), the hedging error converges in law to a random variable of the form

$$\frac{1}{3} \int_0^T s_t dY_t + \frac{1}{\sqrt{6}} \int_0^T \left(a_t^2 - \frac{2}{3} s_t^2\right)^{1/2} \sigma_t^Y dB_t. \tag{1.2}$$

Here B is a Brownian motion independent of all the other quantities and the processes s and a can be interpreted as the asymptotic local conditional skewness and kurtosis of the increments of the process X between two consecutive discretization dates (see next sections for details).

One can remark a crucial difference between the deterministic discretization schemes associated to (1.1) and the random stopping times case leading to (1.2). For deterministic dates, the discretization error asymptotically behaves as a stochastic integral with respect to Brownian motion. Therefore, it is (essentially) centered. In the case of random discretization dates, one may obtain a "biased" asymptotic hedging error because of the presence of the term

$$\int_0^T s_t dY_t.$$

Hence, if s does not vanish and Y has non zero drift, the asymptotic hedging error is no longer centered.

From a practitioner viewpoint, this is quite an interesting property. Indeed, it shows that in the presence of market trends, the trader may actually be compensated for the extra risk arising from discrete trading, provided that the rebalancing dates are chosen in an appropriate way. Of course one may say this is not the option trader's job to try to get a positive expected return with the hedging strategy. However, knowing that there is anyhow a hedging error, it seems reasonable to optimize it to the trader's benefit.

Hence we place ourselves in the asymptotic high frequency regime where n is large and therefore

$$\sup_j (\tau_{j+1}^n - \tau_j^n)$$

is small, meaning that the hedging error should be small. In this setting we address the second question raised above, that is finding the optimal times to rebalance the portfolio. To do so, we simply use an asymptotic expectation-error type criterion. More precisely, we wish to maximize the expectation of the hedging error under a constraint on its L^2 norm. This is quite in the spirit of [15], where the author aims at finding an optimal hedging frequency to optimize the Sharpe ratio. Remark that in our context, the L^2 norm is more meaningful than the variance since the primary goal of the trader is to make the hedging error small. Our asymptotic approach goes as follows. First, we approximate the law of the renormalized hedging error by that in Equation (1.2). Then we find the processes a_t^* and s_t^* which correspond to optimality in terms of our expectation-error criterion for the family of laws given by (1.2). Finally, we show that we can indeed build a discretization rule which leads to the optimal a_t^* and s_t^* in the limiting distribution of the hedging error.

Note that using an asymptotic framework to design optimal discretizations of hedging strategies has been a quite popular approach in the recent years. Such method (although in a slightly different context) is in particular used in [3, 4, 7] in the continuous setting whereas the case with jumps is investigated in [14]. All these works aim at minimizing some form of transaction costs (typically the number of trades) under some constraint on the L^2 norm of the hedging error. Here we also put a constraint on the L^2 norm of the hedging error. However, instead of minimizing transaction costs, we maximize the expectation of the hedging error. Thus our viewpoint is that of a trader giving himself a lower bound on the quality of his hedge (the L^2 norm of the hedging error), but allowing himself to try to take advantage of market drifts provided the constraint is satisfied.

In practice, our work should probably only be considered as a benchmark. Indeed, we somehow make the highly unreasonable assumption that practitioners observe the drift. This is of course not realistic at all since any kind of statistical estimation of the drift is irrelevant in this high frequency setting. However, some practitioners still have views on the market and our work gives them a way to incorporate their beliefs in their hedging strategies.

The paper is organized as follows. In Section 2, we investigate the set of admissible discretization rules, that is those leading to a limiting law of the form (1.2). In particular, we extend the examples provided in [4] by showing that the discretization rules based on hitting times of stochastic barriers are admissible. In Section 3, we consider a first criterion for optimizing the trading times: the modified Sharpe ratio. It enables us to carry out very simple computations. However, the relevance of the modified Sharpe ratio being in fact quite arguable, a more suitable approach in which we consider an expectation-error type criterion is investigated in Section 4. Using tools from linear-quadratic optimal control theory, explicit developments are provided in the Black-Scholes model in Section 5. Finally, the longest proofs are relegated to an appendix.

2 Assumptions and admissible strategies

In this section we detail our assumptions on the processes X and Y together with the admissibility conditions for the sampling schemes.

2.1 Assumptions on the dynamics and admissibility conditions

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space. We write Y for the underlying asset. Let $T > 0$ stand for the maturity of the derivative to be hedged. We assume that the benchmark hedging strategy deduced from a theoretical model simply consists in holding a certain number of units of the underlying asset, denoted by X, and some cash in a self financed way, under zero interest rates. Throughout the paper, we assume both Y and X are Itô processes of the form

$$dY_t = b_t^Y dt + \sigma_t^Y dW_t^Y, \;\; dX_t = b_t^X dt + \sigma_t^X dW_t^X \tag{2.1}$$

on $[0, T]$, where W^X and W^Y are $\mathbb{F}$-Brownian motions which may be arbitrarily correlated, and the coefficients of X and Y satisfy the following technical assumptions.

Assumption 2.1.

- *The processes b^Y, b^X, σ^Y and σ^X are adapted and continuous on $[0, T]$ almost surely.*
- *The volatility process σ^Y of Y is positive on $[0, T]$ almost surely.*
- *The volatility process σ^X of X is positive on $[0, T)$ almost surely.*
- *The instantaneous Sharpe ratio $\rho = b^Y/\sigma^Y$ satisfies*

$$\mathbb{E}\Big[\int_0^T \rho_t^2 dt\Big] < +\infty.$$

Example 2.1 (The Black-Scholes model)**.** The case that $b_t^Y = bY_t$ and $\sigma_t^Y = \sigma Y_t$ with constants b and $\sigma > 0$ corresponds to the Black-Scholes model. The instantaneous Sharp ratio $\rho = b/\sigma$ is a constant. To hedge a call option with payoff $(Y_T - K)_+$ and strike $K > 0$, the standard theory suggests to use the so-called Delta hedging strategy:

$$X_t = \Phi(d_1(t, Y_t)), \;\; d_1(t, y) = \frac{\log(y/K) + \sigma^2(T-t)/2}{\sigma\sqrt{T-t}},$$

where Φ stands for the distribution function of a standard Gaussian random variable. By Itô's formula, we see that X is an Itô process of the form (2.1) with $W^X = W^Y$ and

$$b_t^X = \phi(d_1(t, Y_t)) \left\{ \frac{\partial d_1}{\partial t}(t, Y_t) + \frac{\sigma^2}{2} \frac{\partial^2 d_1}{\partial y^2}(t, Y_t) Y_t^2 + b \frac{\partial d_1}{\partial y}(t, Y_t) Y_t \right\} + \frac{\sigma^2}{2} \Big(\frac{\partial d_1}{\partial y}(t, Y_t) \Big)^2 \phi'(d_1(t, Y_t)) Y_t^2,$$

$$\sigma_t^X = \sigma \phi(d_1(t, Y_t)) \frac{\partial d_1}{\partial y}(t, Y_t) Y_t,$$

with ϕ the density of a standard Gaussian random variable. Almost surely, $Y_T \neq K$ and therefore both b^X and σ^X are continuous on $[0, T]$ and $b_T^X = \sigma_T^X = 0$. Furthermore σ^X is positive on $[0, T)$. Hence Assumption 2.1 is satisfied.

As explained in the introduction, in practice, the trader cannot realize the theoretical strategy X_t which typically implies continuous trading. Hence the quantity

$$\int_0^T X_s dY_s$$

only represents a benchmark terminal wealth and X_t is a benchmark hedging strategy. Thus, we consider that this strategy is discretized over the stopping times

$$0 = \tau_0^n \leq \tau_1^n \leq \cdots \leq \tau_j^n \leq \cdots,$$

so that for given n, almost surely, τ_j^n attains T for j large enough. Such array of stopping times is called a discretization rule. Consequently, if we define the discretized process X^n by

$$X_t^n = X_{\tau_j^n}, \quad t \in [\tau_j^n, \tau_{j+1}^n),$$

the hedging error Z_T^n with respect to the benchmark strategy writes

$$Z_T^n = \int_0^T (X_{s-}^n - X_s) dY_s.$$

We now define the admissibility conditions for our discretization rules which we comment in the next subsection.

Condition 2.1 (Admissibility conditions)**.** *A discretization rule (τ_j^n) is admissible if there exist continuous $\mathbb{F}$-adapted processes a and s satisfying*

$$\mathbb{E}\left[\int_0^T \Big(1 + (\rho_t)^2\Big)(a_t^2 + s_t^2)(\sigma_t^Y)^2 dt \right] < \infty, \tag{2.2}$$

and a positive sequence ε_n tending to zero such that:

- *The first two moments of the renormalized hedging error $\varepsilon_n^{-1} Z_T^n$ converge to those of a random variable of the form*

$$Z_{a,s}^* = \frac{1}{3} \int_0^T s_t dY_t + \frac{1}{\sqrt{6}} \int_0^T \Big(a_t^2 - \frac{2}{3} s_t^2 \Big)^{1/2} \sigma_t^Y dB_t, \tag{2.3}$$

  that is,

$$\mathbb{E}[\varepsilon_n^{-1} Z_T^n] \to \mathbb{E}[Z_{a,s}^*], \quad \mathbb{E}[(\varepsilon_n^{-1} Z_T^n)^2] \to \mathbb{E}[(Z_{a,s}^*)^2], \tag{2.4}$$

  where B is a Brownian motion independent of all the other quantities.

- *Almost surely, the processes a_t and s_t satisfy $a_t^2 \geq s_t^2$, for all $t \in [0, T]$.*

2.2 Comments on the admissibility conditions

Equation 2.2 is simply a technical integrability condition. We now give the interpretation of the sequence ε_n. Recall that for fixed n, we deal with an increasing sequence of stopping times (τ_j^n) over $[0,T]$. Typically, ε_n^2 will represent the order of magnitude of the interarrival time $\tau_{j+1}^n - \tau_j^n$. For example, in the case of equidistant trading times with frequency n/T, ε_n can simply be taken equal to $n^{-1/2}$. In the case of the hitting times based scheme consisting in rebalancing the portfolio each time the process X has varied by ν_n, where ν_n is a deterministic sequence tending to zero, one can choose $\varepsilon_n = \nu_n$ (since the order of magnitude of the time interval between two hitting times is ν_n^2).

The specific form (2.3) may appear rather ad hoc at first sight. However, it is in fact quite natural. Indeed, Proposition 2.1 below, which is proved in Appendix and used to show the main result of the next subsection, indicates that as soon as the quadratic covariations $\varepsilon_n^{-2}\langle Z^n\rangle$ and $\varepsilon_n^{-1}\langle Z^n, Y\rangle$ have regular limits, the form (2.3) appears for the weak limit of the renormalized hedging error. So the idea for this admissibility condition is that in our asymptotic approach, we want to work in regular cases where the renormalized hedging error can be approximated by a random variable of the form (2.3). However, our asymptotic optimality criterion will be based on the first two moments of the renormalized hedging error only. Therefore, we just require these first two moments to be asymptotically close to those of a random variable of the form (2.3) (in particular we do not impose the convergence in law of the renormalized hedging error towards $Z^*_{a,s}$, although this is the underlying idea behind this admissibility condition). We now give Proposition 2.1.

Proposition 2.1. *If there exist a sequence $\varepsilon_n \to 0$ and continuous processes s and a such that*

$$\varepsilon_n^{-2}\langle Z^n\rangle \to \frac{1}{6}\int_0^{\cdot} a_u^2(\sigma_u^Y)^2 du, \tag{2.5}$$

$$\varepsilon_n^{-1}\langle Z^n, Y\rangle \to \frac{1}{3}\int_0^{\cdot} s_u(\sigma_u^Y)^2 du, \tag{2.6}$$

uniformly in probability on $[0,T]$, then $\varepsilon_n^{-1}Z^n$ converges in law to

$$\frac{1}{3}\int_0^{\cdot} s_t dY_t + \frac{1}{\sqrt{6}}\int_0^{\cdot}\left(a_t^2 - \frac{2}{3}s_t^2\right)^{1/2}\sigma_t^Y dB_t \tag{2.7}$$

in $C[0,T]$. In particular the convergence in law of $\varepsilon_n^{-1}Z_T^n$ to $Z^*_{a,s}$ defined by (2.3) *holds. If in addition,*

$$\varepsilon_n^{-4/3}\sup_{j\geq 0}(\tau_{j+1}^n \wedge T_0 - \tau_j^n \wedge T_0) \to 0 \tag{2.8}$$

in probability, for all $T_0 \in [0,T)$, then almost surely $a_t^2 \geq s_t^2$ for all $t \in [0,T]$.

We now consider the processes a_t^2 and s_t appearing in the admissibility conditions. We place ourselves in the situation where Proposition 2.1 can be applied. In that case, an inspection of the proof of this lemma shows that the inequality $a_t^2 \geq s_t^2$ essentially follows from the elementary fact that $\mathbb{E}[\Delta^2]\mathbb{E}[\Delta^4] \geq \mathbb{E}[\Delta^3]^2$ for a general random variable Δ. Indeed, a_t^2 and s_t are respectively related to the local third and fourth conditional moments of the increments of X. Proposition 2.2 below, which is proved in Appendix and used to show the main result in the next subsection, somehow illustrates the connections between a_t^2 and s_t and the conditional moments. Thus we give it now. Let $\Delta_{j,n} = X_{\tau^n_{j+1}} - X_{\tau^n_j}$ be the increment of X between τ^n_j and τ^n_{j+1} and N^n_t be the number of rebalancing times until time t:

$$N^n_t = \max\left\{j \geq 0 | \tau^n_j \leq t\right\}.$$

The following proposition holds.

Proposition 2.2. *Let ε_n be a positive sequence tending to* 0 *and s and a be continuous processes. Assume the following:*

- *The family of random variables*

$$\varepsilon_n^{-4} \sup_{t\in[0,T]} |X^n_t - X_t|^4 \tag{2.9}$$

is uniformly integrable.

- *The following uniform convergences in probability on* $[0, T_0]$ *hold for all* $T_0 \in [0, T)$*:*

$$\begin{aligned}
\varepsilon_n^{-1} \sum_{j=0}^{N^n_\cdot} \kappa_{\tau^n_j} \mathbb{E}\Big[\Delta^3_{j,n} \Big| \mathcal{F}_{\tau^n_j}\Big] &\to -\int_0^\cdot s_u(\sigma^Y_u)^2 du,\\
\varepsilon_n^{-2} \sum_{j=0}^{N^n_\cdot} \kappa_{\tau^n_j} \mathbb{E}\Big[\Delta^4_{j,n} \Big| \mathcal{F}_{\tau^n_j}\Big] &\to \int_0^\cdot a^2_u(\sigma^Y_u)^2 du,
\end{aligned} \tag{2.10}$$

where $\kappa_u = (\sigma^Y_u/\sigma^X_u)^2$.

Then the convergences (2.5), (2.6) *and* (2.8) *hold.*

Proposition 2.2 is useful to obtain the convergences (2.5) and (2.6) for a given discretization rule since it is usually easy to have approximate values of the conditional moments of the increments. We actually apply this approach in the proof of the main result of the next subsection.

2.3 Examples of admissible discretization rules

We show in this section that the most common discretization rules are admissible. We start with hitting times based schemes. We have the following result.

Proposition 2.3 (Hitting times based discretization rule)**.** *Let ε_n be a positive sequence tending to zero and $\underline{l}$ and $\bar{l}$ be two adapted processes which are positive and continuous on* $[0, T]$ *almost surely with*

$$\mathbb{E}\left[\int_0^T \Big(1 + (\rho_t)^2\Big)(\bar{l}_t \vee \underline{l}_t)^2 (\sigma^Y_t)^2 dt\right] < \infty. \tag{2.11}$$

The discretization rule based on the hitting times of $\varepsilon_n \underline{l}_t$ or $\varepsilon_n \bar{l}_t$ by the process X:

$$\tau^n_{j+1} = \inf\left\{t > \tau^n_j : X_t \notin (X_{\tau^n_j} - \varepsilon_n \underline{l}_t, X_{\tau^n_j} + \varepsilon_n \bar{l}_t)\right\} \wedge T \tag{2.12}$$

is admissible. Moreover, we can take

$$s_t = \underline{l}_t - \bar{l}_t, \quad a_t^2 = (s_t)^2 + \underline{l}_t \bar{l}_t \tag{2.13}$$

and we also have the convergence (2.7) *and therefore the convergence in law of Z_T^n towards $Z_{a,s}^*$ defined by* (2.3).

It is interesting to note here that the limit $Z_{a,s}^*$ does not depend on the structure of X.

This result is particularly important since many traders monitor the values of the increments of their so-called delta (which corresponds to the process X) in order to decide when to rebalance their portfolio. Thus they are indeed using hitting times based strategies. This proposition notably extends the weak convergence results in [4] since it shows that not only constant barriers (between τ_j^n and τ_{j+1}^n) but also time varying stochastic barriers can be considered. This will be very useful in the next sections since our optimal discretization rules will correspond to hitting times of such barriers. Furthermore, in the proofs, the assumption that the time varying barriers satisfy (2.11) will enable us to deduce quite easily some relevant integrability properties for the hedging error (which would be harder to obtain with locally constant barriers).

Now remark that under the condition $a_t^2 > s_t^2$, we can always find some positive processes $\bar{l}_t$ and $\underline{l}_t$ such that (2.13) is satisfied. Indeed, it is easy to see that the real numbers $\bar{l}_t$ and $-\underline{l}_t$ can be taken as the roots of the quadratic equation $x^2 + s_t x + s_t^2 - a_t^2 = 0$. Under the condition $a_t^2 > s_t^2$, this equation admits two nonzero roots with different signs. Therefore, another interesting property of hitting times based schemes is the following.

Lemma 2.1. *For any pair of limiting processes s and a satisfying* (2.2) *and $a_t^2 > s_t^2$, we can always build a corresponding admissible discretization rule based on hitting times as in* (2.12)-(2.13).

Consequently, if one has some processes a_t and s_t as targets, Lemma 2.1 implies that a strategy giving rise to these processes in the limiting distribution (2.3) can be found. We will work in this framework in Section 4. Remark that there are infinitely many strategies for which the hedging error converge in law to some $Z_{a,s}^*$ with the same a and s as limiting processes. The hitting time strategy is an efficient one among them, in the sense that it requires the least number of rebalancing in an asymptotic sense, see [4] for the detail.

Another classical discretization rule is given by equidistant trading times. Here, the integrability property (2.4) in the admissibility conditions does not hold in full generality. Compared to the hitting times setting, this is because the deviations of the benchmark strategy are not explicitly controlled by the barriers. Nevertheless, the following example describes a reasonable framework under which such discretization rule is admissible.

Proposition 2.4 (Equidistant sampling discretization rule)**.** *Consider the hedging strategy of a European option with payoff $h(Y_T)$ and replace Assumption 2.1 by that the underlying Y_t follows a diffusion process of the form*

$$dY_t = b(t, Y_t)Y_t dt + \sigma(t, Y_t)Y_t dW_t,$$

with b, σ and h some deterministic functions satisfying the regularity assumptions p.21-23 in [17] (allowing in particular for call and put in the Black-Scholes model). Define the delta hedging portfolio:

$$X_t = \frac{\partial P}{\partial y}(t, Y_t), \text{ with } P(t,y) = \mathbb{E}^{\mathbb{Q}}_{(t,y)}[h(Y_T)],$$

where $\mathbb{E}^{\mathbb{Q}}$ denotes the expectation operator under the risk neutral probability. Let ε_n be a positive sequence tending to zero. Then the equidistant trading times discretization rule:

$$\tau^n_j = j\varepsilon_n^2, \quad j = 0, \dots, n, \dots$$

is admissible (under the original measure). Moreover, we can take

$$s_t = 0, \quad a_t^2 = 3(\sigma_t^X)^2.$$

The proof of Proposition 2.4 follows easily from previous works. We can first obtain the convergence in law towards $Z^*_{a,s}$ using for example the results of [9]. Indeed, up to localization, we can assume that σ^X, σ^Y, b^X and b^Y are bounded. Then the integrability conditions in the mentioned reference are obviously satisfied and the convergence follows. For (2.4), it suffices to use Theorem 2.4.1 in [17] where the convergence of the L^2 norm of the normalized error under the original measure is provided.

Finally, note that the discretization rule based on equidistant trading times will not be of interest for us since the associated s_t process vanishes and so the expectation of the limiting variable is zero.

3 Asymptotic optimality: a preliminary approach

Our viewpoint is that the trader's priority is to get a small hedging error. However, once this error is suitably controlled, he may try to take advantage of the directional views he has on the market. Hence, adopting the asymptotic approximation under which the first two moments of the renormalized hedging error are given by those of $Z^*_{a,s}$, we aim at maximizing $\mathbb{E}[Z^*_{a,s}]$ while keeping $\mathbb{E}[(Z^*_{a,s})^2]$ reasonably small. This very problem is treated in Section 4.

Here, as a first step, we consider the approximation for $\mathbb{E}[(Z^*_{a,s})^2]$ given by $\mathbb{E}[(Z^{*,c}_{a,s})^2]$, where $Z^{*,c}_{a,s}$ denotes the sum of the two integrals with respect to the Brownian motions W^Y and B in the definition of $Z^*_{a,s}$ in Equation (2.3), that is

$$Z^{*,c}_{a,s} = \frac{1}{3}\int_0^T s_t\sigma_t^Y dW_t^Y + \frac{1}{\sqrt{6}}\int_0^T \left(a_t^2 - \frac{2}{3}s_t^2\right)^{1/2} \sigma_t^Y dB_t.$$

To do so, we place ourselves in this section under the additional admissibility condition that the renormalized hedging error weakly converges in the sense of (2.7) and we take s_t and a_t^2 as the processes in the limit (2.7) (so s_t and a_t^2 are uniquely defined). Replacing $\mathbb{E}[(Z^*_{a,s})^2]$ by $\mathbb{E}[(Z^{*,c}_{a,s})^2]$ is technically very convenient but in practice quite arguable since this approximation is meaningful only when the drift is small. However, our aim here is only to have a first rough idea about the form of the optimal discretization rules. Since we wish to get the moment of order one large while that of order two remains controlled, we consider that we want to maximize the so-called modified Sharpe ratio S defined by

$$S = S(a,s) = \frac{\mathbb{E}[Z^*_{a,s}]}{\sqrt{\mathbb{E}[(Z^{*,c}_{a,s})^2]}}.$$

This ratio is said to be modified since we use $\mathbb{E}[(Z^{*,c}_{a,s})^2]$ instead of the variance of $Z^*_{a,s}$.

Hence we are looking for strategies which maximize S. To do so, we now introduce the notion of nearly efficient (modified) Sharpe ratio.

Definition 3.1 (Nearly efficient Sharpe ratio)**.** *The value $S^* \in \mathbb{R}$ is said to be a nearly efficient Sharpe ratio if:*

- *For any admissible discretization rule with associated limiting processes a and s, the associated modified Sharpe ratio $S(a,s)$ satisfies*

$$S(a,s) \leq S^*.$$

- *For any $\eta > 0$, there exists a discretization rule with associated limiting processes a and s such that*

$$S(a,s) \geq S^* - \eta.$$

We only consider nearly efficient ratios since our strategies will not enable us to attain exact efficiency (which would corresponds to $\eta = 0$ in the previous definition). Of course the slight difference between efficient and nearly efficient ratios has no importance in practice.

In our setting, for any limiting variable $Z^*_{a,s}$, we have

$$S(a,s) = \frac{\mathbb{E}\Big[\frac{1}{3}\int_0^T s_t b^Y_t dt\Big]}{\Big(\mathbb{E}\Big[\frac{1}{9}\int_0^T s_t^2(\sigma^Y_t)^2 dt + \frac{1}{6}\int_0^T \big(a_t^2 - \frac{2}{3}s_t^2\big)(\sigma^Y_t)^2 dt\Big]\Big)^{1/2}}.$$

Now, the admissibility condition $a_t^2 \geq s_t^2$ implies

$$S(a,s) \leq \frac{\sqrt{6}}{3}\frac{\mathbb{E}\Big[\int_0^T s_t b^Y_t dt\Big]}{\Big(\mathbb{E}\Big[\int_0^T s_t^2(\sigma^Y_t)^2 dt\Big]\Big)^{1/2}}$$

and Cauchy-Schwarz inequality gives

$$S(a,s) \leq \frac{\sqrt{6}}{3}\Big(\mathbb{E}\Big[\int_0^T \Big(\frac{b^Y_t}{\sigma^Y_t}\Big)^2 dt\Big]\Big)^{1/2}.$$

This provides an upper bound for the modified Sharpe ratio. We now wish to find a discretization rule enabling to (almost) attain this upper bound. To achieve this, our rule must be so that for the associated processes a_t and s_t, the inequalities used above ($a_t^2 \geq s_t^2$ and

Cauchy-Schwarz) become almost equalities. This means that a_t should be close to s_t and s_t essentially proportional to $b_t^Y/(\sigma_t^Y)^2$. Furthermore, we want the product $s_t b_t^Y$ to be essentially positive in order to get a positive modified Sharpe ratio. If we look for this rule among the hitting times based schemes specified by two processes $(\underline{l}_t, \bar{l}_t)$, Lemma 2.1 implies that

- the difference $\underline{l}_t - \bar{l}_t$ should be essentially proportional to $b_t/(\sigma_t^Y)^2$,
- the product $\underline{l}_t \bar{l}_t$ should be negligible compared to $(\underline{l}_t - \bar{l}_t)^2$,
- the term $(\underline{l}_t - \bar{l}_t)b_t^Y$ should be essentially positive.

From these remarks together with Proposition 2.3, we easily deduce the following theorem.

Theorem 3.1. *Suppose that for all $t \leq T$, $b_t^Y \neq 0$. Then the value*

$$\frac{\sqrt{6}}{3}\Big(\mathbb{E}\Big[\int_0^T \Big(\frac{b_t^Y}{\sigma_t^Y}\Big)^2 dt\Big]\Big)^{1/2}$$

is a nearly efficient Sharpe ratio. It is approximately attained by the discretization rule defined for $\lambda > 0$ by

$$\tau_{j+1}^{n,\lambda} = \inf\Big\{t > \tau_j^{n,\lambda}; X_t - X_{\tau_j^{n,\lambda}} = -\frac{b_t^Y}{(\sigma_t^Y)^2}e^{\lambda}\varepsilon_n \text{ or } \frac{b_t^Y}{(\sigma_t^Y)^2}e^{-\lambda}\varepsilon_n\Big\}, \quad \tau_0^n = 0. \tag{3.1}$$

Indeed,

$$\lim_{\lambda \to +\infty} S(\lambda) = \frac{\sqrt{6}}{3}\Big(\mathbb{E}\Big[\int_0^T \Big(\frac{b_t^Y}{\sigma_t^Y}\Big)^2 dt\Big]\Big)^{1/2},$$

where $S(\lambda)$ denotes the modified Sharpe ratio obtained for the law of the variable $Z_{a,s}^*$ associated to the discretization rule (3.1) *with parameter λ.*

This result provides simple and explicit strategies for optimizing the modified Sharpe ratio. It is also very easy to interpret. Indeed, we see that in order to take advantage of the drift, one needs to consider asymmetric barriers. The limitation is that we do not really control accurately the magnitude of the hedging error at maturity.

The asymptotic setting simply means that we require λ to be quite large while $e^{\lambda}\varepsilon_n$ is small. When using such discretization rule in practice, it is reasonable to consider that the trader fixes a maximal value for the asymmetry between the barriers controlled by λ. This way he can choose the parameter λ. Then ε_n is set to match the bound on $\mathbb{E}[(Z_{a,s}^{*,c})^2]$ that the trader does not want to exceed.

4 Asymptotic expectation-error optimization

In this section, we now consider a natural expectation-error type criterion in order to optimize our discretization rules. To do so, we work in an asymptotic setting where we are looking for discretization rules which are optimal in the expectation-error sense for their associated limiting random variable $Z_{a,s}^*$. Before giving our main result, we introduce some definitions inspired by classical portfolio theory.

Definition 4.1 (Non dominated couple)**.** *A couple $(m, v) \in (\mathbb{R}^+)^2$ is said to be non dominated if there exists no admissible discretization rule such that its associated limiting random variable $Z^*_{a,s}$ satisfies*

$$\mathbb{E}[Z^*_{a,s}] \geq m, \ \ \mathbb{E}[(Z^*_{a,s})^2] < v.$$

The set of non dominated couples is called the non domination domain.

Definition 4.2 (Nearly efficient couple)**.** *A couple $(m, v) \in (\mathbb{R}^+)^2$ is said to be nearly efficient if it is non dominated and for any $\eta > 0$, there exists an admissible discretization rule such that its associated limiting random variable $Z^*_{a,s}$ satisfies*

$$\mathbb{E}[Z^*_{a,s}] = m, \ \ \mathbb{E}[(Z^*_{a,s})^2] \leq v + \eta.$$

It is efficient if we can take $\eta = 0$.

We introduce the set $\mathcal{Z}_T$ of random variables of the form

$$Z_{T,s} = \frac{1}{3}\int_0^T s_t dY_t + \frac{1}{3\sqrt{2}}\int_0^T s_t \sigma_t^Y dB_t, \tag{4.1}$$

where B is a Brownian motion independent of $\mathcal{F}$ and s_t is an adapted continuous process such that

$$\mathbb{E}\Big[\int_0^T \Big(1 + (\rho_t)^2\Big) s_t^2 (\sigma_t^Y)^2 dt\Big] < \infty. \tag{4.2}$$

We also define the notions of non dominated and efficient couples with respect to $\mathcal{Z}_T$. The definitions are the same as Definition 4.1 and Definition 4.2 except that we replace "admissible discretization rule" by "process s satisfying (4.2)" and "its associated limiting random variable $Z^*_{a,s}$" by $Z_{T,s}$.

We can now state our main result which enables us to compute efficient discretization rules.

Theorem 4.1. *The following results hold:*

- *The non domination domain coincides with the non domination domain with respect to $\mathcal{Z}_T$.*
- *Let (m^*, v^*) be an efficient couple with respect to $\mathcal{Z}_T$, with associated optimal process s^*. Then (m^*, v^*) is a nearly efficient couple. More precisely, let $\delta > 0$ and $(\underline{l}_t^\delta, \overline{l}_t^\delta)$ be defined by*

$$\underline{l}_t^\delta - \overline{l}_t^\delta = s_t^*, \quad (\underline{l}_t^\delta)^2 - \underline{l}_t^\delta \overline{l}_t^\delta + (\overline{l}_t^\delta)^2 = (s_t^*)^2 + \frac{6\delta}{(\sigma_t^Y)^2},$$

that is

$$\underline{l}_t^\delta = \sqrt{\frac{(s_t^*)^2}{4} + \frac{6\delta}{(\sigma_t^Y)^2}} + \frac{s_t^*}{2}, \quad \overline{l}_t^\delta = \sqrt{\frac{(s_t^*)^2}{4} + \frac{6\delta}{(\sigma_t^Y)^2}} - \frac{s_t^*}{2}. \tag{4.3}$$

Then the hitting times based discretization rule specified through the barriers $(\underline{l}_t^\delta, \overline{l}_t^\delta)$ satisfies

$$\mathbb{E}[Z^*_{a,s}] = m^*, \quad \mathbb{E}[(Z^*_{a,s})^2] = v^* + \delta T.$$

We have therefore reduced the impulse control problem of finding the optimal rebalancing times to a classical expectation-error optimization with continuous dynamics. The solutions of this problem can be obtained by solving for $\mu > 0$

$$\inf_{(s_t)}\Big\{-\mathbb{E}[Z_{T,s}] + \mu\mathbb{E}[(Z_{T,s})^2]\Big\},$$

for which we can apply the theory of linear-quadratic optimal control, see for example [11, 18]. As shown in the next section, we can even obtain closed formulas in the case where the underlying has deterministic drift and volatility. Note that again, our barriers strategies enable us to attain only nearly efficient couples. Indeed, reaching efficient couples would lead to the use of degenerate barriers with $\delta = 0$. This does not make sense in practice, however δ can of course be selected small.

In practice, once he has chosen the target nearly efficient couple he wants to reach, the trader needs to select δ and ε_n. Two ideas enabling to avoid microstructure effects seem natural and easy to implement:

- Fix a minimal time between two rebalancings t_{min}. After a rebalancing at a random time say τ, wait t_{min} and then apply the strategy with $\delta = 0$ (that is rebalance immediately if at $t = \tau + t_{min}$, $X_t - X_\tau$ is not inside the interval $(-\varepsilon_n \underline{l}_t^0, \varepsilon_n \overline{l}_t^0)$ and wait for the exit time otherwise). The parameter ε_n can be chosen according to the average number of transactions the trader is willing to make.

- Fix (roughly) a minimal distance for the closest barrier after a rebalancing. Then compute δ and ε_n according to the general level of volatility σ_t^Y so that they (approximately) lead to this bound and the average number of transactions the trader is willing to make.

5 One explicit example : Black-Scholes model with time varying coefficients

In this section, we explain how our method can be applied in practice through the simple example of the Black-Scholes model with time varying coefficients. So we assume the underlying follows the dynamics

$$dY_t = Y_t(b_t dt + \sigma_t dW_t),$$

where b_t and σ_t are continuous deterministic functions. We also assume b_t and σ_t do not vanish. Using the theory of linear-quadratic optimal control, we give an explicit solution for the problem of designing optimal rebalancing times in this specific setting.

5.1 Explicit formulas

We aim at finding the efficient couples for the controlled random variables of the form $Z_{T,s}$ as in (4.1). Following [18], such problem is classically recast as follows: solving for any $\mu > 0$ the optimization problem

$$\inf_{(s_t, 0\le t\le T)} -\mathbb{E}[Z_{T,s}] + \mu\mathbb{E}[(Z_{T,s})^2] = \inf_{(s_t, 0\le t\le T)} \mu\mathbb{E}\Big[\Big(Z_{T,s} - \frac{1}{2\mu}\Big)^2\Big] - \frac{1}{4\mu}.$$

Let us define the family of processes of the form

$$d\check{Z}_t = s_t Y_t(\tilde{b}_t dt + \tilde{\sigma}_t dW_t), \;\; \check{Z}_0 = 0,$$

with $\tilde{b}_t = b_t/3$, $\tilde{\sigma}_t = \sigma_t/3$ and s_t adapted continuous. Using obvious computations, the independence between the process B in Equation (4.1) and $\mathcal{F}$, and the fact that s_t is $\mathcal{F}$-adapted, we get $\mathbb{E}[\check{Z}_T] = \mathbb{E}[Z_{T,s}]$ and

$$\mu\mathbb{E}\Big[\Big(\check{Z}_T - \frac{1}{2\mu}\Big)^2\Big] = \mu\mathbb{E}\Big[\Big(Z_{T,s} - \frac{1}{2\mu}\Big)^2\Big] - \frac{\mu}{18}\mathbb{E}\Big[\int_0^T (s_t\sigma_t Y_t)^2 dt\Big].$$

Hence, we can equivalently solve

$$\inf_{(s_t, 0\le t\le T)} \mathbb{E}\Big[\mu\tilde{Z}_T^2 + \frac{\mu}{2}\int_0^T (s_t\tilde{\sigma}_t Y_t)^2 dt\Big],$$

with

$$d\tilde{Z}_t = s_t Y_t(\tilde{b}_t dt + \tilde{\sigma}_t dW_t), \;\; \tilde{Z}_0 = -\frac{1}{2\mu}.$$

Using the results of [18] which are summarized in Theorem B.1 in Appendix B, the optimal control s_t^* and optimally controlled process $\tilde{Z}_t^*$ satisfy

$$s_t^* Y_t = -\frac{1}{\tilde{b}_t}\frac{\dot{P}_t}{P_t}\tilde{Z}_t^*,$$

where P_t is the solution of the (ordinary) differential equation

$$\dot{P}_t = \rho_t^2 \frac{P_t^2}{P_t + \mu}, \;\; P_T = 2\mu,$$

with $\rho_t = b_t/\sigma_t$. The solution of this equation is given by

$$P_t = \frac{\mu}{L\Big(\frac{1}{2}\exp\Big(\int_t^T \rho_s^2 ds + \frac{1}{2}\Big)\Big)},$$

with L is the inverse function of $x \mapsto xe^x$. Moreover, the optimal process $\tilde{Z}^*$ satisfies

$$\frac{d\tilde{Z}_t^*}{\tilde{Z}_t^*} = -\frac{\dot{P}_t}{P_t}(dt + \frac{1}{\rho_t}dW_t), \; \tilde{Z}_0^* = -\frac{1}{2\mu}.$$

Therefore, we obtain

$$\mathbb{E}[\tilde{Z}_T^*] = -\frac{1}{2\mu}\frac{P_0}{P_T}.$$

Using Theorem B.1, we get

$$\mathbb{E}\Big[(\tilde{Z}_T^*)^2 + \frac{1}{2}\int_0^T (s_t^* Y_t \tilde{\sigma}_t)^2 dt\Big] = \Big(\frac{1}{2\mu}\Big)^2 \frac{P_0}{P_T}.$$

Consequently, we have that the optimal variable Z_{T,s^*} satisfies

$$\mathbb{E}[Z_{T,s^*}] = \frac{1}{2\mu}\Big(1 - \frac{P_0}{P_T}\Big)$$

and

$$\mathbb{E}\Big[\Big(Z_{T,s^*} - \frac{1}{2\mu}\Big)^2\Big] = \Big(\frac{1}{2\mu}\Big)^2 \frac{P_0}{P_T}.$$

Hence

$$\mathbb{E}[(Z_{T,s^*})^2] = \Big(\frac{1}{2\mu}\Big)^2 (1 - \frac{P_0}{P_T}).$$

We have thus proved the following proposition.

Proposition 5.1. *In the Black-Scholes model with time varying coefficients, the efficient points are the couples of the form*

$$(m, m^2 \frac{P_T}{P_T - P_0}),$$

with $m > 0$ (remark that the ratio $\frac{P_T}{P_T - P_0}$ does not depend on μ). Furthermore, the associated process s_t^* enabling to compute optimal rules according to Theorem 4.1 is explicitly given by

$$\frac{1}{3} s_t^* Y_t = -\frac{1}{b_t} \frac{\dot{P}_t}{P_t} \tilde{Z}_t^*,$$

with

$$\frac{d\tilde{Z}_t^*}{\tilde{Z}_t^*} = -\frac{1}{b_t} \frac{\dot{P}_t}{P_t} \frac{dY_t}{Y_t}, \ \tilde{Z}_0^* = -\frac{1}{2\mu}.$$

Note that in practice, $\tilde{Z}^*$ is not observable. However, it can of course be approximated by a process $\tilde{Z}^{(*)}$ thanks to historical data, using for example a scheme of the form

$$\tilde{Z}_{t_{i+1}}^{(*)} = \tilde{Z}_{t_i}^{(*)} \Big(1 - \frac{1}{b_{t_i}} \frac{\dot{P}_{t_i}}{P_{t_i}} \frac{Y_{t_{i+1}} - Y_{t_i}}{Y_{t_i}}\Big), \quad \tilde{Z}_0^{(*)} = -\frac{1}{2\mu},$$

where the t_i are the observation times of market data.

Appendix A Proofs

In the following C denotes a constant which may vary from line to line. Note that we use several localization procedures in the proofs. We often give them in details since some of them are slightly unusual, in particular because of the fact that σ^X may vanish at maturity.

A.1 Proof of Proposition 2.1

We start by proving in a very standard way the stable convergence of $\varepsilon_n^{-1}Z^n$ in $C[0,T]$, which is stronger than the weak convergence. More precisely, we show that for any bounded continuous function f on $C[0,T]$ and bounded random variable U defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$,

$$\lim_{n\to\infty} \mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_\cdot)] = \mathbb{E}[Uf(Z^*_\cdot)],$$

where Z^* is defined by

$$Z_t^* = \frac{1}{3}\int_0^t s_u dY_u + \frac{1}{\sqrt{6}}\int_0^t \left(a_u^2 - \frac{2}{3}s_u^2\right)^{1/2} \sigma_u^Y dB_u,$$

on an extension of $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ on which B is a Brownian motion independent of all the other quantities. For $K > 0$, we set

$$\alpha^K = \inf\{t > 0; |\rho_t| \geq K\} \wedge T.$$

Since ρ is continuous on $[0,T]$ almost surely,

$$\lim_{K\to\infty} \mathbb{P}[\alpha^K < T] = 0.$$

Now remark that

$$\begin{aligned}
&|\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_\cdot)] - \mathbb{E}[Uf(Z^*_\cdot)]| \\
&\leq |\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_\cdot)] - \mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K})]| + |\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K})] - \mathbb{E}[Uf(Z^*_{\cdot\wedge\alpha^K})]| \\
&\qquad + |\mathbb{E}[Uf(Z^*_{\cdot\wedge\alpha^K})] - \mathbb{E}[Uf(Z^*_\cdot)]| \\
&\leq 4\|f\|_\infty\|U\|_\infty\mathbb{P}[\alpha^K < T] + |\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K})] - \mathbb{E}[Uf(Z^*_{\cdot\wedge\alpha^K})]|.
\end{aligned}$$

Consequently, it suffices to show that for any $K > 0$,

$$\lim_{n\to\infty} |\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K}) - \mathbb{E}[Uf(Z^*_{\cdot\wedge\alpha^K})]| = 0.$$

Let

$$\mathcal{E} = \exp\Big\{ -\int_0^{\alpha^K} \rho_t dW_t^Y - \frac{1}{2}\int_0^{\alpha^K} \rho_t^2 dt\Big\}.$$

Since $\mathbb{E}[\mathcal{E}] = 1$, the measure $\mathbb{Q}$ defined by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}$$

is a probability measure under which $Z^n_{\cdot\wedge\alpha^K}$ is a local martingale. Under $\mathbb{Q}$, the uniform convergences in probability (2.5) and (2.6) on $[0,T]$ remain true. Therefore by Theorem IX.7.3 of [10], we have the stable convergence of $\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K}$ to $Z^*_{\cdot\wedge\alpha^K}$ under $\mathbb{Q}$. Note that $\tilde{U} = U/\mathcal{E}$ is a $\mathbb{Q}$-integrable positive random variable and moreover, for all $A > 0$,

$$
\begin{aligned}
&|\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K})] - \mathbb{E}[Uf(Z^*_{\cdot\wedge\alpha^K})]| \\
\leq &|\mathbb{E}^{\mathbb{Q}}[\tilde{U}f(\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K})] - \mathbb{E}^{\mathbb{Q}}[\tilde{U}f(Z^*_{\cdot\wedge\alpha^K})]| \\
\leq &|\mathbb{E}^{\mathbb{Q}}[(\tilde{U}\wedge A)f(\varepsilon_n^{-1}Z^n_{\cdot\wedge\alpha^K})] - \mathbb{E}^{\mathbb{Q}}[(\tilde{U}\wedge A)f(Z^*_{\cdot\wedge\alpha^K})]| + 2\|f\|_\infty\mathbb{E}^{\mathbb{Q}}[\tilde{U}1_{\{\tilde{U}\geq A\}}].
\end{aligned}
$$

The second term tends to 0 uniformly in n as $A \to \infty$. The first term converges to 0 due to the stable convergence under $\mathbb{Q}$ since $(\tilde{U} \wedge A)$ is a bounded random variable.

Now, we prove $a^2 \geq s^2$ under the additional condition (2.8). Since a and s are continuous, it suffices to show $a_t^2 \geq s_t^2$ for all $t \in [0,T)$. Fix $T_0 < T$ and let

$$
\hat{\alpha}^K = \inf\{u > 0; |b_u^X| \vee \sigma_u^X \vee \sigma_u^Y \geq K \text{ or } \sigma_u^X \leq 1/K\} \wedge T_0 \tag{A.1}
$$

for $K > 0$. Since σ^X is positive and continuous on $[0, T_0]$, we have

$$
\lim_{K\to\infty} \mathbb{P}[\hat{\alpha}^K < T_0] = 0. \tag{A.2}
$$

Therefore, it suffices to show

$$
a^2_{u\wedge\hat{\alpha}^K} \geq s^2_{u\wedge\hat{\alpha}^K} \tag{A.3}
$$

for all $u \geq 0$ and $K > 0$. Fix K and define the probability measure $\hat{\mathbb{Q}}$ by

$$
\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = \exp\Big\{-\int_0^{\hat{\alpha}^K} \frac{b_u^X}{\sigma_u^X} dW_u^X - \frac{1}{2}\int_0^{\hat{\alpha}^K} \Big(\frac{b_u^X}{\sigma_u^X}\Big)^2 du\Big\}.
$$

Under $\hat{\mathbb{Q}}$, $X_{\cdot\wedge\hat{\alpha}^K}$ is a martingale with bounded quadratic variation. Since $\hat{\mathbb{Q}}$ is equivalent to $\mathbb{P}$, it suffices to show (A.3) under $\hat{\mathbb{Q}}$.

By (2.8), there exists a subsequence $\{n(k)\}$ such that

$$
\hat{\mathbb{Q}}\Big[\varepsilon_{n(k)}^{-4/3} \sup_{j\geq 0}(\tau_{j+1}^{n(k)} \wedge T_0 - \tau_j^{n(k)} \wedge T_0) > \frac{1}{k}\Big] < \frac{1}{k}.
$$

Let

$$
T_k = \inf\Big\{u > 0, \varepsilon_{n(k)}^{-4/3} \sup_{j\geq 0}(\tau_{j+1}^{n(k)} \wedge u - \tau_j^{n(k)} \wedge u) > \frac{1}{k}\Big\} \wedge \hat{\alpha}^K.
$$

Then

$$
\lim_{k\to\infty} \hat{\mathbb{Q}}[T_k < \hat{\alpha}^K] = 0
$$

and so,

$$
\begin{aligned}
&\varepsilon_{n(k)}^{-1}\left\langle Z^{n(k)}, Y\right\rangle_{t\wedge T_k} \to \frac{1}{3}\int_0^{t\wedge\hat{\alpha}^K} s_u(\sigma_u^Y)^2 du, \\
&\varepsilon_{n(k)}^{-2}\left\langle Z^{n(k)}\right\rangle_{t\wedge T_k} \to \frac{1}{6}\int_0^{t\wedge\hat{\alpha}^K} a_u^2(\sigma_u^Y)^2 du,
\end{aligned} \tag{A.4}
$$

in probability as $k \to \infty$ for all $t \geq 0$. Let

$$\hat{\tau}_j^k = \tau_j^{n(k)} \wedge T_k$$

for $j \geq 0$. We now give three technical lemmas.

Lemma A.1. *Let $\kappa_u = (\sigma_u^Y / \sigma_u^X)^2$. We have*

$$\frac{1}{3}\varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t\wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^3 \Big| \mathcal{F}_{\hat{\tau}_j^k \wedge t}\Big] - \varepsilon_{n(k)}^{-1} \left\langle Z^{n(k)}, Y \right\rangle_{t\wedge T_k} \to 0,$$

in probability as $k \to \infty$ for all $t \geq 0$.

Proof. By Itô's formula,

$$\frac{1}{3}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^3 \Big| \mathcal{F}_{\hat{\tau}_j^k \wedge t}\Big] = \mathbb{E}^{\hat{\mathbb{Q}}}\Big[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\left\langle X \right\rangle_u \Big| \mathcal{F}_{\hat{\tau}_j^k \wedge t}\Big].$$

We now show that

$$\begin{aligned} &\varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t\wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}}\Big[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\left\langle X \right\rangle_u \Big| \mathcal{F}_{\hat{\tau}_j^k \wedge t}\Big] \\ &- \varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t\wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\left\langle X \right\rangle_u \to 0 \end{aligned} \tag{A.5}$$

and

$$\varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t\wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\left\langle X \right\rangle_u - \varepsilon_{n(k)}^{-1} \left\langle Z^{n(k)}, Y \right\rangle_{t\wedge T_k} \to 0, \tag{A.6}$$

in probability.

By Lenglart inequality for discrete martingales (see e.g., Lemma A.2 of [4]), a sufficient condition for (A.5) is the fact that

$$\varepsilon_{n(k)}^{-2} \sum_{j=0}^{N_{t\wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t}^2 \mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\left\langle X \right\rangle_u\Big)^2 \Big| \mathcal{F}_{\hat{\tau}_j^k \wedge t}\Big] \to 0, \tag{A.7}$$

in probability. To get this convergence, first use successively Hölder inequality, Itô's formula and Burkholder-Davis-Gundy inequality to obtain that

$$\begin{aligned}
&\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}} \kappa^2_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u - X_{\hat\tau^k_j})d\langle X\rangle_u\Big)^2\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\\
&\le C\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}} \kappa^2_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t} - \langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^{3/2}\Big(\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u - X_{\hat\tau^k_j})^4 d\langle X\rangle_u\Big)^{1/2}\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\\
&\le C\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t} - \langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^3\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{1/2}\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u - X_{\hat\tau^k_j})^4 d\langle X\rangle_u\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{1/2}\\
&= C\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t} - \langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^3\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{1/2}\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat\tau^k_{j+1}\wedge t} - X_{\hat\tau^k_j\wedge t})^6|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{1/2}\\
&\le C\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t} - \langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^3\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\\
&\le C\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[|\hat\tau^k_{j+1}\wedge t - \hat\tau^k_j\wedge t|^3\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big].
\end{aligned}$$

Note also that

$$\Big\{j \le N^{n(k)}_{t\wedge T_k}\Big\} = \Big\{\tau^{n(k)}_j \le t\wedge T_k\Big\} \in \mathcal{F}_{\hat\tau^k_j\wedge t}.$$

Then (A.7) follows since

$$\varepsilon^{-2}_{n(k)}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\sum_{j=0}^{\infty}|\hat\tau^k_{j+1}\wedge t - \hat\tau^k_j\wedge t|^3\Big] \le \varepsilon^{-2}_{n(k)}\frac{\varepsilon^{8/3}_{n(k)}}{k^2}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\sum_{j=0}^{\infty}|\hat\tau^k_{j+1}\wedge t - \hat\tau^k_j\wedge t|\Big] \le \frac{\varepsilon^{2/3}_{n(k)}}{k^2}t \to 0.$$

We now turn to (A.6). Note that

$$\varepsilon^{-1}_{n(k)}\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\kappa_{\hat\tau^k_j\wedge t}\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u - X_{\hat\tau^k_j})d\left\langle X\right\rangle_u = \varepsilon^{-1}_{n(k)}\sum_{j=0}^{\infty}\kappa_{\hat\tau^k_j\wedge t}\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u - X_{\hat\tau^k_j})d\left\langle X\right\rangle_u$$

and

$$\varepsilon^{-1}_{n(k)}\left\langle Z^{n(k)}, Y\right\rangle_{t\wedge T_k} = \varepsilon^{-1}_{n(k)}\sum_{j=0}^{\infty}\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u - X_{\hat\tau^k_j})\kappa_u d\left\langle X\right\rangle_u .$$

Therefore, the absolute value of the left hand side of (A.6) is dominated by

$$
\begin{aligned}
&\varepsilon_{n(k)}^{-1}\sum_{j=0}^{\infty}\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}|X_u - X_{\hat{\tau}_j^k}|\,|\kappa_u - \kappa_{\hat{\tau}_j^k\wedge t}|d\langle X\rangle_u\\
&\leq \Big(\varepsilon_{n(k)}^{-2}\sum_{j=0}^{\infty}\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}|X_u - X_{\hat{\tau}_j^k}|^2\,|\kappa_u - \kappa_{\hat{\tau}_j^k\wedge t}|^2 d\langle X\rangle_u\Big)^{1/2}\Big(\sum_{j=0}^{\infty}\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}d\langle X\rangle_u\Big)^{1/2}\\
&\leq \sup_{u\geq 0}\frac{1}{\kappa_{u\wedge T_k}}\sup_{u\geq 0, j\geq 0}\Big|\kappa_{\hat{\tau}_{j+1}^k\wedge u} - \kappa_{\hat{\tau}_j^k\wedge u}\Big|\varepsilon_{n(k)}^{-1}\left\langle Z^{n(k)}\right\rangle_{t\wedge T_k}^{1/2}\langle X\rangle_{t\wedge T_k}^{1/2},
\end{aligned}
$$

which converges to 0 due to (A.4) and the uniform continuity of κ. □

Lemma A.2. *We have*

$$
\frac{1}{6}\varepsilon_{n(k)}^{-2}\sum_{j=0}^{N_{t\wedge T_k}^{n(k)}}\kappa_{\hat{\tau}_j^k\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(X_{\hat{\tau}_{j+1}^k\wedge t} - X_{\hat{\tau}_j^k\wedge t}\Big)^4\Big|\mathcal{F}_{\hat{\tau}_j^k\wedge t}\Big] - \varepsilon_{n(k)}^{-2}\left\langle Z^{n(k)}\right\rangle_{t\wedge T_k}\to 0,
$$

in probability as $k\to\infty$, for all $t\geq 0$.

Proof. The proof is very similar to the previous one. By Itô's formula,

$$
\frac{1}{6}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(X_{\hat{\tau}_{j+1}^k\wedge t} - X_{\hat{\tau}_j^k\wedge t}\Big)^4\Big|\mathcal{F}_{\hat{\tau}_j^k\wedge t}\Big] = \mathbb{E}^{\hat{\mathbb{Q}}}\Big[\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}(X_u - X_{\hat{\tau}_j^k})^2 d\langle X\rangle_u\Big|\mathcal{F}_{\hat{\tau}_j^k\wedge t}\Big].
$$

We now show that

$$
\begin{aligned}
&\varepsilon_{n(k)}^{-2}\sum_{j=0}^{N_{t\wedge T_k}^{n(k)}}\kappa_{\hat{\tau}_j^k\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}(X_u - X_{\hat{\tau}_j^k})^2 d\langle X\rangle_u\Big|\mathcal{F}_{\hat{\tau}_j^k\wedge t}\Big]\\
&-\varepsilon_{n(k)}^{-2}\sum_{j=0}^{N_{t\wedge T_k}^{n(k)}}\kappa_{\hat{\tau}_j^k\wedge t}\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}(X_u - X_{\hat{\tau}_j^k})^2 d\langle X\rangle_u\to 0
\end{aligned}
\tag{A.8}
$$

and

$$
\varepsilon_{n(k)}^{-2}\sum_{j=0}^{N_{t\wedge T_k}^{n(k)}}\kappa_{\hat{\tau}_j^k\wedge t}\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}(X_u - X_{\hat{\tau}_j^k})^2 d\langle X\rangle_u - \varepsilon_{n(k)}^{-2}\left\langle Z^{n(k)}\right\rangle_{t\wedge T_k}\to 0, \tag{A.9}
$$

in probability.

By Lenglart inequality for discrete martingales, a sufficient condition for (A.8) is

$$
\varepsilon_{n(k)}^{-4}\sum_{j=0}^{N_{t\wedge T_k}^{n(k)}}\kappa_{\hat{\tau}_j^k\wedge t}^2\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\int_{\hat{\tau}_j^k\wedge t}^{\hat{\tau}_{j+1}^k\wedge t}(X_u - X_{\hat{\tau}_j^k})^2 d\langle X\rangle_u\Big)^2\Big|\mathcal{F}_{\hat{\tau}_j^k\wedge t}\Big]\to 0, \tag{A.10}
$$

in probability. To get this convergence, first use successively Hölder inequality, Itô's formula and Burkholder-Davis-Gundy inequality to obtain that

$$
\begin{aligned}
&\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}} \kappa^2_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u-X_{\hat\tau^k_j})^2 d\langle X\rangle_u\Big)^2\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\\
&\le \sum_{j=0}^{N^{n(k)}_{t\wedge T_k}} \kappa^2_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t}-\langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^{4/3}\Big(\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u-X_{\hat\tau^k_j})^6 d\langle X\rangle_u\Big)^{2/3}\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\\
&\le C\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t}-\langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^4\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{1/3}\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\int_{\hat\tau^k_j\wedge t}^{\hat\tau^k_{j+1}\wedge t}(X_u-X_{\hat\tau^k_j})^6 d\langle X\rangle_u\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{2/3}\\
&= C\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t}-\langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^4\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{1/3}\Big(\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t})^8|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^{2/3}\\
&\le C\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(\langle X\rangle_{\hat\tau^k_{j+1}\wedge t}-\langle X\rangle_{\hat\tau^k_j\wedge t}\Big)^4\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\\
&\le C\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[|\hat\tau^k_{j+1}\wedge t-\hat\tau^k_j\wedge t|^4\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big].
\end{aligned}
$$

Then, observe that

$$\varepsilon^{-4}_{n(k)}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\sum_{j=0}^{\infty}|\hat\tau^k_{j+1}\wedge t-\hat\tau^k_j\wedge t|^4\Big]\le\frac{t}{k^3}\to 0,$$

which gives (A.8). The proof for (A.9) is obtained in the same way as that for (A.6). □

We finally give the following almost straightforward result, which is easily deduced from simplified versions of the proofs of the previous lemma.

Lemma A.3. *We have*

$$\sum_{j=0}^{N^{n(k)}_{t\wedge T_k}}\kappa_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t}\Big)^2\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]-\langle Y\rangle_{t\wedge T_k}\to 0,$$

in probability as $k\to\infty$ for all $t\ge 0$.

We are now ready to complete the proof of Proposition 2.1. From (A.4) and Lemmas A.1, A.2 and A.3, we have for all $0 \le v \le t$ the following convergences in probability as $k \to \infty$:

$$\sum_{j=N^{n(k)}_{v\wedge T_k}+1}^{N^{n(k)}_{t\wedge T_k}} \kappa_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t}\Big)^2\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big] \to \int_{v\wedge\hat\alpha^K}^{t\wedge\hat\alpha^K}(\sigma^Y_y)^2du,$$

$$\varepsilon^{-1}_{n(k)}\sum_{j=N^{n(k)}_{v\wedge T_k}+1}^{N^{n(k)}_{t\wedge T_k}} \kappa_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t}\Big)^3\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big] \to \int_{v\wedge\hat\alpha^K}^{t\wedge\hat\alpha^K}s_u(\sigma^Y_y)^2du,$$

$$\varepsilon^{-2}_{n(k)}\sum_{j=N^{n(k)}_{v\wedge T_k}+1}^{N^{n(k)}_{t\wedge T_k}} \kappa_{\hat\tau^k_j\wedge t}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[\Big(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t}\Big)^4\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big] \to \int_{v\wedge\hat\alpha^K}^{t\wedge\hat\alpha^K}a_u^2(\sigma^Y_y)^2du.$$

Since

$$\Big(\mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t})^3\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^2 \le \mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t})^2\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t})^4\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big],$$

we have

$$\begin{aligned}&\Big(\varepsilon^{-1}_{n(k)}\sum_{j=N^{n(k)}_{v\wedge T_k}+1}^{N^{n(k)}_{t\wedge T_k}} \kappa_{\hat\tau^k_j}\mathbb{E}^{\hat{\mathbb{Q}}}\Big[(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t})^3\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\Big]\Big)^2\\ &\le \varepsilon^{-2}_{n(k)}\sum_{j=N^{n(k)}_{v\wedge T_k}+1}^{N^{n(k)}_{t\wedge T_k}} \kappa_{\hat\tau^k_j}\mathbb{E}^{\hat{\mathbb{Q}}}\left[(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t})^4\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\right]\sum_{j=N^{n(k)}_{v\wedge T_k}+1}^{N^{n(k)}_{t\wedge T_k}} \kappa_{\hat\tau^k_j}\mathbb{E}^{\hat{\mathbb{Q}}}\left[(X_{\hat\tau^k_{j+1}\wedge t}-X_{\hat\tau^k_j\wedge t})^2\Big|\mathcal{F}_{\hat\tau^k_j\wedge t}\right].\end{aligned}$$

This implies that for all $0 \le v \le t$,

$$\Big(\int_{v\wedge\hat\alpha^K}^{t\wedge\hat\alpha^K}s_u(\sigma^Y_u)^2du\Big)^2 \le \int_{v\wedge\hat\alpha^K}^{t\wedge\hat\alpha^K}a_u^2(\sigma^Y_u)^2du\int_{v\wedge\hat\alpha^K}^{t\wedge\hat\alpha^K}(\sigma^Y_u)^2du.$$

Thus we obtain (A.3).

A.2 Proof of Proposition 2.2

In this proof, using a classical localization procedure together with Girsanov theorem, we can assume that $b^X = 0$ and that σ^X and σ^Y are bounded on $[0, T]$. We start with two technical lemmas and their proof.

Lemma A.4. *For any $p \in [0, 2)$,*

$$\varepsilon_n^{-p}\sup_{j\ge 0}(\langle X\rangle_{\tau^n_{j+1}}-\langle X\rangle_{\tau^n_j}) \to 0,$$

in probability.

Proof. Let $K > 0$ and

$$\gamma_K^n = \inf\{t > 0; \varepsilon_n^{-1}|X_t - X_t^n| \geq K\} \wedge T. \tag{A.11}$$

Using the tightness of the family (2.9), we get

$$\lim_{K\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}[\gamma_K^n < T] = 0. \tag{A.12}$$

Therefore, it is enough to show that for any $K > 0$,

$$\varepsilon_n^{-p} \sup_{j\geq 0}(\langle X\rangle_{\tau_{j+1}^n\wedge\gamma_K^n} - \langle X\rangle_{\tau_j^n\wedge\gamma_K^n}) \to 0,$$

in probability. Take an integer $m > 2/(2-p)$. Since

$$\sup_{j\geq 0}(\langle X\rangle_{\tau_{j+1}^n\wedge\gamma_K^n} - \langle X\rangle_{\tau_j^n\wedge\gamma_K^n}) \leq \Big(\sum_{j=0}^{\infty}(\langle X\rangle_{\tau_{j+1}^n\wedge\gamma_K^n} - \langle X\rangle_{\tau_j^n\wedge\gamma_K^n})^m\Big)^{1/m},$$

the statement of the lemma follows from the fact that

$$\begin{aligned}
\mathbb{E}\Big[\sum_{j=0}^{\infty}(\langle X\rangle_{\tau_{j+1}^n\wedge\gamma_K^n} - \langle X\rangle_{\tau_j^n\wedge\gamma_K^n})^m\Big] &\leq C\mathbb{E}\Big[\sum_{j=0}^{\infty}\Big(\sup_{t\geq 0}|X_{\tau_{j+1}^n\wedge\gamma_K^n\wedge t} - X_{\tau_j^n\wedge\gamma_K^n\wedge t}|\Big)^{2m}\Big] \\
&\leq C\varepsilon_n^{2m-2}\mathbb{E}\Big[\sum_{j=0}^{\infty}\Big(\sup_{t\geq 0}|X_{\tau_{j+1}^n\wedge\gamma_K^n\wedge t} - X_{\tau_j^n\wedge\gamma_K^n\wedge t}|\Big)^{2}\Big] \\
&\leq C\varepsilon_n^{2m-2}\mathbb{E}\Big[\sum_{j=0}^{\infty}(\langle X\rangle_{\tau_{j+1}^n\wedge\gamma_K^n} - \langle X\rangle_{\tau_j^n\wedge\gamma_K^n})\Big] \\
&\leq C\varepsilon_n^{2m-2}.
\end{aligned}$$

Here we have used that $\mathbb{E}[\langle X\rangle_T] < \infty$. The result follows using Hölder inequality. □

Lemma A.5. *For any $p \in [0,2)$ and $T_0 \in [0,T)$,*

$$\varepsilon_n^{-p}\sup_{j\geq 0}(\tau_{j+1}^n \wedge T_0 - \tau_j^n \wedge T_0) \to 0,$$

in probability. In particular, the convergence in probability (2.8) *holds for all $T_0 \in [0,T)$.*

Proof. Let $T_0 \in [0,T)$, $K > 0$ and

$$\hat{\gamma}_K = \inf\{t > 0; \sigma_t^X \leq 1/K\} \wedge T_0. \tag{A.13}$$

Using the continuity and the positivity of σ^X on $[0,T)$, we get

$$\lim_{K\to\infty} \mathbb{P}[\hat{\gamma}_K < T_0] = 0.$$

Therefore, it is enough to show that for any $K > 0$,

$$\varepsilon_n^{-p}\sup_{j\geq 0}(\tau_{j+1}^n \wedge \hat{\gamma}_K - \tau_j^n \wedge \hat{\gamma}_K) \to 0,$$

in probability. This follows from Lemma A.4 since

$$\sup_{j\geq 0}(\tau_{j+1}^n \wedge \hat{\gamma}_K - \tau_j^n \wedge \hat{\gamma}_K) \leq C\sup_{j\geq 0}(\langle X\rangle_{\tau_{j+1}^n} - \langle X\rangle_{\tau_j^n}).$$

□

We now give the end of the proof of Proposition 2.2. Define γ_K^n by (A.11). It suffices to show that for any $K > 0$

$$\sup_{t\geq 0}\left|\varepsilon_n^{-2}\langle Z^n\rangle_{t\wedge\gamma_K^n} - \frac{1}{6}\int_0^{t\wedge\gamma_K^n} a_u^2(\sigma_u^Y)^2du\right| \to 0,$$
$$\sup_{t\geq 0}\left|\varepsilon_n^{-1}\langle Z^n, Y\rangle_{t\wedge\gamma_K^n} - \frac{1}{3}\int_0^{t\wedge\gamma_K^n} s_u(\sigma_u^Y)^2du\right| \to 0,$$

in probability. Since

$$\varepsilon_n^{-1}\sup_{t\geq 0}|X_{t\wedge\gamma_K^n} - X^n_{t\wedge\gamma_K^n}| \leq K, \tag{A.14}$$

the families $\varepsilon_n^{-2}\langle Z^n\rangle_{\cdot\wedge\gamma_K^n}$ and $\varepsilon_n^{-1}\langle Z^n, Y\rangle_{\cdot\wedge\gamma_K^n}$ are equicontinuous. So we just need to prove that for any $t \in [0, T)$,

$$\varepsilon_n^{-2}\langle Z^n\rangle_{t\wedge\gamma_K^n} - \frac{1}{6}\int_0^{t\wedge\gamma_K^n} a_u^2(\sigma_u^Y)^2du \to 0,$$
$$\varepsilon_n^{-1}\langle Z^n, Y\rangle_{t\wedge\gamma_K^n} - \frac{1}{3}\int_0^{t\wedge\gamma_K^n} s_u(\sigma_u^Y)^2du \to 0,$$

in probability. Let

$$\beta_M^n = \inf\{u > 0; (1/\sigma_u^X) \geq M\} \wedge t \wedge \gamma_K^n$$

for $M > 0$. Since $t < T$, Lemma A.4 gives that

$$\lim_{M\to\infty}\sup_{n\in\mathbb{N}}\mathbb{P}[\beta_M^n < t\wedge\gamma_K^n] = 0.$$

Therefore it is enough to show that for any $M > 0$,

$$\varepsilon_n^{-2}\langle Z^n\rangle_{\beta_M^n} - \frac{1}{6}\int_0^{\beta_M^n} a_u^2(\sigma_u^Y)^2du \to 0,$$
$$\varepsilon_n^{-1}\langle Z^n, Y\rangle_{\beta_M^n} - \frac{1}{3}\int_0^{\beta_M^n} s_u(\sigma_u^Y)^2du \to 0,$$

in probability. From the assumptions of Proposition 2.2, we have

$$\varepsilon_n^{-1}\sum_{j=0}^{N^n_{\beta_M^n}}\kappa_{\tau_j^n}\mathbb{E}\left[\Delta_{j,n}^3\middle|\mathcal{F}_{\tau_j^n}\right] + \int_0^{\beta_M^n} s_u(\sigma_u^Y)^2du \to 0,$$
$$\varepsilon_n^{-2}\sum_{j=0}^{N^n_{\beta_M^n}}\kappa_{\tau_j^n}\mathbb{E}\left[\Delta_{j,n}^4\middle|\mathcal{F}_{\tau_j^n}\right] - \int_0^{\beta_M^n} a_u^2(\sigma_u^Y)^2du \to 0,$$

in probability. Moreover, by Itô's formula,

$$\frac{1}{3}\varepsilon_n^{-1}\sum_{j=0}^{N^n_{\beta_M^n}}\kappa_{\tau_j^n}\mathbb{E}\left[\Delta_{j,n}^3\middle|\mathcal{F}_{\tau_j^n}\right] = \varepsilon_n^{-1}\sum_{j=0}^{N^n_{\beta_M^n}}\kappa_{\tau_j^n}\mathbb{E}\Big[\int_{\tau_j^n}^{\tau_{j+1}^n}(X_u - X^n_{\tau_j^n})d\langle X\rangle_u\,\Big|\mathcal{F}_{\tau_j^n}\Big],$$
$$\frac{1}{6}\varepsilon_n^{-2}\sum_{j=0}^{N^n_{\beta_M^n}}\kappa_{\tau_j^n}\mathbb{E}\left[\Delta_{j,n}^4\middle|\mathcal{F}_{\tau_j^n}\right] = \varepsilon_n^{-2}\sum_{j=0}^{N^n_{t\wedge\beta^K}}\kappa_{\tau_j^n}\mathbb{E}\Big[\int_{\tau_j^n}^{\tau_{j+1}^n}(X_u - X^n_{\tau_j^n})^2d\langle X\rangle_u\,\Big|\mathcal{F}_{\tau_j^n}\Big].$$

Now remark that the following convergences in probability hold:

$$\sup_{t\geq 0}\Big|\varepsilon_n^{-1}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa_{\tau^n_j}\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})d\langle X\rangle_u-\varepsilon_n^{-1}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa_{\tau^n_j}\mathbb{E}\Big[\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})d\langle X\rangle_u\,\Big|\mathcal{F}_{\tau^n_j}\Big]\Big|\to 0,$$
$$\sup_{t\geq 0}\Big|\varepsilon_n^{-2}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa_{\tau^n_j}\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})^2d\langle X\rangle_u-\varepsilon_n^{-2}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa_{\tau^n_j}\mathbb{E}\Big[\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})^2d\langle X\rangle_u\,\Big|\mathcal{F}_{\tau^n_j}\Big]\Big|\to 0. \tag{A.15}$$

Indeed, as seen in the proofs of Lemmas A.1 and A.2, the convergences in probability in (A.15) are deduced from the following ones:

$$\begin{aligned}&\varepsilon_n^{-2}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa^2_{\tau^n_j}\mathbb{E}\Big[\Big(\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})d\langle X\rangle_u\Big)^2\Big|\mathcal{F}_{\tau^n_j}\Big]\to 0,\\ &\varepsilon_n^{-4}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa^2_{\tau^n_j}\mathbb{E}\Big[\Big(\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})^2d\langle X\rangle_u\Big)^2\Big|\mathcal{F}_{\tau^n_j}\Big]\to 0.\end{aligned} \tag{A.16}$$

Since $Q_n=\varepsilon_n^{-4}\sup_{t\in[0,T]}|X^n_t-X_t|^4$ is uniformly integrable and

$$\sum_{j=0}^{\infty}(\langle X\rangle_{\tau^n_{j+1}}-\langle X\rangle_{\tau^n_j})^2$$

is bounded and converges to 0 in probability by Lemma A.4, we have

$$\mathbb{E}\Big[\sum_{j=0}^{N^n_{\beta^n_M}}\kappa^2_{\tau^n_j}\Big(\int_{\tau^n_j}^{\tau^n_{j+1}}\varepsilon_n^{-k}(X_u-X^n_{\tau^n_j})^kd\langle X\rangle_u\Big)^2\Big]\leq C\mathbb{E}\Big[Q_n^{k/2}\sum_{j=0}^{\infty}(\langle X\rangle_{\tau^n_{j+1}}-\langle X\rangle_{\tau^n_j})^2\Big]\to 0$$

for $k=1,2$, which gives (A.16).

We also have

$$\sup_{t\geq 0}\Big|\varepsilon_n^{-1}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa_{\tau^n_j}\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})d\langle X\rangle_u-\varepsilon_n^{-1}\sum_{j=0}^{\infty}\kappa_{\tau^n_j}\int_{\tau^n_j\wedge\beta^n_M}^{\tau^n_{j+1}\wedge\beta^n_M}(X_u-X^n_{\tau^n_j})d\langle X\rangle_u\Big|\to 0,$$
$$\sup_{t\geq 0}\Big|\varepsilon_n^{-2}\sum_{j=0}^{N^n_{\beta^n_M}}\kappa_{\tau^n_j}\int_{\tau^n_j}^{\tau^n_{j+1}}(X_u-X^n_{\tau^n_j})^2d\langle X\rangle_u-\varepsilon_n^{-2}\sum_{j=0}^{\infty}\kappa_{\tau^n_j}\int_{\tau^n_j\wedge\beta^n_M}^{\tau^n_{j+1}\wedge\beta^n_M}(X_u-X^n_{\tau^n_j})^2d\langle X\rangle_u\Big|\to 0, \tag{A.17}$$

in probability. These two convergences follow using that

$$\sup_{j\geq 0, t\in[0,T]} \int_{\tau_j^n\wedge t}^{\tau_{j+1}^n\wedge t} \varepsilon_n^{-i}|X_u - X_u^n|^i d\langle X\rangle_u \to 0$$

in probability for $i = 1, 2$, which is deduced from (A.12) and the fact that

$$\sup_{j\geq 0, t\geq 0} \int_{\tau_j^n\wedge\gamma_K^n\wedge t}^{\tau_{j+1}^n\wedge\gamma_K^n\wedge t} \varepsilon_n^{-i}|X_u - X_u^n|^i d\langle X\rangle_u \leq K^i \sup_{j\geq 0}(\langle X\rangle_{\tau_{j+1}^n} - \langle X\rangle_{\tau_j^n}) \to 0,$$

in probability, by Lemma A.4.

Finally, remark that the uniform continuity of κ and (A.14) imply

$$\begin{aligned}
&\sup_{t\geq 0}\left|\varepsilon_n^{-1}\sum_{j=0}^{\infty}\kappa_{\tau_j^n}\int_{\tau_j^n\wedge\beta_M^n}^{\tau_{j+1}^n\wedge\beta_M^n}(X_u - X_{\tau_j^n}^n)d\langle X\rangle_u + \varepsilon_n^{-1}\langle Z^n, Y\rangle_{\beta_M^n}\right| \to 0,\\
&\sup_{t\geq 0}\left|\varepsilon_n^{-2}\sum_{j=0}^{\infty}\kappa_{\tau_j^n}\int_{\tau_j^n\wedge\beta_M^n}^{\tau_{j+1}^n\wedge\beta_M^n}(X_u - X_{\tau_j^n}^n)^2 d\langle X\rangle_u - \varepsilon_n^{-2}\langle Z^n\rangle_{\beta_M^n}\right| \to 0,
\end{aligned} \tag{A.18}$$

in probability. Then Proposition 2.2 is eventually obtained from (A.15) together with (A.17) and (A.18).

A.3 Proof of Proposition 2.3

A.3.1 Proof of the convergence in law to (2.3)

We start with the stable convergence in law of the renormalized hedging error. Such convergence being stable against localization procedures, we can assume without loss of generality that $|b^X|$, σ^X, $|b^Y|$, σ^Y, $1/\sigma^Y$, $\bar{l}$, $1/\bar{l}$, $\underline{l}$ and $1/\underline{l}$ are bounded by a constant $K > 0$. Then in particular we have $\varepsilon_n^{-1}\sup_{t\in[0,T]}|X_t^n - X_t| \leq K$.

By Lemma A.5, we have (2.8) for all $T_0 \in [0, T)$. Furthermore $\varepsilon_n^{-1}\langle Z^n, Y\rangle$ and $\varepsilon_n^{-2}\langle Z^n\rangle$ are equicontinuous. Therefore the uniform convergences in probability (2.5) and (2.6) follow from the corresponding convergences in probability at each $t \in [0, T)$.

Fix $T_0 \in [0, T)$ and define $\hat{\alpha}^K$ by (A.1). Then we have (A.2) and so, we can assume without loss of generality that $1/\sigma^X \leq K$ in order to show the convergences (2.5) and (2.6) on $[0, T_0]$. Also, thanks to the Girsanov-Maruyama transformation, we can assume $b^X = 0$. Define for $\delta > 0$ and $t \in [0, T_0]$

$$w_t(\delta) = \sup\{|\bar{l}_u - \bar{l}_v| + |\underline{l}_u - \underline{l}_v|; 0 \leq u \leq t,\ 0 \leq v \leq t,\ |u - v| \leq \delta\}.$$

Since $\bar{l}$ and $\underline{l}$ are continuous and bounded, we have

$$\mathbb{E}[w_{T_0}(\delta)] \to 0$$

as $\delta \to 0$. Let

$$T_K^n = \inf\{t > 0; w_t(\varepsilon_n) \geq K\mathbb{E}[w_{T_0}(\varepsilon_n)]\} \wedge T_0.$$

Note that

$$\sup_{n\in\mathbb{N}} \mathbb{P}[T_K^n < T_0] \leq \sup_{n\in\mathbb{N}} \mathbb{P}[w_{T_0}(\varepsilon_n) \geq K\mathbb{E}[w_{T_0}(\varepsilon_n)]] \leq \frac{1}{K} \to 0,$$

as $K \to \infty$. On the set $\{T_K^n < T_0\}$, we can replace $\bar{l}$ and $\underline{l}$ by $\bar{l}_{\cdot\wedge T_K^n}$ and $\underline{l}_{\cdot\wedge T_K^n}$ respectively. This means that we can assume without loss of generality that $w_{T_0}(\varepsilon_n) \leq K\mathbb{E}[w_{T_0}(\varepsilon_n)]$. Now in order to apply Proposition 2.2, it remains to show (2.10).

Part 1: Technical lemma

We give here a first technical lemma.

Lemma A.6. *The sequence $\varepsilon_n^2 N_{T_0}^n$ is tight.*

Proof. Since

$$|X_{\tau_{j+1}^n\wedge T_0} - X_{\tau_j^n\wedge T_0}|^2 \geq \frac{\varepsilon_n^2}{K^2},$$

we have

$$\varepsilon_n^2 N_{T_0}^n \leq K^2 \sum_{j=0}^{N_{T_0}^n} (X_{\tau_{j+1}^n\wedge T_0} - X_{\tau_j^n\wedge T_0})^2 \to K^2 \langle X\rangle_{T_0},$$

in probability by Lemma A.5. □

Part 2: Approximation lemma

We give here an important result. Let $\tilde{\tau}_{j+1}^n$ be the exit time of fixed barriers defined by

$$\tilde{\tau}_{j+1}^n = \inf\left\{t > \tau_j^n : X_t \notin (X_{\tau_j^n} - \varepsilon_n \underline{l}_{\tau_j^n}, X_{\tau_j^n} + \varepsilon_n \bar{l}_{\tau_j^n})\right\} \wedge T_0. \tag{A.19}$$

We have the following lemma.

Lemma A.7. *We have*

$$\sum_{j=0}^{N_{T_0}^n} \mathbb{E}\left[(\tilde{\tau}_{j+1}^n - \tau_{j+1}^n)|\mathcal{F}_{\tau_j^n}\right] \to 0,$$

in probability.

Proof. Since the sequence $\varepsilon_n^2 N_{T_0}^n$ is tight, it is enough to show that

$$\frac{1}{\varepsilon_n^2} \sup_{j\leq N_{T_0}^n} \mathbb{E}\left[\tilde{\tau}_{j+1}^n - \tau_{j+1}^n \middle| \mathcal{F}_{\tau_j^n}\right] \to 0.$$

We write

$$\frac{1}{\varepsilon_n^2} \sup_{j\leq N_{T_0}^n} \mathbb{E}\left[\tilde{\tau}_{j+1}^n - \tau_{j+1}^n \middle| \mathcal{F}_{\tau_j^n}\right] = R_1 + R_2,$$

with

$$R_1 = \frac{1}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{E}\Big[(\tilde{\tau}_{j+1}^n - \tau_{j+1}^n) 1_{\{\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n \geq \tau_j^n + \varepsilon_n\}} \Big| \mathcal{F}_{\tau_j^n}\Big],$$

$$R_2 = \frac{1}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{E}\Big[(\tilde{\tau}_{j+1}^n - \tau_{j+1}^n) 1_{\{\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n < \tau_j^n + \varepsilon_n\}} \Big| \mathcal{F}_{\tau_j^n}\Big].$$

We first treat R_1. We have

$$R_1 \leq \frac{T}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{P}\Big[\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n \geq \tau_j^n + \varepsilon_n \Big| \mathcal{F}_{\tau_j^n}\Big].$$

Since $\bar{l}$, $\underline{l}$ and σ^X are bounded from below by $1/K$, using the Dambis, Dubins-Schwartz theorem we get that there exists some $C > 0$ such that

$$\mathbb{P}\Big[\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n \geq \tau_j^n + \varepsilon_n \Big| \mathcal{F}_{\tau_j^n}\Big] \leq \mathbb{P}[\rho^n \geq C\varepsilon_n],$$

with ρ^n the first exit time of $[-\varepsilon_n/K, \varepsilon_n/K]$ by a Brownian motion starting from zero. Using the well-known bound $\mathbb{E}[(\rho^n)^k] \leq C\varepsilon_n^{2k}$ for $k \in \mathbb{N}$, Markov's inequality gives the convergence to zero of R_1.

We now turn to R_2. Recall that $w_{T_0}(\varepsilon_n) \leq K\mathbb{E}[w_{T_0}(\varepsilon_n)] = \delta_n \to 0$. Then, we have

$$(\tilde{\tau}_{j+1}^n - \tau_{j+1}^n) 1_{\{\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n < \tau_j^n + \varepsilon_n\}} \leq \hat{J}_{j+1}^n - \check{J}_{j+1}^n,$$

with

$$\hat{J}_{j+1}^n = \inf\Big\{t \geq \tau_j^n ; X_{\tau_j^n + t} - X_{\tau_j^n} \notin \Big(-\varepsilon_n(\underline{l}_{\tau_j^n} + \delta_n), \varepsilon_n(\bar{l}_{\tau_j^n} + \delta_n)\Big)\Big\},$$

$$\check{J}_{j+1}^n = \inf\Big\{t \geq \tau_j^n ; X_{\tau_j^n + t} - X_{\tau_j^n} \notin \Big(-\varepsilon_n(\underline{l}_{\tau_j^n} - \delta_n), \varepsilon_n(\bar{l}_{\tau_j^n} - \delta_n)\Big)\Big\}.$$

Using again the Dambis, Dubins-Schwarz theorem and the various boundedness assumptions, we get

$$\mathbb{E}\Big[\mathbb{E}\Big[\hat{J}_{j+1}^n - \check{J}_{j+1}^n \Big| \mathcal{F}_{\check{J}_{j+1}^n}\Big] \Big| \mathcal{F}_{\tau_j^n}\Big] \leq C\varepsilon_n^2 \delta_n.$$

Consequently,

$$\mathbb{E}[R_2] \leq C\delta_n,$$

which gives the result. □

Part 3: Proof of (2.10)

Here we prove (2.10), which completes the proof of the convergence in law of $\varepsilon_n^{-1} Z_T^n$ with the help of Proposition 2.1 and Proposition 2.2. As already seen, by Itô's formula, we have

$$\mathbb{E}[\Delta_{j,n}^4 | \mathcal{F}_{\tau_j^n}] = 6\mathbb{E}\Big[\int_{\tau_j^n}^{\tau_j^n + 1} (X_t - X_{\tau_j^n})^2 d\langle X \rangle_t \Big| \mathcal{F}_{\tau_j^n}\Big] = A_j,$$

$$\mathbb{E}[\Delta_{j,n}^3 | \mathcal{F}_{\tau_j^n}] = 3\mathbb{E}\Big[\int_{\tau_j^n}^{\tau_j^n + 1} (X_t - X_{\tau_j^n}) d\langle X \rangle_t \Big| \mathcal{F}_{\tau_j^n}\Big] = B_j.$$

Therefore, we obtain

$$
\begin{aligned}
\varepsilon_n^{-2}\sum_{j=0}^{N_t^n}\kappa_{\tau_j^n}A_j &= \varepsilon_n^{-2}\sum_{j=0}^{N_t^n}\kappa_{\tau_j^n}\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^4|\mathcal{F}_{\tau_j^n}\Big]+R_t\\
\varepsilon_n^{-1}\sum_{j=0}^{N_t^n}\kappa_{\tau_j^n}B_j &= \varepsilon_n^{-1}\sum_{j=0}^{N_t^n}\kappa_{\tau_j^n}\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^3|\mathcal{F}_{\tau_j^n}\Big]+R_t'
\end{aligned}
$$

where

$$
\begin{aligned}
R_t &= 6\varepsilon_n^{-2}\sum_{j=0}^{N_t^n}\kappa_{\tau_j^n}\mathbb{E}\Big[\int_{\tilde\tau_{j+1}^n}^{\tau_{j+1}^n}(X_u-X_u^n)^2(\sigma_u^X)^2du\Big|\mathcal{F}_{\tau_j^n}\Big],\\
R_t' &= 3\varepsilon_n^{-1}\sum_{j=0}^{N_t^n}\kappa_{\tau_j^n}\mathbb{E}\Big[\int_{\tilde\tau_{j+1}^n}^{\tau_{j+1}^n}(X_u-X_u^n)(\sigma_u^X)^2du\Big|\mathcal{F}_{\tau_j^n}\Big].
\end{aligned}
$$

Since $\varepsilon_n^{-1}\sup_t|X_t-X_t^n|\le K$ and $\sigma^X\le K$, R and R' converge to 0 uniformly in probability on $[0,T_0]$ by Lemma A.7. Using that for $b_1>0$ and $b_2>0$, the probability that a Brownian motion starting from zero hits level b_1 before level $-b_2$ is equal to $b_2/(b_2+b_1)$, we get

$$
\varepsilon_n^{-2}\frac{\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^4\Big|\mathcal{F}_{\tau_j^n}\Big]}{\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^2\Big|\mathcal{F}_{\tau_j^n}\Big]}=a_{\tau_j^n}^2,\quad \varepsilon_n^{-1}\frac{\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^3\Big|\mathcal{F}_{\tau_j^n}\Big]}{\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^2\Big|\mathcal{F}_{\tau_j^n}\Big]}=-s_{\tau_j^n},
$$

where

$$
a^2=\bar{l}^2+\underline{l}^2-\bar{l}\underline{l},\quad s=\underline{l}-\bar{l}.
$$

Then, to complete the proof, it suffices to show that the convergences

$$
\begin{aligned}
\sum_{j=0}^{N_\cdot^n}\kappa_{\tau_j^n}a_{\tau_j^n}^2\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^2|\mathcal{F}_{\tau_j^n}\Big] &\to \int_0^\cdot a_u^2(\sigma_u^Y)^2du,\\
\sum_{j=0}^{N_\cdot^n}\kappa_{\tau_j^n}s_{\tau_j^n}\mathbb{E}\Big[(X_{\tilde\tau_{j+1}^n}-X_{\tau_j^n})^2|\mathcal{F}_{\tau_j^n}\Big] &\to \int_0^\cdot s_u(\sigma_u^Y)^2du
\end{aligned}
$$

hold uniformy in probability on $[0,T_0]$. This follows from Lemma A.4 in [4] together with Lemma A.4.

A.3.2 Proof of (2.4)

Here we prove a moment convergence result. Thus the localization procedure does not apply here. We set

$$
\begin{aligned}
A_n &= \varepsilon_n^{-1}\int_0^T(X_t^n-X_t)b_t^Y dt,\\
B_n &= \varepsilon_n^{-1}\int_0^T(X_t^n-X_t)\sigma_t^Y dW_t^Y.
\end{aligned}
$$

We have

$$(\varepsilon_n^{-1} Z_T^n)^2 = (A_n + B_n)^2 \leq 2(A_n^2 + B_n^2).$$

Thus it is enough to prove the uniform integrability of (A_n^2) and (B_n^2) to obtain the result. For (A_n^2), we have

$$\sup_n (A_n)^2 \leq \Big(\int_0^T (\bar{l}_t \vee \underline{l}_t) \left| b_t^Y \right| dt \Big)^2 \leq \int_0^T (\bar{l}_t \vee \underline{l}_t)^2 (\rho_t)^2 (\sigma_t^Y)^2 dt.$$

The right hand side of the last inequality being integrable, this gives the result for $(A_n)^2$. We now turn to (B_n^2). The sequence (B_n^2) is non negative integrable and converges in law towards an integrable limit. Thus the uniform integrability is equivalent to the convergence in expectation, see for example [2]. Since

$$\varepsilon_n^{-2} \langle Z^n \rangle_T \to \frac{1}{6} \int_0^T a_t^2 (\sigma_t^Y)^2 dt$$

and

$$\varepsilon_n^{-2} \langle Z^n \rangle_T \leq \int_0^T (\bar{l}_t \vee \underline{l}_t)^2 (\sigma_t^Y)^2 dt,$$

we readily obtain

$$\mathbb{E}[B_n^2] = \mathbb{E}[\varepsilon_n^{-2} \langle Z^n \rangle_T] \to \frac{1}{6} \mathbb{E}[\int_0^T a_t^2 (\sigma_t^Y)^2 dt],$$

which concludes the proof.

A.4 Proof of Theorem 4.1

We start with the first part of Theorem 4.1. Let (m, v) be a non dominated couple. Suppose it is a dominated couple with respect to $\mathcal{Z}_T$. This means there exists a process s_t^* such that the associated expectation, say $m' = \mathbb{E}[Z_{T,s^*}]$, is larger than m and the expected error, say $v' = \mathbb{E}[(Z_{T,s^*})^2]$, is strictly smaller than v. From Lemma 2.1, for any η we can find an admissible strategy with limiting variable $Z^*_{s^*+\eta,s^*}$. Clearly, we can find η small enough, such that $\mathbb{E}[Z^*_{s^*+\eta,s^*}] = m'$ and

$$v' \leq \mathbb{E}[(Z^*_{s^*+\eta,s^*})^2] < v.$$

Consequently (m, v) is a dominated couple, which is absurd. Conversely, any point which is non dominated with respect to $\mathcal{Z}_T$ is non dominated since $a_t^2 \geq s_t^2$.

For the second part, it remains to show that the proposed discretization rules indeed lead to nearly efficient couples. The fact that they are admissible is clear from Proposition 2.3. Recall now that for the suggested rule

$$a_t^2 = (s_t^*)^2 + \frac{6\delta}{(\sigma_t^Y)^2}.$$

This equality gives that the limiting variable $Z^*_{a,s}$ associated to this discretization rule satisfies

$$\mathbb{E}[Z^*_{a,s}] = \frac{1}{3} \mathbb{E}\Big[\int_0^T s_t^* dY_t \Big]$$

and

$$\mathbb{E}[(Z^*_{a,s})^2] = \frac{1}{9}\mathbb{E}\Big[\Big(\int_0^T s^*_t dY_t\Big)^2\Big] + \frac{1}{6}\mathbb{E}\Big[\int_0^T \Big(a_t^2 - \frac{2}{3}(s^*_t)^2\Big)(\sigma^Y_t)^2 dt\Big]$$
$$= \frac{1}{9}\mathbb{E}\Big[\Big(\int_0^T s^*_t dY_t\Big)^2\Big] + \frac{1}{18}\mathbb{E}\Big[\int_0^T \Big((s^*_t)^2(\sigma^Y_t)^2\Big)dt\Big] + \delta T.$$

The couple

$$\Big(\frac{1}{3}\mathbb{E}\Big[\int_0^T s^*_t dY_t\Big], \frac{1}{9}\mathbb{E}\Big[\Big(\int_0^T s^*_t dY_t\Big)^2\Big] + \frac{1}{18}\mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}(s^*_t)^2(\sigma^Y_t)^2\Big)dt\Big]\Big)$$

being non dominated, we obtain the result.

Appendix B Linear-quadratic optimal control

We give here a summary of useful formulas from [18]. Consider a controlled system governed by the following linear SDE:

$$\begin{cases} dX_t = (A_t X_t + B_t u_t + f_t)dt + \sum_{j=1}^m D^j_t u_t dW^j_t, \\ X_0 = x \in \mathbb{R}^n, \end{cases} \tag{B.1}$$

where x is the initial state and $W = (W^1, \cdots, W^m)$ is a m-dimensional Brownian motion on a given filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ and $u \in L^2_{\mathcal{F}}([0,T], \mathbb{R}^m)$ is a control. For each control u, the associated cost is

$$J(u) = \mathbb{E}\left[\int_0^T \frac{1}{2}\Big(X_t' Q_t X_t + u_t' R_t u_t\Big)dt + \frac{1}{2}X_T' H X_T\right]. \tag{B.2}$$

We suppose that all the parameters are deterministic and continuous on $[0,T]$ and H belongs to S^n_+ the set of $n \times n$ symmetric positive matrices. We introduce the following matrix Riccati equation

$$\begin{cases} \dot{P}_t = -P_t A_t - A_t' P_t - Q_t + P_t B_t K_t^{-1} B_t' P_t, \\ P_T = H, \\ K_t = R_t + \sum_{j=1}^m D^{j'}_t P_t D^j_t > 0, \quad \forall t \in [0,T], \end{cases} \tag{B.3}$$

along with an equation

$$\begin{cases} \dot{g}_t = -A_t' g_t + P_t B_t K_t^{-1} B_t' g_t - P_t f_t, \\ g_T = 0. \end{cases} \tag{B.4}$$

Then following result is given in [18].

Theorem B.1. *If* (B.3) *and* (B.4) *admit solutions $P \in C([0,T], S^n_+)$ and $g \in C([0,T], \mathbb{R}^n)$ respectively, then the stochastic linear-quadratic control problem* (B.1)-(B.2) *has an optimal feedback control*

$$u^*(t,x) = -K_t^{-1} B_t'(P_t X_t + g_t).$$

Moreover, the optimal cost value is

$$J^* = \frac{1}{2}\int_0^T \Big(2 f_t' g_t - g_t B_t K_t^{-1} B_t' g_t\Big)dt + \frac{1}{2}x' P_0 x + x g_0.$$

Acknowledgements

We thank Philippe Amzelek and Joe Bonnaud from BNP-Paribas for inspiring discussions.

References

[1] D. Bertsimas, L. Kogan, and A. W. Lo. When is time continuous? *Journal of Financial Economics*, 55(2):173–204, 2000.

[2] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Wiley, 2009.

[3] M. Fukasawa. Asymptotically efficient discrete hedging. *Stochastic Analysis with Financial Applications*, pages 331–346, 2011.

[4] M. Fukasawa. Discretization error of stochastic integrals. *The Annals of Applied Probability*, 21(4):1436–1465, 2011.

[5] M. Fukasawa. Efficient discretization of stochastic integrals. *Finance and Stochastics*, 18(1):175–208, 2014.

[6] S. Geiss. Quantitative approximation of certain stochastic integrals. *Stochastics: An International Journal of Probability and Stochastic Processes*, 73(3-4):241–270, 2002.

[7] E. Gobet, N. Landon, et al. Almost sure optimal hedging strategy. *The Annals of Applied Probability*, 24(4):1652–1690, 2014.

[8] E. Gobet and E. Temam. Discrete time hedging errors for options with irregular payoffs. *Finance and Stochastics*, 5(3):357–367, 2001.

[9] T. Hayashi and P. A. Mykland. Evaluating hedging errors: an asymptotic approach. *Mathematical finance*, 15(2):309–343, 2005.

[10] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Berlin: Springer-Verlag, 2003.

[11] A. E. Lim and X. Y. Zhou. Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, 27(1):101–120, 2002.

[12] C. Y. Robert and M. Rosenbaum. On the microstructural hedging error. *SIAM Journal on Financial Mathematics*, 1(1):427–453, 2010.

[13] H. Rootzen. Limit distributions for the error in approximations of stochastic integrals. *The Annals of Probability*, pages 241–251, 1980.

[14] M. Rosenbaum, P. Tankov, et al. Asymptotically optimal discretization of hedging strategies with jumps. *The Annals of Applied Probability*, 24(3):1002–1048, 2014.

[15] A. Sepp. When you hedge discretely: Optimization of sharpe ratio for delta-hedging strategy under discrete hedging and transaction costs. *Available at SSRN 1865998*, 2013.

[16] P. Tankov and E. Voltchkova. Asymptotic analysis of hedging errors in models with jumps. *Stochastic processes and their applications*, 119(6):2004–2027, 2009.

[17] R. Zhang. *Couverture approchée des options Européennes*. PhD thesis, Ecole des Ponts ParisTech, 1999.

[18] X. Y. Zhou and D. Li. Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Applied Mathematics and Optimization*, 42(1):19–33, 2000.