



Optimal discretization of stochastic integrals for degenerate semimartingale

Emmanuel Gobet, Uladzislau Stazhynski

► To cite this version:

Emmanuel Gobet, Uladzislau Stazhynski. Optimal discretization of stochastic integrals for degenerate semimartingale. 2015. hal-01241190v1

HAL Id: hal-01241190

<https://polytechnique.hal.science/hal-01241190v1>

Preprint submitted on 10 Dec 2015 (v1), last revised 8 Nov 2017 (v2)

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

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

Optimal discretization of stochastic integrals for degenerate semimartingale

Emmanuel Gobet* Uladzislau Stazhynski†

First version: May 28, 2015.
This version: December 10, 2015

Abstract

We study the optimal discretization error of stochastic integrals, driven by a multidimensional continuous Brownian semimartingale. As a main difference compared to *Gobet-Landon, Annals of Applied Probability 2014 (DOI: 10.1214/13-AAP959)*, the martingale part of the semimartingale may be degenerate. In this setting we establish a pathwise lower bound for the renormalized quadratic variation of the error and we provide a sequence of discretization stopping times, which is asymptotically optimal. The latter is defined as hitting times of random ellipsoids by the semimartingale at hand.

Keywords: discretization of stochastic integrals, hitting times, random ellipsoids, almost sure convergence.

MSC2010: 60G40, 60F15, 60H05.

1 Introduction

Statement of the problem. In this work we consider the problem of finding a finite sequence of optimal stopping times $\mathcal{T}^n = \{0 = \tau_0^n < \tau_1^n < \dots < \tau_{N_T^n}^n = T\}$ which minimizes the renormalized quadratic variation of the discretization error of the stochastic integral

$$Z_s^n = \int_0^s v(t, S_t) \cdot dS_t - \sum_{\tau_{i-1}^n < s} v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n}), \quad (1.1)$$

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

†Ecole Normale Supérieure, 45 rue d'Ulm, 75230 Paris cedex 5, France. Email: stazhinsky.vlad@yandex.ru

where S is a d -dimensional continuous Brownian semimartingale and $v(t, x)$ is a \mathbb{R}^d -valued continuous function. Here $T \in (0, +\infty)$ is fixed. The number of stopping times N_T^n is allowed to be random.

The almost sure minimization of Z_T^n is hopeless since after suitable renormalization and under some mild assumptions on the model, Z_T^n weakly converges to a mixture of Gaussian random variables (see [Roo80][KP91][JP12]). Alternatively we aim at minimizing a.s. the product

$$N_T^n \langle Z^n \rangle_T. \quad (1.2)$$

The choice of this minimization criterion is inspired by the fact that in many particular cases with deterministic discretization times, we have $\mathbb{E}(\langle Z^n \rangle_T) \sim \text{Const}/N_T^n$ as $N_T^n \rightarrow +\infty$.

Background results. The study of such minimization problems has been initiated by [Fuk11b] in dimension $d = 1$, but instead of (1.2) the author considers a criterion in expectation for both terms, ie. $\mathbb{E}(N_T^n) \mathbb{E}(\langle Z^n \rangle_T)$. However, if $n \rightarrow +\infty$ denotes an asymptotic parameter (defined later), observe that

$$\liminf_{n \rightarrow +\infty} \mathbb{E}(N_T^n) \mathbb{E}(\langle Z^n \rangle_T) \underset{\text{Cauchy-Schwarz ineq.}}{\geq} \liminf_{n \rightarrow +\infty} \left[\mathbb{E} \left(\sqrt{N_T^n \langle Z^n \rangle_T} \right) \right]^2 \quad (1.3)$$

$$\underset{\text{Fatou lemma}}{\geq} \left[\mathbb{E} \left(\sqrt{\liminf_{n \rightarrow +\infty} N_T^n \langle Z^n \rangle_T} \right) \right]^2. \quad (1.4)$$

Since the solution to the problem of a.s. minimizing (1.2) exists (see Theorem 5.1) and is such that N_T^n and $\langle Z^n \rangle_T$ are asymptotically proportional (see the limits (5.14) and (5.15)), the above inequalities can be turned into equalities (with a little of technical work) and therefore, we get for free a solution to minimizing asymptotically $\mathbb{E}(N_T^n) \mathbb{E}(\langle Z^n \rangle_T)$, however with substantially more information.

As an extra motivation for this theoretical study, we refer to the recent work of Hairer et al [HHJ15], which highlights that discretization schemes for stochastic differential equations using deterministic grid may surprisingly converge very slowly in L_2 -norm. Actually any slow rate is possible [JMY15]. These amazing results give a strong incentive for studying discretization problems with stochastic grids and pathwise criterion. Applications of the current results to pathwise-optimal discretization of SDEs are left to future research.

The ω -wise minimization of (1.2) has been addressed in a multi-dimensional setting $d \geq 1$, under some model conditions, in [GL14a]. In this reference, the problem is originally considered in a financial setting: Z^n is interpreted as the hedging error of the discrete Delta-hedging strategy of an option with underlying asset S , maturity T and price function $u(t, S_t)$ (then we must take $v(t, x) = \nabla_x u(t, x)$). In this case $(\tau_i^n)_{1 \leq i \leq N_T^n}$ read as trading dates. For more details on this interpretation, see [Fuk11a] and the introduction of [GL14a].

Our contributions. In the current work, we revisit the mathematical analysis of [GL14a] with significant improvements and extensions. First, we allow S to be a more general Brownian semimartingale $S = A + M$, while in [GL14a] S is essentially a local Brownian martingale ($A = 0, M = \int_0^\cdot \sigma_s dB_s$). Second, the martingale part of S can be degenerate in our setting, while a stronger a.s. ellipticity (on σ) is considered in [GL14a]. Third, we provide a strategy \mathcal{T}^n attaining the lower bound, while in [GL14a], only a μ -optimal strategy (with μ small) is designed.

The first improvement aforementioned is achieved by establishing that admissible strategies for the semimartingale S and for its local martingale part M are the same: see Theorem 3.4. This allows to transfer some general results of [GL14a] to our extended setting. The second improvement is important for applications and requires a quite delicate analysis. Actually the possible degeneracy lets us lose some continuity property (in particular because we need to consider the inverse σ^{-1}) and some convergence properties. To overcome these issues, we assume that in a sense, σ_t and $D_x v(t, S_t)$ are not zero simultaneously: for a precise statement, see Assumption (H_C) or a weaker Assumption (H_A). The third improvement consists in providing an optimal strategy: we know that heuristically, the natural candidate is a sequence of hitting times by S of random ellipsoids which characteristics depend on v and S . However, in general and in particular because of the degenerate setting, this strategy is not admissible in the sense of our definitions (Section 2). Alternatively, we prove that a small perturbation makes the strategy admissible, without altering its asymptotic optimality.

Organisation of the paper. In Section 2, we define the model and the admissible strategies under study. In Section 3, we state and establish crucial properties of admissible strategies. The minimization of (1.2) is studied in Section 4, while exhibiting a optimal strategy is made in Section 5. We also present a few examples and a numerical experiment. Technical results are postponed to Appendix.

Notation used throughout this work.

- We denote by $x \cdot y$ the scalar product between two vectors x and y and by $|x| = (x \cdot x)^{\frac{1}{2}}$ the Euclidean norm of x . The induced norm of a $m \times d$ -matrix is denoted by $|A| := \sup_{x \in \mathbb{R}^d, |x|=1} |Ax|$.
- The transposition of a matrix A is denoted by A^T ; we denote by $\text{Tr}(A)$ the trace of a square matrix A ; Id_d stands for the identity matrix of size d .
- $\mathcal{S}^d(\mathbb{R})$, $\mathcal{S}_+^d(\mathbb{R})$ and $\mathcal{S}_{++}^d(\mathbb{R})$ are respectively the sets of symmetric, symmetric non-negative definite and symmetric positive-definite $d \times d$ ma-

trices with real coefficients.

- For $A \in \mathcal{S}^d(\mathbb{R})$ we denote $\Lambda(A) = (\lambda_1(A), \dots, \lambda_d(A))$ the eigenvalues of A placed in decreasing order, we set $\lambda_{\min}(A) := \lambda_d(A)$ and $\lambda_{\max}(A) := \lambda_1(A)$.
- We denote by $\text{Diag}(a_1, \dots, a_d)$ the square matrix of size d with diagonal entries a_1, \dots, a_d .
- For the partial derivatives of a function $f(t, x)$ we write

$$D_t f(t, x) = \frac{\partial f}{\partial t}(t, x), \quad D_{x_i} f(t, x) = \frac{\partial f}{\partial x_i}(t, x), \quad D_{x_i x_j}^2 f(t, x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(t, x).$$

- For a \mathbb{R}^d -valued semimartingale S we denote $\langle S \rangle_t$ its matrix of cross-variations $(\langle S^i, S^j \rangle_t)_{1 \leq i, j \leq d}$.
- We sometimes write f_t for $f(t, S_t)$ where S is a semimartingale and f is some function.
- For a given sequence of stopping times \mathcal{T}^n , the last stopping time before $t \leq T$ is defined by $\phi(t) = \max\{\tau_j^n : \tau_j^n < t\}$. We omit to indicate the dependence on n . Furthermore for a process $(f_t)_{0 \leq t \leq T}$ we write $\Delta f_t := f_t - f_{\phi(t-)}$. Besides we set $\Delta_t := t - \phi(t-)$ and $\Delta \tau_i^n := \tau_i^n - \tau_{i-1}^n$.
- C_0 stands for a a.s. finite non-negative random variable, which may change from line to line.

2 Model and strategies

2.1 Probabilistic model: assumptions

Let $T > 0$ and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ be a filtered probability space supporting a d -dimensional Brownian motion $B = (B^i)_{1 \leq i \leq d}$ defined on $[0, T]$, where $(\mathcal{F}_t)_{0 \leq t \leq T}$ is the \mathbb{P} -augmented natural filtration of B and $\mathcal{F} = \mathcal{F}_T$. Let

$$(\alpha, \theta_\sigma) \in \left(\frac{1}{2}, 1\right] \times (0, 1] \quad (2.1)$$

be some regularity parameters and let $(S_t)_{0 \leq t \leq T}$ be a d -dimensional continuous semimartingale of the form

$$S_t = A_t + M_t, \quad 0 \leq t \leq T, \quad (2.2)$$

where the processes A and M satisfy the following hypotheses.

(**H_A**) The process A is continuous, adapted and of finite variation, and satisfies

$$|A_t - A_s| \leq C_0 |t - s|^\alpha \quad \forall s, t \in [0, T] \quad \text{a.s.} \quad (\mathbf{H}_A)$$

(**H_M**) The process M is a continuous local martingale of the form

$$M_t = \int_0^t \sigma_s dB_s, \quad 0 \leq t \leq T, \quad (\mathbf{H}_M)$$

where σ is a continuous adapted $d \times d$ -matrix valued process, such that the value σ_t is a.s. non-zero for any $t \in [0, T]$, and

$$|\sigma_t - \sigma_s| \leq C_0 |t - s|^{\theta_\sigma/2} \quad \forall s, t \in [0, T] \quad \text{a.s.}$$

Furthermore, we assume that the function v , involved in (1.1), is a $\mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d)$ function with values in \mathbb{R}^d . For applications like in [GL14a], we shall allow its derivatives in uniform norm (in space) to explode as $t \rightarrow T$, whilst remaining bounded a.s. in an infinitesimal tube centered at $(t, S_t)_{0 \leq t < T}$. This is stated precisely in what follows.

(**H_v**) Let $\mathcal{D} \in \{D_{x_j}, D_{x_j x_k}^2, D_t : 1 \leq j, k \leq d\}$, then

$$\mathbb{P} \left(\lim_{\delta \rightarrow 0} \sup_{0 \leq t < T} \sup_{|x - S_t| \leq \delta} |\mathcal{D}v(t, x)| < +\infty \right) = 1. \quad (\mathbf{H}_v)$$

2.2 Class $\mathcal{T}^{\text{adm.}}$ of admissible sequences of strategies

Now we define the class of strategies under consideration. As the optimality in our problem is achieved asymptotically as a parameter $n \rightarrow +\infty$, a strategy is naturally indexed by $n \in \mathbb{N}$: a *strategy* is a finite sequence of increasing stopping times

$$\mathcal{T}^n := \{\tau_0^n = 0 < \dots < \tau_i^n < \dots < \tau_{N_T^n}^n = T\}, \quad \text{with } N_T^n < +\infty \text{ a.s.}$$

We now define the appropriate asymptotic framework. Let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a sequence of positive deterministic real numbers converging to 0 as $n \rightarrow +\infty$ and such that

$$\sum_{n \geq 0} \varepsilon_n^2 < +\infty.$$

In the following, all convergences are taken as $n \rightarrow +\infty$. The above summability enables to derive a.s. convergence results: alternatively, had we assumed only $\varepsilon_n \rightarrow 0$, using a subsequence-based argument (see [GL14b, Section 2.2]) we would get convergences in probability.

On the one hand the parameter ε_n controls the oscillations of S between two successive stopping times in \mathcal{T}^n .

($\mathbf{A}_S^{\text{osc.}}$) The following non-negative random variable is a.s. finite:

$$\sup_{n \geq 0} \left(\varepsilon_n^{-2} \sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} |S_t - S_{\tau_{i-1}^n}|^2 \right) < +\infty.$$

Here the lower argument in the assumption ($\mathbf{A}_S^{\text{osc.}}$) refers explicitly to the process at hand. On the other hand $\varepsilon_n^{-2\rho_N}$ (for some $\rho_N \geq 1$) upper bounds up to a constant the number of stopping times in the strategy \mathcal{T}^n .

(\mathbf{A}_N) The following non-negative random variable is a.s. finite:

$$\sup_{n \geq 0} (\varepsilon_n^{2\rho_N} N_T^n) < +\infty.$$

In the above, ρ_N is a given parameter satisfying

$$1 \leq \rho_N < \left(1 + \frac{\theta_\sigma}{2}\right) \wedge \frac{4}{3} \wedge \left(\frac{1}{2} + \alpha\right). \quad (2.3)$$

where (α, θ_σ) are given in (2.1).

Definition 1. A sequence of strategies $\mathcal{T} := \{\mathcal{T}^n : n \geq 0\}$ is admissible for the process S if it fulfills the hypotheses ($\mathbf{A}_S^{\text{osc.}}$) and (\mathbf{A}_N). The set of admissible sequences is denoted by $\mathcal{T}_S^{\text{adm.}}$.

The larger ρ_N , the wider the class of strategies under consideration.

3 General results for admissible strategies

This section gathers preliminary results, needed to establish the subsequent main results. In Subsection 3.1, we recall without proof some estimates about the mesh size $\sup_{1 \leq i \leq N_T^n} \Delta \tau_i^n$ of the time grid \mathcal{T}^n simultaneously for any n , as well as bounds on (local) martingales depending on n . This is preparatory for Subsection 3.2 where we establish an important result: in our setting, admissible sequences of strategies for S and M are the same. Last in Subsection 3.3, we establish the a.s. convergence of weighted quadratic variations under some mild assumptions, which are crucial to derive our new optimality results.

3.1 Control of $\Delta \tau^n$ and martingale increments

We start from a simple and efficient criterion for a.s. convergence of continuous local martingales.

Lemma 3.1 ([GL14a, Corollary 2.1]). *Let $p > 0$, and let $\{(K_t^n)_{0 \leq t \leq T} : n \geq 0\}$ be a sequence of continuous scalar local martingales vanishing at zero. Then*

$$\sum_{n \geq 0} \langle K^n \rangle_T^{p/2} < +\infty \text{ a.s.} \iff \sum_{n \geq 0} \sup_{0 \leq t \leq T} |K_t^n|^p < +\infty \text{ a.s..}$$

The useful application is the sense \Rightarrow : by controlling the summability of quadratic variations, we obtain the non trivial a.s. convergence of $\sup_{0 \leq t \leq T} |K_t^n|$ to 0. This kind of reasoning is used in this work.

The next two lemmas yield controls of $\Delta\tau_i$ and of martingales increments for an admissible sequence of strategies. In view of the Brownian motion scaling property one might guess that an admissible sequence of strategies $\mathcal{T} = \{\mathcal{T}^n : n \geq 0\}$ yields stopping times increments of magnitude roughly equal to ε_n^2 . More generally, we can study in a similar way the increments of martingales. Here we give a rigorous statement of these heuristics.

Lemma 3.2 ([GL14a, Corollary 2.2]). *Assume (\mathbf{H}_M) and let $\mathcal{T} = \{\mathcal{T}^n : n \geq 0\}$ be a sequence of strategies. Let $\rho > 0$, then the following hold:*

(i) *Assume \mathcal{T} satisfies $(\mathbf{A}_M^{\text{osc.}})$, then*

$$\sup_{n \geq 0} (\varepsilon_n^{\rho-1} \sup_{1 \leq i \leq N_T^n} \Delta\tau_i^n) < +\infty \text{ a.s..}$$

(ii) *Assume \mathcal{T} satisfies $(\mathbf{A}_M^{\text{osc.}})$ -(\mathbf{A}_N), then*

$$\sup_{n \geq 0} (\varepsilon_n^{\rho-2} \sup_{1 \leq i \leq N_T^n} \Delta\tau_i^n) < +\infty \text{ a.s..}$$

Lemma 3.3 ([GL14a, Corollary 2.3]). *Assume (\mathbf{H}_M) . Let $((K_t^n)_{0 \leq t \leq T})_{n \geq 0}$ be a sequence of \mathbb{R}^d -valued continuous local martingales such that $\langle K^n \rangle_t = \int_0^t \kappa_r^n dr$ for a measurable adapted κ^n satisfying the following inequality: there exist a non-negative a.s. finite random variable C_κ and a deterministic parameter $\theta \geq 0$ such that*

$$0 \leq |\kappa_r^n| \leq C_\kappa (|\Delta M_r|^{2\theta} + |\Delta_r|^\theta) \quad \forall 0 \leq r < T, \quad \forall n \geq 0, \text{ a.s..}$$

Finally, let $\rho > 0$, then the following assertions hold.

(i) *Assume \mathcal{T} satisfies $(\mathbf{A}_M^{\text{osc.}})$, then*

$$\sup_{n \geq 0} (\varepsilon_n^{\rho-(1+\theta)/2} \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta K_t^n|) < +\infty \text{ a.s..}$$

(ii) *Assume \mathcal{T} satisfies $(\mathbf{A}_M^{\text{osc.}})$ -(\mathbf{A}_N), then*

$$\sup_{n \geq 0} (\varepsilon_n^{\rho-(1+\theta)} \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta K_t^n|) < +\infty \text{ a.s..}$$

3.2 The admissible sequences of strategies for S and M coincide

We now aim at proving the following Theorem.

Theorem 3.4. *Let S be a semimartingale of the form (2.2) and satisfying (\mathbf{H}_A) -(\mathbf{H}_M). Then a sequence of strategies $\mathcal{T} = \{\mathcal{T}_n : n \geq 0\}$ is admissible for S if and only if it is admissible for M with the same parameter ρ_N : in other words, if \mathcal{T} satisfies (\mathbf{A}_N) ,*

$$(\mathbf{A}_M^{\text{osc.}}) \Leftrightarrow (\mathbf{A}_S^{\text{osc.}}).$$

Rephrased differently, defining admissible sequence of strategies based on the martingale M is robust to perturbation by adding to M a finite variation process A , satisfying α -Hölder regularity with $\alpha > 1/2$.

Proof. For convenience in the proof, we adopt the short notation

$$|\Delta\tau^n|_\infty := \sup_{1 \leq i \leq N_T^n} \Delta\tau_i^n, \quad |\Delta U|_\infty := \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta U_t|,$$

for any process U .

Proof of \Rightarrow . Suppose first that $\mathcal{T} = \{\mathcal{T}_n : n \geq 0\}$ is admissible for S . Let us prove that it is admissible for M , i.e. the assumption $(\mathbf{A}_M^{\text{osc.}})$ is satisfied. We proceed in several steps.

\triangleright *Step 1. Preliminary bound.* From $|M_t - M_s| \leq |S_t - S_s| + |A_t - A_s|$ and (\mathbf{H}_A) , we get

$$|\Delta M|_\infty \leq |\Delta S|_\infty + C_0 |\Delta\tau^n|_\infty^\alpha \leq C_0 (\varepsilon_n + |\Delta\tau^n|_\infty^\alpha). \quad (3.1)$$

Using Itô's formula and (\mathbf{H}_M) , we obtain that for any $0 \leq s < t \leq T$

$$0 \leq t - s \leq C_E^{-1} \int_s^t \text{Tr}(\sigma_r \sigma_r^\top) dr = C_E^{-1} \sum_{j=1}^d (\langle S^j \rangle_t - \langle S^j \rangle_s) \quad (3.2)$$

$$= C_E^{-1} \sum_{j=1}^d \left((S_t^j - S_s^j)^2 - 2 \int_s^t (S_r^j - S_s^j) dS_r^j \right), \quad (3.3)$$

where $C_E := \inf_{t \in [0, T]} \text{Tr}(\sigma_t \sigma_t^\top) > 0$ a.s.. Hence

$$\Delta t \leq C_E^{-1} \left(C_0 \varepsilon_n^2 + 2 \sum_{j=1}^d \left| \int_{\phi(t)}^t \Delta S_r^j dA_r^j \right| + 2 \sum_{j=1}^d \left| \int_{\phi(t)}^t \Delta S_r^j dM_r^j \right| \right). \quad (3.4)$$

Using that A is of finite variation and $(\mathbf{A}_S^{\text{osc.}})$, we get the crude estimate

$$\sum_{j=1}^d \left| \int_{\phi(t)}^t \Delta S_r^j dA_r^j \right| \leq C_0 \varepsilon_n. \quad (3.5)$$

Now consider the local martingale $K_t^{n,j} = \varepsilon_n^{\frac{2}{p}-1} \left(\int_0^t \Delta S_r^j dM_r^j \right)$ for some $p > 0$. We have

$$\sum_{n \geq 0} \langle K^{n,j} \rangle_T^{\frac{p}{2}} = \sum_{n \geq 0} \varepsilon_n^{2-p} \left(\int_0^T |\Delta S_r^j|^2 d\langle M^j \rangle_r \right)^{\frac{p}{2}} \leq C_0 \sum_{n \geq 0} \varepsilon_n^2 < +\infty \quad \text{a.s.},$$

which by Lemma 3.1 implies that $\sum_{n \geq 0} \sup_{0 \leq t \leq T} |K_t^{n,j}|^p < +\infty$ a.s., and thus $\sup_{n \geq 0} \sup_{0 \leq t \leq T} |K_t^{n,j}| < +\infty$ a.s.. This reads

$$\sup_{0 \leq t \leq T} \left| \int_0^t \Delta S_r^j dM_r^j \right| \leq C_0 \varepsilon_n^{1-\frac{2}{p}} = C_0 \varepsilon_n^{1-\delta}, \quad (3.6)$$

where $\delta = 2/p$ is an arbitrary positive number. Plugging this and (3.5) into (3.4) yields

$$|\Delta \tau^n|_\infty \leq C_0(\varepsilon_n^2 + \varepsilon_n + \varepsilon_n^{1-\delta}) \leq C_0 \varepsilon_n^{1-\delta}. \quad (3.7)$$

The above is analogous to Lemma 3.2-(i) but under the assumption $(\mathbf{A}_S^{\text{osc}})$. Combined with (3.1), we then deduce

$$|\Delta M|_\infty \leq C_0 \varepsilon_n^{\alpha(1-\delta)} \quad (3.8)$$

for any given $\delta \in (0, 1)$.

▷ *Step 2.* We prove the following lemma, which gives the basis for a continuation argument (Step 3): once we have estimated $|\Delta M|_\infty$ with some order w.r.t. ε_n , we obtain automatically a slightly better order, up to reaching the order 1, as required by $(\mathbf{A}_M^{\text{osc}})$.

Lemma 3.5. *Suppose that for some $\beta > 0$*

$$\sup_{n \geq 0} \left(\varepsilon_n^{-\beta} \sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} |\Delta M_t|^2 \right) < +\infty \quad \text{a.s.} \quad (3.9)$$

Then for any $\rho > 0$

$$\sup_{n \geq 0} \left(\varepsilon_n^{-(\beta-\rho)} \sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} \left| \sum_{j=1}^d \Delta \langle M^j \rangle_t \right| \right) < +\infty \quad \text{a.s.}$$

Proof. Let $p > 0$. Consider the following two sequences of processes:

$$U_t^n = \varepsilon_n^{2-\beta p+2\rho_N} \sum_{\tau_{i-1}^n < t} \left| \sum_{j=1}^d \Delta \langle M^j \rangle_{\tau_i^n \wedge t} \right|^p,$$

$$V_t^n = \varepsilon_n^{2-\beta p+2\rho_N} \sum_{\tau_{i-1}^n < t} \sup_{s \in (\tau_{i-1}^n, \tau_i^n \wedge t]} |\Delta M_s|^{2p}.$$

We aim at proving that $\sum_{n \geq 0} U_T^n < +\infty$ a.s. using Lemma A.1 in Appendix. First, $\sum_{n \geq 0} V_T^n$ converges a.s.: indeed using (A_N) and (3.9) we obtain

$$\sum_{n \geq 0} V_T^n \leq C_0 \sum_{n \geq 0} \varepsilon_n^{2-\beta p+2\rho_N} N_T^n \sup_{1 \leq i \leq N_T^n} \sup_{s \in (\tau_{i-1}^n, \tau_i^n]} |\Delta M_s|^{2p} \leq C_0 \sum_{n \geq 0} \varepsilon_n^2 < +\infty.$$

Second observe that for any n , $t \mapsto V_t^n$ is a.s. non-decreasing. Last it remains to verify the relation of domination of Lemma A.1-(iii). Let $k \in \mathbb{N}$, let θ_k be defined as in the quoted lemma. On the set $\{\tau_{i-1}^n < t \wedge \theta_k\}$ from a conditional version of the multidimensional BDG inequality we have

$$\mathbb{E} \left(\left| \sum_{j=1}^d \Delta \langle M^j \rangle_{\tau_i^n \wedge t \wedge \theta_k} \right|^p \middle| \mathcal{F}_{\tau_{i-1}^n} \right) \leq c_p \mathbb{E} \left(\sup_{\tau_{i-1}^n < s \leq \tau_i^n \wedge t \wedge \theta_k} |\Delta M_s|^{2p} \middle| \mathcal{F}_{\tau_{i-1}^n} \right).$$

Then it follows that

$$\begin{aligned} \mathbb{E} (U_{t \wedge \theta_k}^n) &= \varepsilon_n^{2-\beta p+2\rho_N} \sum_{i=1}^{+\infty} \mathbb{E} \left(1_{\tau_{i-1}^n < t \wedge \theta_k} \mathbb{E} \left(\left| \sum_{j=1}^d \Delta \langle M^j \rangle_{\tau_i^n \wedge t \wedge \theta_k} \right|^p \middle| \mathcal{F}_{\tau_{i-1}^n} \right) \right) \\ &\leq c_p \mathbb{E} (V_{t \wedge \theta_k}^n). \end{aligned}$$

Hence by Lemma A.1, we obtain that $\sum_{n \geq 0} U_T^n$ converges a.s. and thus $\sup_{n \geq 0} U_T^n < +\infty$ a.s..

Now write $\varepsilon_n^{2-\beta p+2\rho_N} \sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} \left| \sum_{j=1}^d \Delta \langle M^j \rangle_t \right|^p \leq U_T^n$, which implies

$$\sup_{n \geq 0} \left(\varepsilon_n^{(2+2\rho_N)/p-\beta} \sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} \left| \sum_{j=1}^d \Delta \langle M^j \rangle_t \right| \right) < +\infty \quad \text{a.s..}$$

To conclude, choose $p = \frac{2+2\rho_N}{\rho}$ to get the desired result. \square

\triangleright *Step 3. Continuation scheme.* Take $\delta > 0$, as in (3.8), set $d_0 = \alpha(1 - \delta)$ and $\rho_0 = \frac{(2\alpha - 1)d_0}{2\alpha} > 0$. Consider the sequence $(d_m)_{m \geq 0}$ given by $d_{m+1} = 2\alpha d_m - \alpha \rho_0$ for $m \geq 0$. Assume for a while that

$$d_{m+1} - d_m \geq \alpha \rho_0, \quad (3.10)$$

and let us show by induction that, for any $m \geq 0$,

$$|\Delta M|_\infty \leq C_0 \varepsilon_n^{\min(d_m, 1)}. \quad (3.11)$$

The case $m = 0$ stems directly from (3.8). Now suppose that (3.11) holds for m . If $d_m \geq 1$, since $d_{m+1} \geq d_m$ owing to (3.10), (3.11) is valid for $m + 1$. If $d_m < 1$, then we have $|\Delta M|_\infty \leq C_0 \varepsilon_n^{d_m}$ and using Lemma 3.5 we obtain

$$\left| \sum_{j=1}^d \Delta \langle M^j \rangle \right|_\infty \leq C_0 \varepsilon_n^{2d_m - \rho_0}.$$

Consequently (3.2) gives $|\Delta\tau^n|_\infty \leq C_0 \varepsilon_n^{2d_m - \rho_0}$ which, combined with (3.1), yields

$$|\Delta M|_\infty \leq C_0 \varepsilon_n^{\min(1, \alpha(2d_m - \rho_0))}.$$

This finishes the proof of (3.11) for $m + 1$. It remains to show (3.10) by induction. For $m = 0$ we get $d_1 = 2\alpha d_0 - \alpha \rho_0$ and thus

$$d_1 - d_0 = (2\alpha - 1)d_0 - \frac{(2\alpha - 1)d_0}{2} = \frac{(2\alpha - 1)d_0}{2} = \alpha \rho_0.$$

Suppose that (3.10) is true for all $m < k$ and let us extend to $m = k$. We write

$$d_{m+1} - d_m = (2\alpha - 1)d_m - \frac{(2\alpha - 1)d_0}{2} \geq (2\alpha - 1)d_0 - \frac{(2\alpha - 1)d_0}{2} = \alpha \rho_0,$$

using that $d_m \geq d_0$ by the induction assumption. We are done.

▷ *Step 4. Conclusion.* In view of (3.10), $(d_m)_{m \geq 0}$ becomes larger than 1 for some m , for which (3.11) simply writes $|\Delta M|_\infty \leq C_0 \varepsilon_n$. $(\mathbf{A}_M^{\text{osc.}})$ is proved. \square

Proof of \Leftarrow . Now suppose that the sequence \mathcal{T} is admissible for M . Let us prove the admissibility of \mathcal{T} for the process S . Again it is enough to verify the assumption $(\mathbf{A}_S^{\text{osc.}})$. Similarly to the decomposition (3.1), we have

$$|\Delta S|_\infty \leq |\Delta M|_\infty + |\Delta A|_\infty \leq C_0(\varepsilon_n + |\Delta\tau^n|_\infty^\alpha).$$

From Lemma 3.2-(ii), for any $\gamma > 0$, we have $|\Delta\tau^n|_\infty \leq C_0 \varepsilon_n^{2-\gamma}$ a.s.. Since $\alpha > 1/2$, we can choose γ such that $(2 - \gamma)\alpha > 1$ and for such γ we deduce $|\Delta S|_\infty \leq C_0(\varepsilon_n + \varepsilon_n^{(2-\gamma)\alpha}) \leq C_0 \varepsilon_n$. The proof is complete. \square

Remark 1.

- Theorem 3.4 implies that if a sequence of strategies fulfills (\mathbf{A}_N) , we do not need to emphasize anymore the dependence of the assumption $(\mathbf{A}_M^{\text{osc.}})$ on a particular process M or S ; in that case, we will write simply $(\mathbf{A}^{\text{osc.}})$ and will refer to admissible sequence of strategies $\mathcal{T}^{\text{adm.}} := \mathcal{T}_M^{\text{adm.}} = \mathcal{T}_S^{\text{adm.}}$.
- In addition, we can use all the results for admissible sequences of strategies based on the local martingale M and $(\mathbf{A}_M^{\text{osc.}})$ (as those from [GL14a]): in particular, for any admissible sequences of strategies (for M or S), we have $\sup_{1 \leq i \leq N_T^n} |\Delta\tau_i^n| \leq C_0 \varepsilon_n^{2-\gamma}$ for any $\gamma > 0$.

A direct consequence of Lemma 3.2-(ii), (\mathbf{H}_A) and Theorem 3.4 is the following.

Corollary 3.6. *Let S be a semimartingale of the form (2.2) and satisfying (\mathbf{H}_A) -(\mathbf{H}_M). If $\mathcal{T} \in \mathcal{T}^{\text{adm.}}$, then for any $\rho > 0$,*

$$\sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta A_t| \leq C_0 \varepsilon_n^{2\alpha - \rho}.$$

3.3 Convergence results for quadratic variation

We first recall a convergence result about weighted discrete quadratic M -variations corresponding to $\mathcal{T} = \{\mathcal{T}^n, n \geq 0\}$.

Proposition 3.7. *[GL14a, Proposition 2.3] Assume (\mathbf{H}_M) and let \mathcal{T} be a sequence of strategies satisfying $(\mathbf{A}_M^{\text{osc.}})$. Let $(H_t)_{0 \leq t < T}$ be a continuous adapted $d \times d$ -matrix process such that $\sup_{t \in [0, T]} |H_t| < +\infty$ a.s., and let $(K_t)_{0 \leq t \leq T}$ be a \mathbb{R}^d -valued continuous local martingale such that $\langle K \rangle_t = \int_0^t \kappa_r dr$ with $\sup_{t \in [0, T]} |\kappa_t| < +\infty$ a.s.. Then*

$$\sum_{\tau_{i-1}^n < T} \Delta K_{\tau_i^n}^\top H_{\tau_{i-1}^n} \Delta K_{\tau_i^n} \xrightarrow{\text{a.s.}} \int_0^T \text{Tr}(H_t d\langle K \rangle_t).$$

We now establish an extension to the semimartingale S .

Proposition 3.8. *Let S be a semimartingale of the form (2.2) and satisfying (\mathbf{H}_A) - (\mathbf{H}_M) , and let \mathcal{T} be a sequence of strategies satisfying $(\mathbf{A}_S^{\text{osc.}})$. Let $(H_t)_{0 \leq t < T}$ as in Proposition 3.7. Then*

$$\sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n}^\top H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \xrightarrow{\text{a.s.}} \int_0^T \text{Tr}(H_t d\langle M \rangle_t).$$

Proof. From Itô's lemma, the difference between the above left hand side and the right one is equal to

$$\int_0^T \Delta S_t^\top (H_{\varphi(t)} + H_{\varphi(t)}^\top) dS_t + \int_0^T \text{Tr}([H_{\varphi(t)} - H_t] d\langle M \rangle_t). \quad (3.12)$$

Due to (\mathbf{H}_M) , the second term is bounded by $C_0 \int_0^T |H_{\varphi(t)} - H_t| dt$: it converges to 0 by an application of the dominated convergence theorem. Indeed, H is continuous and bounded on $[0, T)$ and the mesh size goes to 0 under $(\mathbf{A}_S^{\text{osc.}})$ (see (3.7) which is established under $(\mathbf{A}_S^{\text{osc.}})$ and without using (\mathbf{A}_N)). Next, decompose the first term of (3.12) into stochastic integrals w.r.t. A and M . On the one hand, A is of finite variation, thus

$$\left| \int_0^T \Delta S_t^\top (H_{\varphi(t)} + H_{\varphi(t)}^\top) dA_t \right| \leq C_0 \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta S_t| \sup_{t \in [0, T]} |H_t| \xrightarrow{\text{a.s.}} 0 \quad (3.13)$$

in view of $(\mathbf{A}_S^{\text{osc.}})$. On the other hand, $\int_0^T \Delta S_t^\top (H_{\varphi(t)} + H_{\varphi(t)}^\top) dM_t \xrightarrow{\text{a.s.}} 0$ by proceeding very similarly to the proof of (3.6). \square

In the next theorems we identify an important admissible sequence of strategies, namely hitting times by S of random ellipsoids parametrized by a matrix process $(H_t)_{0 \leq t < T}$ (or a perturbation of it). This extends [GL14a, Proposition 2.4] to hitting times of S and to possibly degenerate H . This more general construction of ellipsoids is a significant improvement, and crucial for the subsequent optimality results.

Theorem 3.9. *Let S be a semimartingale of the form (2.2) and satisfying (\mathbf{H}_A) -(\mathbf{H}_M), and let $(H_t)_{0 \leq t < T}$ be a continuous adapted symmetric non-negative definite $d \times d$ matrix process, such that a.s.*

$$0 < \inf_{0 \leq t < T} \lambda_{\min}(H_t) \leq \sup_{0 \leq t < T} \lambda_{\max}(H_t) < +\infty.$$

The strategy \mathcal{T}^n given by

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n := \inf\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^\top H_{\tau_{i-1}^n} (S_t - S_{\tau_{i-1}^n}) \geq \varepsilon_n^2\} \wedge T, \end{cases}$$

defines a admissible sequence of strategies.

The proof is given later. The condition $\sup_{0 \leq t < T} \lambda_{\max}(H_t) < +\infty$ ensures that none of the corresponding ellipsoids $\mathcal{E}_t := \{x^\top H_t x \leq c\}$ with $c > 0$ are flat in some directions, it allows to derive a bound on the number of hitting times N_T^n as in (\mathbf{A}_N) . The non-degeneracy condition $\lambda_{\min}(H_t) > 0$ (i.e. \mathcal{E}_t is bounded) is important to control the increments ΔS as in $(\mathbf{A}_S^{\text{osc}})$. Without this latter condition, we need to perturb the above sequence of strategies. To this purpose, let $\chi(\cdot)$ be a smooth function such that

$$\mathbf{1}_{(-\infty, 1/2]} \leq \chi(\cdot) \leq \mathbf{1}_{(-\infty, 1]}, \quad (3.14)$$

and for $\mu > 0$ set $\chi_\mu(x) = \chi(x/\mu)$.

Theorem 3.10. *Let S be a semimartingale of the form (2.2) and satisfying (\mathbf{H}_A) -(\mathbf{H}_M). Assume that ρ_N defined in (2.3) is such that $\rho_N > 1$, and let $\delta \in (0, 2(\rho_N - 1)]$. Let $(H_t)_{0 \leq t < T}$ be an adapted symmetric non-negative definite $d \times d$ matrix process, such that*

(i) *there exists a random variable C_H , positive and finite a.s., such that*

$$\lambda_{\max}(H_t) \leq C_H, \quad \forall t \in [0, T), \quad \text{a.s..}$$

(notice that H is not necessarily continuous).

Define a sequence of processes H^n by

$$H_t^{(n)} = H_t + \varepsilon_n^\delta \chi_{\varepsilon_n^\delta}(\lambda_{\min}(H_t)) \text{Id}_d.$$

Then the strategy \mathcal{T}^n defined by

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n := \inf\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^\top H_{\tau_{i-1}^n}^{(n)} (S_t - S_{\tau_{i-1}^n}) \geq \varepsilon_n^{2+\delta}\} \wedge T, \end{cases} \quad (3.15)$$

forms a sequence $\mathcal{T} = \{\mathcal{T}^n : n \geq 0\}$ satisfying the assumption $(\mathbf{A}_S^{\text{osc}})$. If in addition the following convergence holds

(ii)

$$\sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n}^\top H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \xrightarrow{\text{a.s.}} \int_0^T \text{Tr}(H_t d\langle M \rangle_t),$$

then the sequence \mathcal{T} satisfies also the assumption (\mathbf{A}_N) , that is $\mathcal{T} \in \mathcal{T}^{\text{adm}}$.

Proof of Theorem 3.10. First let us prove that \mathcal{T}_n is a.s. of finite size for any $n \in \mathbb{N}$. The definition of H_t^n implies that

$$\lambda_{\max}(H_t^{(n)}) \leq C_H + \sup_{n \geq 0} \varepsilon_n^\delta < +\infty, \quad \forall t \in [0, T) \quad \text{a.s.}$$

Define the event $\mathcal{N}^n := \{\omega : N_T^n(\omega) = +\infty\}$. For $\omega \in \mathcal{N}^n$ the infinite sequence $(\tau_i^n(\omega))$ is increasing and bounded, thus converges. Hence on $\mathcal{N}^n \cap E_S$, with

$$E_S = \{(S_t)_{t \in [0, T]} \text{ is continuous and } C_H < +\infty\},$$

we have

$$\begin{aligned} 0 < \varepsilon_n^{2+\delta} &= (S_{\tau_i^n} - S_{\tau_{i-1}^n})^\top H_{\tau_{i-1}^n}^{(n)} (S_{\tau_i^n} - S_{\tau_{i-1}^n}) \\ &\leq \left(C_H + \sup_{n \geq 0} \varepsilon_n^\delta \right) |S_{\tau_i^n} - S_{\tau_{i-1}^n}|^2 \xrightarrow{i \rightarrow +\infty} 0, \end{aligned}$$

which is impossible. Hence $\mathbb{P}(\mathcal{N}^n \cap E_S) = 0$, but $\mathbb{P}(E_S) = 1$ thus $\mathbb{P}(\mathcal{N}^n) = 0$.

Next we show that \mathcal{T} satisfies $(\mathbf{A}_S^{\text{osc}})$. From the definition of $H_t^{(n)}$ it is straightforward that

$$\lambda_{\min}(H_t^{(n)}) \geq \frac{\varepsilon_n^\delta}{2}, \quad \forall t \in [0, T).$$

Thus

$$\begin{aligned} &\varepsilon_n^{-2} \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta S_t|^2 \\ &\leq \left(\inf_{t \in [0, T)} \lambda_{\min}(H_t^{(n)}) \right)^{-1} \varepsilon_n^{-2} \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} (\Delta S_t^\top H_t^{(n)} \Delta S_t) \leq 2\varepsilon_n^{-\delta} \varepsilon_n^{-2} \varepsilon_n^{2+\delta} = 2, \end{aligned}$$

which validates the assumption $(\mathbf{A}_S^{\text{osc}})$.

Finally assume that in addition (ii) holds and let us show that the sequence of strategies \mathcal{T} satisfies the assumption (\mathbf{A}_N) . Writing $N_T^n = 1 + \sum_{1 \leq i \leq N_T^n - 1} 1$ and using $2 + \delta \leq 2\rho_N$, we observe that (for n large enough so that $\varepsilon_n \leq 1$)

$$\varepsilon_n^{2\rho_N} N_T^n \leq \varepsilon_n^{2+\delta} N_T^n \leq \varepsilon_n^{2+\delta} + \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} + \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top (H_{\tau_{i-1}^n}^{(n)} - H_{\tau_{i-1}^n}) \Delta S_{\tau_i^n}. \quad (3.16)$$

Now by (ii) we have

$$\sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \text{Tr}(H_t d\langle M \rangle_t) \stackrel{a.s.}{<} +\infty$$

(the contribution $i = N_T^n$ does not change the convergence). Besides from the definition of $H^{(n)}$ we get

$$\left| \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top (H_{\tau_{i-1}^n}^{(n)} - H_{\tau_{i-1}^n}) \Delta S_{\tau_i^n} \right| \leq \varepsilon_n^\delta \sum_{\tau_i^n < T} |\Delta S_{\tau_i^n}|^2 \xrightarrow{a.s.} 0, \quad (3.17)$$

using $\delta > 0$ and Proposition 3.8 (valid since $(\mathbf{A}_S^{\text{osc.}})$ is in force now). We have proved that the r.h.s. of (3.16) converges a.s. to a finite random variable, which completes the verification of the assumption (\mathbf{A}_N) . \square

Proof of Theorem 3.9. This is an adaptation of the previous proof. First, with the same arguments we prove that \mathcal{T}_n is a.s. of finite size for any $n \in \mathbb{N}$. Second, the verification of $(\mathbf{A}_S^{\text{osc.}})$ stems from

$$\varepsilon_n^{-2} \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta S_t|^2 \leq \left(\inf_{t \in [0, T)} \lambda_{\min}(H_t) \right)^{-1}.$$

Third, for n large enough so that $\varepsilon_n \leq 1$, we write

$$\varepsilon_n^{2\rho_N} N_T^n \leq \varepsilon_n^2 N_T^n \leq \varepsilon_n^2 + \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top H_{\tau_{i-1}^n} \Delta S_{\tau_i^n}$$

and we conclude to (\mathbf{A}_N) using Proposition 3.8 and the continuity and boundedness of H . \square

4 Asymptotic lower bound on the discretization error

Let S be a semimartingale of the form (2.2) and let v be the function appearing in the discretization error (1.1), and satisfying (\mathbf{H}_v) . The main result of the

section is Theorem 4.2: this is an extension to the semimartingale case of the asymptotic lower bound on the discretization error, proved in [GL14a, Theorem 3.1] in the martingale case.

The discretization error Z^n defined in (1.1) can be decomposed into a martingale part and a finite variation part:

$$Z_s^n = \int_0^s (v(t, S_t) - v(\phi(t), S_{\phi(t)})) \cdot dM_t + \int_0^s (v(t, S_t) - v(\phi(t), S_{\phi(t)})) \cdot dA_t.$$

The analysis is partially derived from a smart representation of $\langle Z^n \rangle_T$ as a sum of squared random variables and an adequate application of Cauchy-Schwarz inequality. The derivation of such a representation is based on applying the Itô formula to a *suitable function* and identifying the bounded variation term. While it is straightforward in dimension one, a multidimensional version of this result requires to solve the following matrix equation.

Lemma 4.1. *Let c be a $d \times d$ -matrix with real-valued entries. Then the equation*

$$2 \operatorname{Tr}(x)x + 4x^2 = cc^\top \quad (4.1)$$

admits exactly one solution $x(c) \in \mathcal{S}_+^d(\mathbb{R})$. Moreover, the mapping $c \mapsto x(c)$ is continuous.

The proof of the above lemma directly follows from [GL14a, Lemma 3.1] applied for $(cc^\top)^{1/2}$ (i.e. the symmetric non-negative definite square root of cc^\top). Now we state the main result.

Theorem 4.2 (Lower bound). *Assume (\mathbf{H}_A) , (\mathbf{H}_M) , (\mathbf{H}_v) and let \mathcal{T} be an admissible sequence of strategies (satisfying (\mathbf{A}_N) and $(\mathbf{A}^{\text{osc}})$). Let X be the continuous adapted symmetric non-negative definite matrix process solution of (4.1) with $c = \sigma^\top(D_x v)^\top \sigma$, i.e.*

$$X_t := x(\sigma_t^\top(D_x v_t)^\top \sigma_t), \quad \text{for } 0 \leq t < T. \quad (4.2)$$

Then we have

$$\liminf_{n \rightarrow +\infty} N_T^n \langle Z^n \rangle_T \geq \left(\int_0^T \operatorname{Tr}(X_t) dt \right)^2 \quad \text{a.s..}$$

Proof. The martingale part of the discretization error can be written

$$\int_0^s (v(t, S_t) - v(\phi(t), S_{\phi(t)})) \cdot dM_t =: \int_0^s (D_x v_{\phi(t)} \Delta S_t) dM_t + R_s^n. \quad (4.3)$$

Therefore the quadratic variation of Z^n is given by

$$\begin{aligned} \langle Z^n \rangle_T &= \int_0^T \Delta S_t^\top (D_x v_{\phi(t)})^\top d\langle M \rangle_t D_x v_{\phi(t)} \Delta S_t + e_{1,T}^n \\ &= \int_0^T \Delta M_t^\top (D_x v_{\phi(t)})^\top d\langle M \rangle_t D_x v_{\phi(t)} \Delta M_t + e_{1,T}^n + e_{0,T}^n, \end{aligned} \quad (4.4)$$

where

$$\begin{aligned} e_{0,T}^n &:= \int_0^T \Delta A_t^\top (D_x v_{\phi(t)})^\top d\langle M \rangle_t D_x v_{\phi(t)} (\Delta S_t + \Delta M_t), \\ e_{1,T}^n &:= \langle R^n \rangle_T + 2 \left\langle \int_0^T (D_x v_{\phi(t)} \Delta M_t) \cdot dM_t, R^n \right\rangle_T. \end{aligned}$$

Now in the first contribution of $\langle Z^n \rangle_T$ in (4.4), we seek an expression involving only the Brownian motion B and not the local martingale M : hence we replace ΔM_t by $\sigma_{\phi(t)} \Delta B_t$ and $d\langle M \rangle_t$ by $\sigma_{\phi(t)} \sigma_{\phi(t)}^\top dt$, which leads to

$$\langle Z^n \rangle_T = \int_0^T \Delta B_t^\top (\sigma_{\phi(t)}^\top (D_x v_{\phi(t)})^\top \sigma_{\phi(t)} \sigma_{\phi(t)}^\top D_x v_{\phi(t)} \sigma_{\phi(t)}) \Delta B_t dt + e_{0,T}^n + e_{1,T}^n + e_{2,T}^n,$$

where

$$\begin{aligned} e_{2,T}^n &:= \int_0^T \Delta M_t^\top (D_x v_{\phi(t)})^\top \Delta (\sigma_t \sigma_t^\top) D_x v_{\phi(t)} \Delta M_t dt \\ &\quad + \int_0^T (\Delta M_t + \sigma_{\phi(t)} \Delta B_t)^\top (D_x v_{\phi(t)})^\top \sigma_{\phi(t)} \sigma_{\phi(t)}^\top D_x v_{\phi(t)} (\Delta M_t - \sigma_{\phi(t)} \Delta B_t) dt. \end{aligned}$$

Denote $C_t = \sigma_t^\top (D_x v_t)^\top \sigma_t$. We seek a smart representation of the main term of $\langle Z^n \rangle_T$ in the form

$$\sum_{\tau_{i-1}^n < T} (\Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n})^2, \quad (4.5)$$

where X is a suitable measurable adapted symmetric $d \times d$ -matrix process. For such a process X , the Itô formula on each interval $[\tau_{i-1}^n, \tau_i^n]$ yields

$$\begin{aligned} \sum_{\tau_{i-1}^n < T} (\Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n})^2 &= \int_0^T \Delta B_t^\top (2 \operatorname{Tr}(X_{\phi(t)}) X_{\phi(t)} + 4 X_{\phi(t)}^2) \Delta B_t dt \\ &\quad + 4 \int_0^T \Delta B_t^\top X_{\phi(t)} \Delta B_t \Delta B_t^\top X_{\phi(t)} dB_t. \end{aligned}$$

Now take X as stated in the theorem. Clearly $X_t \in \mathcal{S}_+^d(\mathbb{R})$ owing to Lemma 4.1. The continuity of the mapping $c \mapsto x(c)$ also ensures that X is continuous and adapted, as $\sigma^\top (D_x v)^\top \sigma$ is. Then a simplified representation of $\langle Z^n \rangle_T$ readily follows:

$$\langle Z^n \rangle_T = \sum_{\tau_{i-1}^n < T} (\Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n})^2 + e_{0,T}^n + e_{1,T}^n + e_{2,T}^n + e_{3,T}^n, \quad (4.6)$$

where

$$e_{3,T}^n := -4 \int_0^T \Delta B_t^\top X_{\phi(t)} \Delta B_t \Delta B_t^\top X_{\phi(t)} dB_t.$$

Using Cauchy-Schwarz inequality and $X_t \in \mathcal{S}_+^d(\mathbb{R})$, we obtain

$$N_T^n \sum_{\tau_{i-1}^n < T} (\Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n})^2 \geq \left(\sum_{\tau_{i-1}^n < T} \Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n} \right)^2.$$

The process X_t is a.s. continuous on $[0, T)$, with $\sup_{t \in [0, T)} |X_t| < +\infty$ a.s., and thus the assumptions of Proposition 3.7 are satisfied for $(H, K) = (X, B)$. Therefore

$$\sum_{\tau_i^n < T} \Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \text{Tr}(X_t) dt.$$

To summarize we have obtained that

$$\liminf_{n \rightarrow +\infty} (N_T^n \langle Z^n \rangle_T - N_T^n (e_{0,T}^n + e_{1,T}^n + e_{2,T}^n + e_{3,T}^n)) \geq \left(\int_0^T \text{Tr}(X_t) dt \right)^2 \quad \text{a.s..}$$

To complete the proof, it is enough to show that $N_T^n (e_{0,T}^n + e_{1,T}^n + e_{2,T}^n + e_{3,T}^n) \xrightarrow{a.s.} 0$. In view of the assumption (\mathbf{A}_N) it is sufficient to prove that

$$\varepsilon_n^{-2\rho_N} e_{i,T}^n \xrightarrow{a.s.} 0 \quad \text{for } i = 0, 1, 2, 3. \quad (4.7)$$

Contribution $e_{0,T}^n$. Owing to Corollary 3.6, we obtain immediately that

$$|e_{0,T}^n| \leq C_0 \int_0^T |\Delta A_t| (|\Delta S_t| + |\Delta M_t|) dt \leq C_0 \varepsilon_n^{1+2\alpha-\rho},$$

for any $\rho > 0$, which implies $\varepsilon_n^{-2\rho_N} e_{0,T}^n \rightarrow 0$ since $\rho_N < \frac{1}{2} + \alpha$.

Contribution $e_{1,T}^n$. To handle it, we need the following lemma; its proof follows that of [GL14a, Lemma 3.2], with minor adaptations (see Appendix A.1).

Lemma 4.3. *Under the assumptions (\mathbf{H}_A) , (\mathbf{H}_M) , (\mathbf{H}_v) , $(\mathbf{A}^{\text{osc.}})$ and (\mathbf{A}_N) , we have $\varepsilon_n^{2-4\rho_N} \langle R^n \rangle_T \xrightarrow{a.s.} 0$, where R^n is defined in (4.3).*

Now to show that $\varepsilon_n^{-2\rho_N} e_{1,T}^n \rightarrow 0$, use the above lemma and $(\mathbf{A}_M^{\text{osc.}})$ to get

$$\varepsilon_n^{-2\rho_N} |e_{1,T}^n| \leq \varepsilon_n^{-2\rho_N} (\langle R^n \rangle_T + 2C_0 \varepsilon_n (\langle R^n \rangle_T)^{1/2}) = o(\varepsilon_n^{2\rho_N-2}) + o(1) \xrightarrow{a.s.} 0.$$

Contributions $e_{2,T}^n$ and $e_{3,T}^n$. The proof is similar to that of [GL14a, Theorem 3.1], we skip the details. \square

5 Optimal strategy

5.1 Preliminaries, pseudo-inverses

Now our main purpose is to provide, in notation of Theorem 4.2, an optimal discretization strategy, i.e. an admissible strategy \mathcal{T} for which

$$\lim_{n \rightarrow +\infty} N_T^n \langle Z^n \rangle_T = \left(\int_0^T \text{Tr}(X_t) dt \right)^2 \quad \text{a.s..}$$

Notice that an existence result is proved in [GL14a, Theorem 3.3], only under the conditions that σ is invertible, that $v(t, x) = \nabla_x u(t, x)$ with

$$\inf_{0 \leq t < T} \lambda_{\min}(D_{xx}^2 u(t, S_t)) > 0 \quad \text{a.s.}$$

and that $A = 0$ (martingale case). Our aim here is to relax these three conditions, and to extend the ideas of this aforementioned theorem to our general setting.

Actually, the main difficulty comes from the possible degeneracy of σ . First recall the definition and some properties of pseudo-inverse matrix (a.k.a. Moore-Penrose generalized inverse).

Definition 2 (pseudo-inverse of a matrix). *Let \mathcal{M} be a real-valued $d \times d$ -matrix. Consider the singular value decomposition of \mathcal{M}*

$$\mathcal{M} = U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V^\top,$$

where U, V are both orthogonal matrices, and D is a diagonal matrix containing the (positive) singular values of \mathcal{M} on its diagonal. Then the pseudo-inverse of \mathcal{M} is the $d \times d$ -matrix defined as

$$\mathcal{M}^\dagger = V \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^\top.$$

We recall the following well-known properties, which can be easily checked from Definition 2:

$$\begin{cases} \mathcal{M}\mathcal{M}^\dagger\mathcal{M} = \mathcal{M}, & \mathcal{M}^\dagger\mathcal{M}\mathcal{M}^\dagger = \mathcal{M}^\dagger, \\ \text{the matrices } \mathcal{M}\mathcal{M}^\dagger \text{ and } \mathcal{M}^\dagger\mathcal{M} \text{ are symmetric.} \end{cases} \quad (5.1)$$

5.2 Main result

We wish to design optimal stopping times in terms of the process S to allow better tractability. Inspired by [GL14a], a good candidate is then the sequence $\{\mathcal{T}^n : n \geq 0\}$ where \mathcal{T}^n is defined as:

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n := \inf\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^\top \Lambda_{\tau_{i-1}^n} (S_t - S_{\tau_{i-1}^n}) \geq \varepsilon_n^2\} \wedge T, \end{cases} \quad (5.2)$$

where $\Lambda_t := (\sigma_t^{-1})^\top X_t \sigma_t^{-1}$ with X given by (4.2).

Such a sequence turns out to be optimal when S is a martingale and under some additional assumptions (see [GL14a, Theorem 3.3]). The problems with this definition can arise if σ_t is not invertible, or if Λ_t is degenerate for some values of t (then we have difficulties to verify $(\mathbf{A}^{\text{osc}})$). To overcome these problems we use σ_t^\dagger instead of σ_t^{-1} . Furthermore we take $\Lambda_t^{(n)}$ equal to a small perturbation of Λ_t depending on ε_n , such that $\Lambda_t^{(n)}$ is always non-degenerate.

We need one additional assumption.

(H_Λ) Let $(X_t)_{0 \leq t < T}$ be defined in (4.2) and consider the $\mathcal{S}_+^d(\mathbb{R})$ -valued process defined by

$$\Lambda_t := (\sigma_t^\dagger)^\top X_t \sigma_t^\dagger, \quad \forall t \in [0, T]. \quad (\mathbf{H}_\Lambda)$$

There exists a non-negative random variable $c_{(5.3)}$, finite a.s., such that

$$0 \leq \text{Tr}(\Lambda_t) \leq c_{(5.3)}, \quad \forall t \in [0, T], \quad \text{a.s.} \quad (5.3)$$

Note that σ^\dagger may be discontinuous, so Λ may be too. Recall (see (3.14)) that $\chi(\cdot)$ stands for a smooth function such that $\mathbf{1}_{(-\infty, 1/2]} \leq \chi(\cdot) \leq \mathbf{1}_{(-\infty, 1]}$, and for $\mu > 0$, we set $\chi_\mu(x) = \chi(x/\mu)$. Now we state the precise definition of an optimal sequence of strategies.

Theorem 5.1 (Optimal strategy). *Assume that (\mathbf{H}_A) , (\mathbf{H}_M) , (\mathbf{H}_v) , (\mathbf{H}_Λ) are in force. Let ρ_N satisfy (2.3) with $\rho_N > 1$, and let $\delta \in (0, 2(\rho_N - 1)]$. For each $n \in \mathbb{N}$, define the process $(\Lambda_t^{(n)} : t < T)$ by*

$$\Lambda_t^{(n)} = \Lambda_t + \varepsilon_n^\delta \chi_{\varepsilon_n^\delta}(\lambda_{\min}(\Lambda_t)) \text{Id}_d$$

where Λ is given in (\mathbf{H}_Λ) , and define the strategy $\mathcal{T}_{\varepsilon_n^\delta}^n$ by

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n := \inf\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^\top \Lambda_{\tau_{i-1}^n}^{(n)} (S_t - S_{\tau_{i-1}^n}) \geq \varepsilon_n^{2+\delta}\} \wedge T. \end{cases} \quad (5.4)$$

Then the sequence of strategies $\mathcal{T} = \{\mathcal{T}_{\varepsilon_n^\delta}^n : n \geq 0\}$ is admissible for the parameter ρ_N and is asymptotically optimal, i.e.

$$\lim_{n \rightarrow +\infty} N_T^n \langle Z^n \rangle_T = \left(\int_0^T \text{Tr}(X_t) dt \right)^2 \quad \text{a.s.}$$

To conclude this subsection, we provide a condition simpler than (\mathbf{H}_Λ) , the proof is postponed to the end of this section.

Proposition 5.2. *Assume that (\mathbf{H}_A) , (\mathbf{H}_M) , (\mathbf{H}_v) are in force, and assume that $v \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d)$ so that $D_x v_t$ and X_t can be defined continuously up to $t = T$. If the matrix*

$$C_t := \sigma_t^\top (D_x v_t)^\top \sigma_t \neq 0 \quad (\mathbf{H}_C)$$

for all $t \in [0, T]$ a.s., then (\mathbf{H}_Λ) holds.

5.3 Examples

5.3.1 About the assumptions (\mathbf{H}_Λ) and (\mathbf{H}_C)

Recall that under our assumptions, X is a.s. uniformly bounded on $[0, T]$. Thus in order to satisfy (\mathbf{H}_Λ) , it is enough to have σ^\dagger a.s. uniformly bounded on $[0, T]$. We provide a (non-exhaustive) list of such examples.

- a) σ_t is invertible for any t a.s.: then $\sigma_t^\dagger = \sigma_t^{-1}$ is clearly bounded on $[0, T]$.
- b) We can also afford degenerate cases: for instance if σ_t is constant in time (but possibly with $\text{rank}(\sigma_t) < d$), then σ_t^\dagger is also constant in time (and thus bounded).
- c) The previous principle can be generalized to the time-dependent case $\sigma_t = \begin{pmatrix} \Sigma_t & 0 \\ 0 & 0 \end{pmatrix}$ where Σ_t is a square matrix, a.s. invertible at any time: indeed $\sigma_t^\dagger = \begin{pmatrix} \Sigma_t^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ is bounded on $[0, T]$.

Now, we argue that checking (\mathbf{H}_C) may be sometimes much simpler than the verification of (\mathbf{H}_A) . Let us give a non-trivial example where σ^\dagger is not continuous a.s. For the i -th component of S , take a squared δ_i -dimensional radial Ornstein-Uhlenbeck process with parameter $-\lambda_i$, which is the strong solution to

$$S_t^i = S_0^i + \int_0^t (\delta_i - \lambda_i S_s^i) ds + 2 \int_0^t \sqrt{S_s^i} dB_s^i,$$

where $S_0^i > 0$, $\delta_i \geq 0$, $\lambda_i \in \mathbb{R}$ (see [GJY03]). The matrix σ_t is diagonal and its i -th element is equal to $2\sqrt{S_t^i}$. It is easy to check that (\mathbf{H}_A) and (\mathbf{H}_M) hold (in particular $\sigma_t \neq 0$ for all t a.s.). The pseudo-inverse σ_t^\dagger is diagonal with i -th element equal to $[2\sqrt{S_t^i}]^{-1} \mathbf{1}_{S_t^i > 0}$. Assume now that one of the δ_i is strictly smaller than 2: then the associated component S^i has a positive probability to hit 0 before T . As a consequence, with positive probability, σ^\dagger is unbounded on $[0, T]$ and it is not clear anymore to check directly (\mathbf{H}_A) . Alternatively, assume (again to simplify) that $D_x v_t \in \mathcal{S}_{++}^d(\mathbb{R})$. Then $C_t \neq 0$: indeed, $C_t \in \mathcal{S}_+^d(\mathbb{R})$ and $\text{Tr}(C_t) = \text{Tr}(D_x v_t \sigma_t \sigma_t^\top) > 0$ since $\sigma_t \sigma_t^\top \neq 0$ and $D_x v_t$ is invertible.

5.3.2 A numerical example

We consider a two-dimensional example, defined by

$$S_t = \begin{pmatrix} B_t^1 + 0.3B_t^2 \\ \int_0^t B_s^1 ds \end{pmatrix}.$$

It corresponds to a constant (degenerate) matrix

$$\sigma_t = \begin{pmatrix} 1 & 0.3 \\ 0 & 0 \end{pmatrix}.$$

For the function v we take

$$v(t, x) = \begin{pmatrix} \cos(3x_1) \\ \cos(3x_2) \end{pmatrix},$$

and we set $T = 1$. According to the previous paragraph, (\mathbf{H}_A) is satisfied and an optimal sequence of strategies is given by Theorem 5.1. To assess the efficiency of an arbitrary admissible sequence of strategies we set

$$\alpha_n := \frac{N_T^n \langle Z^n \rangle_T}{\left(\int_0^T \text{Tr}(X_t) dt \right)^2} \quad \text{and} \quad \beta_n := \sqrt{N_T^n} Z_T^n.$$

From Theorem 4.2 we must have $\liminf_{n \rightarrow +\infty} \alpha_n \geq 1$ a.s., while for the optimal sequence the equality holds. The normalized error β_n is also important in practice, however we cannot in general asymptotically control a.s. this quantity. But it is easy to believe that the values of β_n are smaller for strategies where the corresponding values of α_n are smaller, at least in mean. We will illustrate this heuristics in the following.

Namely, we denote by $\alpha_{n,opt}, \beta_{n,opt}$ the values of α_n, β_n for the optimal strategy defined in Theorem 5.1, for several simulations. For a given number of grid times N_T^n of the optimal strategy, we also consider the deterministic strategy with N_T^n equidistant times and denote by $\alpha_{n,det}, \beta_{n,det}$ the values of α_n, β_n for this strategy. This way of proceeding is aimed at comparing the performance of these two strategies with the same number of time points.

To compute the hitting times we use a thin uniform time mesh with $\bar{n} = 10000$ points. Figure 1 shows the pairs $(\alpha_{n,opt}, \beta_{n,opt})$ and $(\alpha_{n,det}, \beta_{n,det})$ for ten simulations, with $\varepsilon_n = 0.04$ and $\varepsilon_n^\delta = 0.1$. As expected from Theorems 4.2 and 5.1, we observe the inequality $\alpha_{n,opt} < \alpha_{n,det}$ and the limit $\alpha_{n,opt} \approx 1$. Moreover, the inequality $|\beta_{n,opt}| < |\beta_{n,det}|$ holds as well for most of the simulations.

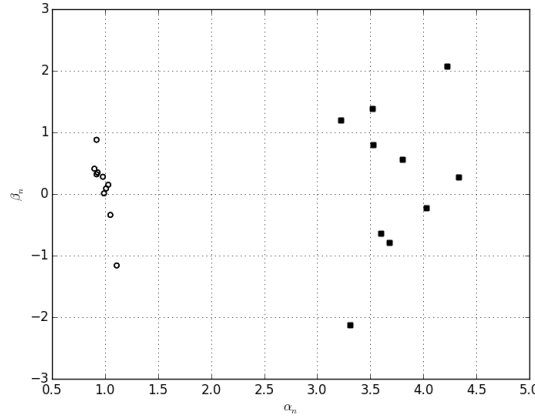


Figure 1: The pairs $(\alpha_{n,det}, \beta_{n,det})$ and $(\alpha_{n,opt}, \beta_{n,opt})$ are represented by squares and circles respectively.

5.4 Proof of Theorem 5.1

The proof is divided into several steps. Assumptions of Theorem 5.1 are in force in all this subsection.

5.4.1 Step 1: a reverse relation between X and Λ

Proposition 5.3. *The following equality holds*

$$X_t = (\sigma_t)^\top \Lambda_t \sigma_t, \quad \forall t \in [0, T) \quad \text{a.s.} \quad (5.5)$$

Proof. We are going to establish the above relation for any given t , with probability 1: however, the reader can check that the negligible set can be the same for all t (as for the definitions of σ, X, Λ) because the arguments used are of deterministic nature.

If σ_t is invertible, $\sigma_t^\dagger = \sigma_t^{-1}$ and obviously $X_t = (\sigma_t)^\top \Lambda_t \sigma_t$ in view of the definition (**H_Λ**).

Now assume that $\text{rank}(\sigma_t) < d$. By (5.1) we have

$$\sigma_t \sigma_t^\dagger \sigma_t = \sigma_t \quad (5.6)$$

and the matrix $\sigma_t^\dagger \sigma_t$ is symmetric. We choose an orthonormal basis $(e_i)_{1 \leq i \leq d}$ under which the matrix $\sigma_t^\dagger \sigma_t$ is diagonal, i.e.

$$\sigma_t^\dagger \sigma_t = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_d \end{pmatrix}$$

for some $\alpha_1, \dots, \alpha_d$. If $\sigma_t^1, \dots, \sigma_t^d$ are the column vectors of σ_t (in the basis $(e_i)_{1 \leq i \leq d}$), then from (5.6) we get

$$(\alpha_1 \sigma_t^1, \dots, \alpha_d \sigma_t^d) = (\sigma_t^1, \dots, \sigma_t^d). \quad (5.7)$$

For any $1 \leq i \leq d$ if $\sigma_t^i \neq 0$ then we must have $\alpha_i = 1$. On the other hand $k := \text{rank}(\sigma_t^\dagger \sigma_t) \leq \text{rank}(\sigma_t) < d$. Hence by permuting the basis elements and using (5.6) we can write $\sigma_t^\dagger \sigma_t$ and σ_t in the form:

$$\sigma_t^\dagger \sigma_t = \begin{pmatrix} \text{Id}_k & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma_t = \begin{pmatrix} \sigma_{1,t}^1 & \dots & \sigma_{1,t}^k & 0 & \dots & 0 \\ \sigma_{2,t}^1 & \dots & \sigma_{2,t}^k & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ \sigma_{d,t}^1 & \dots & \sigma_{d,t}^k & 0 & \dots & 0 \end{pmatrix}. \quad (5.8)$$

We want to show that $X_t = (\sigma_t)^\top \Lambda_t \sigma_t$ which by the definition of Λ_t is equivalent to

$$X_t = (\sigma_t^\dagger \sigma_t)^\top X_t (\sigma_t^\dagger \sigma_t) = (\sigma_t^\dagger \sigma_t) X_t (\sigma_t^\dagger \sigma_t). \quad (5.9)$$

In view of (5.8) and since X is symmetric non-negative definite, the equality (5.9) is equivalent to the following system of equations:

$$e_i^\top X_t e_i = 0 \quad \text{for } i = k+1, \dots, d, \quad (5.10)$$

where (e_i) are the vectors of the basis. We now prove (5.10). Let $i \in \{k+1, \dots, d\}$. From the definition of X_t we get

$$2 \operatorname{Tr}(X_t) X_t + 4 X_t^2 = \sigma_t^\top \tilde{C}_t \sigma_t, \quad (5.11)$$

where $\tilde{C}_t = (D_x v_t)^\top \sigma_t \sigma_t^\top D_x v_t$. From (5.8) it is clear that $\sigma_t e_i = 0$, thus Equation (5.11) yields

$$2 \operatorname{Tr}(X_t) e_i^\top X_t e_i + 4 e_i^\top X_t^2 e_i = 0.$$

Both X_t and X_t^2 are in $\mathcal{S}_+^d(\mathbb{R})$, thus both above terms are non-negative, therefore they are equal to 0. Either $\operatorname{Tr}(X_t) = 0$ (implying $X_t = 0$ and (5.10)), or $\operatorname{Tr}(X_t) > 0$ and $e_i^\top X_t e_i = 0$. In any case, (5.10) holds and we are done. \square

5.4.2 Step 2: verification of $(\mathbf{A}_S^{\text{osc.}})$

The stopping times (5.4) define a sequence of strategy satisfying $(\mathbf{A}_S^{\text{osc.}})$: this is a consequence of Theorem 3.10-(i) with $H = \Lambda$. Indeed the existence of the finite random variable C_H stems from (5.3).

5.4.3 Step 3: verification of (\mathbf{A}_N)

We aim at showing

Proposition 5.4. *We have the following convergence*

$$\sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \xrightarrow{\text{a.s.}} \int_0^T \operatorname{Tr}(\Lambda_t d\langle M \rangle_t) = \int_0^T \operatorname{Tr}(X_t) dt.$$

Then, in view of Theorem 3.10-(ii), we conclude that the sequence of strategies $\mathcal{T} = \{\mathcal{T}_{\varepsilon_n^\delta}^n : n \geq 0\}$ satisfies (\mathbf{A}_N) . Combined with Step 2, we have proved that this is an admissible sequence.

Observe that the above result is not a particular case of Proposition (3.7) since we do not know if Λ is continuous in time (it is likely not for degenerate σ). To handle this difficulty, we are going to leverage the reverse relation between X and Λ (Step 1), and the continuity of X .

Proof of Proposition 5.4. By Itô's lemma like for (3.12) and using that Λ is symmetric, we obtain

$$\sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} = 2 \int_0^T \Delta S_t^\top \Lambda_{\phi(t)} dS_t + \int_0^T \operatorname{Tr}(\Lambda_{\phi(t)} d\langle M \rangle_t). \quad (5.12)$$

Then

$$\begin{aligned} \int_0^T \text{Tr}(\Lambda_{\phi(t)} d\langle M \rangle_t) &= \int_0^T \text{Tr}(\sigma_t^\top \Lambda_{\phi(t)} \sigma_t) dt = \int_0^T \text{Tr}(\sigma_{\phi(t)}^\top \Lambda_{\phi(t)} \sigma_{\phi(t)}) dt \\ &+ \int_0^T \text{Tr}((\sigma_t - \sigma_{\phi(t)})^\top \Lambda_{\phi(t)} (\sigma_t + \sigma_{\phi(t)})) dt. \end{aligned}$$

Observe that the first term on the r.h.s. above is equal to $\int_0^T \text{Tr}(X_{\phi(t)}) dt$ owing to Proposition 5.3: since X is a.s. bounded continuous and the time step goes to 0 (see (3.7) valid under $(\mathbf{A}_S^{\text{osc.}})$), we easily obtain $\int_0^T \text{Tr}(X_{\phi(t)}) dt \xrightarrow{\text{a.s.}} \int_0^T \text{Tr}(X_t) dt$.

The second term tends to 0 a.s. thanks to the continuity of σ and the uniform bound (5.3) on Λ . We have proved

$$\int_0^T \text{Tr}(\Lambda_{\phi(t)} d\langle M \rangle_t) \xrightarrow{\text{a.s.}} \int_0^T \text{Tr}(X_t) dt.$$

To complete the proof, in view of (5.12) it remains to show that

$$\int_0^T \Delta S_t^\top \Lambda_{\phi(t)} dS_t \xrightarrow{\text{a.s.}} 0.$$

The a.s.-convergence to 0 of the contribution $\int_0^T \Delta S_t^\top \Lambda_{\phi(t)} dA_t$ is proved as for (3.13), using $(\mathbf{A}_S^{\text{osc.}})$ and (\mathbf{H}_Λ) . The second contribution $K_T^n := \int_0^T \Delta S_t^\top \Lambda_{\phi(t)} dM_t$ is a local martingale, which bracket is bounded by ε_n^2 up to a random finite constant (use again $(\mathbf{A}_S^{\text{osc.}})$ and (\mathbf{H}_Λ)). Consequently, an application of Lemma 3.1 with $p = 2$, ensures that $K_T^n \xrightarrow{\text{a.s.}} 0$. We are done. \square

5.4.4 Final step: completion of proof of Theorem 5.1

So far, we have showed that the strategy $\mathcal{T} = \{\mathcal{T}_{\varepsilon_n^\delta}^{(n)} : n \geq 0\}$ is admissible. We now prove that

$$\lim_{n \rightarrow +\infty} N_T^n \langle Z^n \rangle_T = \left(\int_0^T \text{Tr}(X_t) dt \right)^2 \quad \text{a.s..}$$

First, proceeding as (3.16), we write that $\varepsilon_n^{2+\delta} N_T^n$ equals

$$\varepsilon_n^{2+\delta} + \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} + \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top (\Lambda_{\tau_{i-1}^n}^{(n)} - \Lambda_{\tau_{i-1}^n}) \Delta S_{\tau_i^n}. \quad (5.13)$$

The first term converges to 0, as well as the last term (proceeding as for (3.17)), while the second one converges a.s. to $\int_0^T \text{Tr}(\Lambda_t d\langle M \rangle_t)$ (Proposition 5.4). To summarize, we have justified

$$\lim_{n \rightarrow +\infty} \varepsilon_n^{2+\delta} N_T^n = \int_0^T \text{Tr}(\Lambda_t d\langle M \rangle_t) = \int_0^T \text{Tr}(X_t) dt \quad \text{a.s..} \quad (5.14)$$

Thus it remains to show that

$$\lim_{n \rightarrow +\infty} \varepsilon_n^{-(2+\delta)} \langle Z^n \rangle_T = \int_0^T \text{Tr}(X_t) dt \quad \text{a.s.} \quad (5.15)$$

Starting from (4.6), write $\langle Z^n \rangle_T$ in the form

$$\langle Z^n \rangle_T = \sum_{\tau_{i-1}^n < T} (\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n})^2 + e_{0,T}^n + e_{1,T}^n + e_{2,T}^n + e_{3,T}^n + e_{4,T}^n + e_{5,T}^n,$$

where $e_{0,T}^n, e_{1,T}^n, e_{2,T}^n, e_{3,T}^n$ are defined as in the proof of Theorem 4.2 and the other terms are defined as follows:

$$\begin{aligned} e_{4,T}^n &:= \sum_{\tau_{i-1}^n < T} (\Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n})^2 - \sum_{\tau_{i-1}^n < T} (\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n})^2, \\ e_{5,T}^n &:= \sum_{\tau_{i-1}^n < T} (\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n})^2 - \sum_{\tau_{i-1}^n < T} (\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n})^2. \end{aligned}$$

First notice that for each $i \leq N_T^n - 1$ we have $\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n} = \varepsilon_n^{2+\delta}$, thus

$$\begin{aligned} &\varepsilon_n^{-(2+\delta)} \sum_{\tau_{i-1}^n < T} (\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n})^2 \\ &= \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n} + \varepsilon_n^{-(2+\delta)} (\Delta S_T^\top \Lambda_{\tau_{N_T^n-1}^n}^{(n)} \Delta S_T)^2 \\ &\xrightarrow{\text{a.s.}} \int_0^T \text{Tr}(\Lambda_t d\langle M \rangle_t) = \int_0^T \text{Tr}(X_t) dt, \end{aligned}$$

where the last convergence is derived similarly to that of (5.13).

Moreover, from (4.7) in the proof of Theorem 4.2, we already have (for ε_n small enough so that $\varepsilon_n \leq 1$ and since $2 + \delta \leq 2\rho_N$)

$$\varepsilon_n^{-(2+\delta)} e_{i,T}^n \leq \varepsilon_n^{-2\rho_N} e_{i,T}^n \xrightarrow{\text{a.s.}} 0 \quad \text{a.s. for } i = 0, 1, 2, 3.$$

To complete the proof of Theorem 5.1, it remains only to prove that

$$\varepsilon_n^{-(2+\delta)} e_{i,T}^n \xrightarrow{\text{a.s.}} 0 \quad \text{a.s. for } i = 4, 5.$$

We start with $e_{5,T}^n$:

$$\begin{aligned} |\varepsilon_n^{-(2+\delta)} e_{5,T}^n| &\leq \sum_{\tau_{i-1}^n < T} (\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} + \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}) \\ &\quad \times |\Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} - \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}| \varepsilon_n^{-(2+\delta)} \\ &\leq \sum_{\tau_{i-1}^n < T} \varepsilon_n^\delta \chi_{\varepsilon_n^\delta}(\lambda_{\min}(\Lambda_{\tau_{i-1}^n})) |\Delta S_{\tau_i^n}|^2 |2\varepsilon_n^{-2-\delta} \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}| \\ &\leq 2\varepsilon_n^\delta \sum_{\tau_{i-1}^n < T} |\Delta S_{\tau_i^n}|^2 \xrightarrow{\text{a.s.}} 0 \end{aligned}$$

thanks to Proposition 3.8.

Finally, we analyse $e_{4,T}^n$. From its definition, Proposition 5.3 and (\mathbf{H}_Λ) , we get

$$\begin{aligned}
& |\varepsilon_n^{-(2+\delta)} e_{4,T}^n| \\
& \leq \varepsilon_n^{-(2+\delta)} \sum_{\tau_{i-1}^n < T} \left| \Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n} - \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \right| \left(\Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n} + \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \right) \\
& \leq \varepsilon_n^{-(2+\delta)} c_{(5.3)} \sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} |\Delta S_t + \sigma_{\phi(t)} \Delta B_t| \left| \int_{\phi(t)}^t \Delta \sigma_s dB_s + \Delta A_t \right| \\
& \quad \times \left(\Delta B_{\tau_i^n}^\top X_{\tau_{i-1}^n} \Delta B_{\tau_i^n} + \Delta S_{\tau_i^n}^\top \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \right). \tag{5.16}
\end{aligned}$$

Now we apply twice Lemma 3.3-(ii), first taking $\theta = 0$ and second taking $\theta = \theta_\sigma$: it readily follows that for any given $\rho > 0$, we have a.s. for any $n \in \mathbb{N}$

$$\sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} (|\Delta M_t| + |\sigma_{\phi(t)} \Delta B_t|) \leq C_0 \varepsilon_n^{1-\rho}, \tag{5.17}$$

$$\sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} \left| \int_{\phi(t)}^t \Delta \sigma_s dB_s \right| \leq C_0 \varepsilon_n^{1+\theta_\sigma-\rho}. \tag{5.18}$$

Moreover by Corollary 3.6 we have

$$\sup_{1 \leq i \leq N_T^n} \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} |\Delta A_t| \leq C_0 \varepsilon_n^{2\alpha-\rho}.$$

The last factor in the r.h.s. of (5.16) converges a.s. to a finite random variable (Propositions 3.7 and 5.4). Combining this with the above estimates, the inequality (5.16) becomes

$$|\varepsilon_n^{-(2+\delta)} e_{4,T}^n| \leq C_0 \varepsilon_n^{-2-\delta} \varepsilon_n^{1-\rho} (\varepsilon_n^{1+\theta_\sigma-\rho} + \varepsilon_n^{2\alpha-\rho}).$$

It is now easy to see that, since we have chosen $\delta < \theta_\sigma$ and $\delta < 2\alpha - 1$, we can take ρ small enough so that $\varepsilon_n^{-(2+\delta)} e_{4,T}^n \rightarrow 0$. The proof is finished. \square

5.5 Proof of Proposition 5.2

Consider the equation solved by X_t (see (4.1) and (4.2)), and multiply it by σ_t^\dagger from the right and by $(\sigma_t^\dagger)^\top$ from the left: it gives

$$2 \operatorname{Tr}(X_t) (\sigma_t^\dagger)^\top X_t \sigma_t^\dagger + 4 (\sigma_t^\dagger)^\top X_t^2 \sigma_t^\dagger = (\sigma_t \sigma_t^\dagger)^\top \tilde{C}_t (\sigma_t \sigma_t^\dagger)$$

where $\tilde{C}_t = (D_x v_t)^\top \sigma_t \sigma_t^\top D_x v_t$. Take the trace, use that $(\sigma_t^\dagger)^\top X_t^2 \sigma_t^\dagger \in \mathcal{S}_+^d(\mathbb{R})$ and (\mathbf{H}_Λ) , in order to obtain

$$2 \operatorname{Tr}(X_t) \operatorname{Tr}(\Lambda_t) \leq \operatorname{Tr} \left((\sigma_t \sigma_t^\dagger)^\top \tilde{C}_t (\sigma_t \sigma_t^\dagger) \right).$$

Recall the inequality $\text{Tr}(\mathcal{S}\mathcal{S}') \leq \text{Tr}(\mathcal{S})\text{Tr}(\mathcal{S}')$ for any non-negative definite symmetric matrices \mathcal{S} and \mathcal{S}' . Thus, $\text{Tr}((\sigma_t\sigma_t^\dagger)^\top \tilde{C}_t(\sigma_t\sigma_t^\dagger)) \leq d^2 \text{Tr}(\tilde{C}_t)$ where we have used the easy inequality $\text{Tr}(\sigma_t\sigma_t^\dagger) \leq d$. Note that the above inequalities are of deterministic nature and therefore they hold for any t with probability 1 (the full set is the one allowing to define $X, \Lambda, \sigma, \tilde{C}$). Invoking (\mathbf{H}_M) and (\mathbf{H}_v) to control \tilde{C} , we deduce that there exists a non-negative random variable \tilde{c} , finite a.s., such that

$$\text{Tr}(X_t) \text{Tr}(\Lambda_t) \leq \tilde{c}, \quad \forall t \in [0, T] \quad \text{a.s.} \quad (5.19)$$

Owing to the condition (\mathbf{H}_C) , $X_t \neq 0$ for any $t \in [0, T]$ a.s., and by continuity of X_t , we get that $\inf_{t \in [0, T]} \text{Tr}(X_t) > 0$ a.s. and we conclude to (\mathbf{H}_Λ) thanks to (5.19). \square

A Appendix

A.1 Proof of the Lemma 4.3

In view of (\mathbf{H}_v) there exists $\Omega_{\mathcal{D}}$ with $\mathbb{P}(\Omega_{\mathcal{D}}) = 1$ such that for every $\omega \in \Omega_{\mathcal{D}}$ there is $\delta(\omega) > 0$ such that, for any $\mathcal{A} \in \{D_{x_j}, D_{x_j x_k}^2, D_t : 1 \leq j, k \leq d\}$,

$$\sup_{0 \leq t < T} \sup_{|x - S_t(\omega)| \leq \delta(\omega)} |\mathcal{A}v(t, x)| < +\infty.$$

Since $\sup_{1 \leq i \leq N_T^n} \Delta\tau_i^n \xrightarrow{\text{a.s.}} 0$ and S is continuous on $[0, T]$, there exists a set $\Omega_{\mathcal{C}}$ of full measure such that, for every $\omega \in \Omega_{\mathcal{C}}$, for n large enough we have

$$\sup_{0 \leq s, t \leq T, |t-s| \leq \sup_{1 \leq i \leq N_T^n} \Delta\tau_i^n} |S_t - S_s| \leq \delta(\omega).$$

Hence for $\omega \in \Omega_{\mathcal{C}} \cap \Omega_{\mathcal{D}}$, for n large enough, by a Taylor formula we obtain

$$\sup_{t \in (\tau_{i-1}^n, \tau_i^n]} |v(t, S_t) - v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n})| \leq C_0 \left(\Delta\tau_i^n + \sup_{t \in (\tau_{i-1}^n, \tau_i^n]} |\Delta S_t|^2 \right).$$

Plugging this estimate into $\langle R^n \rangle_T$ we obtain that a.s., for n large enough,

$$\varepsilon_n^{2-4\rho_N} \langle R^n \rangle_T \leq C_0 \varepsilon_n^{2-4\rho_N} \sum_{\tau_{i-1}^n < T} \left((\Delta\tau_i^n)^3 + \Delta\tau_i^n \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta S_t|^4 \right).$$

We deduce that $\varepsilon_n^{2-4\rho_N} \langle R^n \rangle_T \xrightarrow{\text{a.s.}} 0$ since

- for any $\rho > 0$, $\varepsilon_n^{2-4\rho_N} \sum_{\tau_{i-1}^n < T} (\Delta\tau_i^n)^3 \leq \varepsilon_n^{2-4\rho_N} N_T^n \sup_{1 \leq i \leq N_T} (\Delta\tau_i^n)^3 \leq C_0 \varepsilon_n^{8-6\rho_N-\rho}$ by using Lemma 3.2-(ii), thus it converges to 0 since $\rho_N < 4/3$,
- $\varepsilon_n^{2-4\rho_N} \sum_{\tau_{i-1}^n < T} \Delta\tau_i^n \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta S_t|^4 \leq C_0 \varepsilon_n^{6-4\rho_N} T \rightarrow 0 \quad \text{a.s.}$

We are done. \square

A.2 Almost sure convergence using domination in expectation

The next result allows to prove the a.s. convergence of a dominated process U using that of a dominating process V , the domination relation being in expectation. Its use is crucial in our analysis.

Lemma A.1 ([GL14a, Lemma 2.2]). *Let \mathcal{C}_0^+ be the set of non-negative continuous adapted processes, vanishing at $t = 0$. Let $(U^n)_{n \geq 0}$ and $(V^n)_{n \geq 0}$ be two sequences of processes in \mathcal{C}_0^+ . Assume that*

- (i) $t \mapsto V_t^n$ is a non-decreasing function on $[0, T]$, a.s.;
- (ii) the series $\sum_{n \geq 0} V_T^n$ converges a.s.;
- (iii) there is a constant $c \geq 0$ such that, for every $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $t \in [0, T]$, we have

$$\mathbb{E}[U_{t \wedge \theta_k}^n] \leq c \mathbb{E}[V_{t \wedge \theta_k}^n]$$

with the stopping time $\theta_k := \inf\{s \in [0, T] : \bar{V}_s \geq k\}$ ¹ setting $\bar{V}_t = \sum_{n \geq 0} V_t^n$.

Then for any $t \in [0, T]$, the series $\sum_{n \geq 0} U_t^n$ converges a.s.. As a consequence, $U_t^n \xrightarrow{\text{a.s.}} 0$.

References

- [Fuk11a] M. Fukasawa. Asymptotically efficient discrete hedging. *Stochastic analysis with Financial Applications, Progress in Probability*, 65:331–346, 2011.
- [Fuk11b] M. Fukasawa. Discretization error of stochastic integrals. *Annals of Applied Probability*, 21:1436–1465, 2011.
- [GJY03] A. Göing-Jaesche and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.
- [GL14a] E. Gobet and N. Landon. Almost sure optimal hedging strategy. *Ann. Appl. Probab.*, 24(4):1652–1690, 2014.
- [GL14b] E. Gobet and N. Landon. Optimization of joint p -variations of Brownian semimartingales. *Electronic Journal of Probability*, 19(36), 2014.

¹with the usual convention $\inf \emptyset = +\infty$.

- [HHJ15] M. Hairer, M. Hutzenthaler, and A. Jentzen. Loss of regularity for Kolmogorov equations. *Annals of Probability*, 43(2):468–527, 2015.
- [JMY15] A. Jentzen, T. Müller-Gronbach, and L. Yaroslavtseva. On stochastic differential equations with arbitrary slow convergence rates for strong approximation. *arXiv:1506.02828*, 2015.
- [JP12] J. Jacod and P. Protter. *Discretization of Processes*. Stochastic Modelling and Applied Probability 67. Springer Heidelberg, 2012.
- [KP91] T.G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070, 1991.
- [Roo80] H. Rootzen. Limit distributions for the error in approximations of stochastic integrals. *The Annals of Probability*, 8(2):241–251, 1980.