Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2018, Vol. 54, No. 3, 1556–1582
https://doi.org/10.1214/17-AIHP848

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Optimal discretization of stochastic integrals driven by general Brownian semimartingale[1]

Emmanuel Gobet and Uladzislau Stazhynski

CMAP, Ecole Polytechnique and CNRS, Université Paris Saclay, Route de Saclay, 91128 Palaiseau cedex, France.
E-mail: emmanuel.gobet@polytechnique.edu; uladzislau.stazhynski@polytechnique.edu

Received 2 January 2016; revised 20 March 2017; accepted 12 June 2017

Abstract. We study the optimal discretization error of stochastic integrals, driven by a multidimensional continuous Brownian semimartingale. 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. In comparison with previous available results, we allow a quite large class of semimartingales (relaxing in particular the non degeneracy conditions usually requested) and we prove that the asymptotic lower bound is attainable.

Résumé. Nous étudions l'erreur de discrétisation optimale d'intégrale stochastique, dirigée par une semimartingale brownienne continue multidimensionnelle. Dans ce cadre, nous déterminons une borne inférieure trajectorielle pour la variation quadratique de l'erreur renormalisée et nous fournissons une suite de temps d'arrêt de discrétisation, suite qui est asymptotiquement optimale. Cette dernière est définie explicitement à partir des temps d'atteinte d'ellipsoïdes aléatoires par la semimartingale sous-jacente. En comparaison avec les précédents résultats, nous considérons une très grande classe de semimartingales (relâchant en particulier les conditions de non dégénérescence qui étaient habituellement requises) et nous prouvons que la borne inférieure asymptotique est atteignable.

MSC: 60G40; 60F15; 60H05

Keywords: Discretization of stochastic integrals; Hitting times; Random ellipsoids; Almost sure convergence

1. Introduction

Statement of the problem

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

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

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.

[1]This work is part of the Chair *Financial Risks* of the *Risk Foundation* and the *Finance for Energy Market Research Centre*.

The almost sure minimization of Z^n_T is hopeless since after suitable renormalization and under some mild assumptions on the model, Z^n_T weakly converges to a mixture of Gaussian random variables (see [11,14,17]). Alternatively we aim at minimizing a.s. the product

$$N^n_T \langle Z^n \rangle_T. \tag{1.2}$$

The choice of this minimization criterion is inspired by the fact that in many particular cases with deterministic discretization times, we have $\mathbb{E}(\langle Z^n\rangle_T) \sim \mathrm{Const}/N^n_T$ as $N^n_T \to +\infty$. For example, in the one-dimensional Brownian motion case with $v(t,x) = x$ the value of $\mathbb{E}(\langle Z^n\rangle_T)$ for the regular mesh of size n may be calculated exactly and is equal to $\frac{1}{2n}$. For more general S and v satisfying fractional regularity conditions [5], the error $\mathbb{E}(\langle Z^n\rangle_T)$ is still of magnitude Cst/n by appropriately choosing n deterministic times on $[0,T]$.

Background results

The problem of optimizing the discretization times was initially considered in a different framework: simulation of diffusion processes. In [10] the authors study the optimal discretization times for the simulation of a one-dimensional diffusion X via the Euler/Milshtein schemes, where the discretization times adapt to the local properties of every single trajectory. They consider three different schemes and analyze their L_2 errors (in time and ω):

(a) *A simplified Adaptive scheme* $\hat{X}^{**}_h$, for which the sequence of discretization times $(\tau_i)_{1\le i\le \nu}$ is such that each τ_i is a measurable function of the previously simulated values of the Brownian motions $W_{\tau_1},\dots,W_{\tau_{i-1}}$, and Euler and Milshtein schemes with two appropriate time scales are combined to approximate X. This method is of varying cardinality since the number ν of times is random. Observe that $(\tau_i)_i$ are stopping times but they belong to the subclass of strongly predictable times (see [11, Chapter 14]), along which moments of martingale increments are easier to compute.
(b) *An Adaptive scheme* $\hat{X}^{*}_h$ *with discretization times of fixed cardinality*. To control the number of times, a first monitoring of an approximation of X is considered in order to decide where to refine the discretization whilst maintaining a given number of time points. Therefore, the discretization times are somehow anticipative and they are not stopping times.
(c) *An Adaptive scheme* $\hat{X}_h$ *with path-independent step-size Control*, as a variant of $\hat{X}^{*}_h$ where the monitoring is made in mean and not on the specific path X to simulate.

In [10, Theorem 1], the authors prove the asymptotic superiority of $\hat{X}^{**}_h$ over the two other schemes and [10, Theorem 2] states the asymptotic optimality of each scheme within its own class. For the latter optimality result, the criterion used for the optimization is the renormalized L_2-error. Despite the similarities between our current work and theirs, there are significant differences that we shall stress. First, we consider discretization of stochastic integrals and not of diffusion processes, therefore the objectives are quite different. Second, we study the case of general multi-dimensional continuous Brownian semimartingale whereas [10] handles the case of diffusion in $d = 1$ and [16, Chapter III] deals with $d \ge 1$ under commutative noise assumption. Third, we allow optimization over a quite large class of stopping times, see examples of Remark 1 illustrating this fact.

Besides, the study of minimization problems for stochastic integral discretization has been initiated by [4] in dimension $d = 1$, but instead of (1.2) the author considers a criterion in expectation for both terms, i.e. $\mathbb{E}(N^n_T)\mathbb{E}(\langle Z^n\rangle_T)$. However, if $n \to +\infty$ denotes an asymptotic parameter (defined later), observe that

$$\liminf_{n\to+\infty} \mathbb{E}\big(N^n_T\big)\mathbb{E}\big(\langle Z^n\rangle_T\big) \underset{\text{Cauchy–Schwarz ineq.}}{\ge} \liminf_{n\to+\infty} \Big[\mathbb{E}\Big(\sqrt{N^n_T\langle Z^n\rangle_T}\Big)\Big]^2 \tag{1.3}$$

$$\underset{\text{Fatou lemma}}{\ge} \Big[\mathbb{E}\Big(\sqrt{\liminf_{n\to+\infty} N^n_T\langle Z^n\rangle_T}\Big)\Big]^2. \tag{1.4}$$

Since the solution to the problem of a.s. minimizing (1.2) exists (see Theorem 5.1) and is such that N^n_T 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 $\mathbb{E}(N^n_T)\mathbb{E}(\langle Z^n\rangle_T)$, however with substantially more information.

The pathwise minimization of (1.2) has been addressed in a multi-dimensional setting $d \geq 1$, in [6]: the authors assume that S is a local martingale and the lower bound is achieved under stringent conditions of v (essentially its Jacobian matrix $D_x v$ is invertible). These assumptions are restrictive and we aim at relaxing the hypotheses and strengthening the optimality results. This requires to develop new arguments presented in this work.

As an extra motivation for this theoretical study, we refer to the recent work of Hairer et al. [9], 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 [12]. 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.

Our contributions

In the current work, we prove optimality results in a much larger setting than previously afforded in the literature.

- First, we allow S to be a general Brownian semimartingale $S = A + M$, while in [6] S is essentially a local Brownian martingale ($A = 0$, $M = \int_0^{\cdot} \sigma_s \,\mathrm{d}B_s$). Actually, considering the existence of the *finite variation term* A modifies a priori significantly the definition of admissible discretization strategies (see the definition ($\mathbf{A}_S^{\text{osc.}}$) later) and restricts the set of available tools to analyze them. Our first contribution is to establish that admissible strategies for the semimartingale S and for its local martingale part M are the same: see Theorem 3.4. This is a non-trivial result. This allows to transfer a priori estimates available in the martingale case (Lemmas 3.2 and 3.3) to our extended setting, this is instrumental for the subsequent analysis.
- Second, the martingale part of S can be degenerate in our setting, whereas a stronger a.s. ellipticity (on σ) is considered in [6]. This allows to consider partially degenerate models like

$$S_t = \left(\tilde{S}_t, \int_0^t \tilde{S}_s \,\mathrm{d}s\right)$$

or other SDEs with vanishing diffusion coefficient (see Section 5.3 for examples). Also $D_x v(t, S_t)$ may be not invertible in our work. This second set of improvements requires a quite delicate analysis, which constitutes the core of this work. 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 ($\mathbf{H}_C$) or a weaker Assumption ($\mathbf{H}_\Lambda$). These are quite mild conditions.

The ability to treat the non-elliptic case is fundamental for applications as well:

(a) Regarding financial applications, see for example [3,6], minimizing $\langle Z^n \rangle_T$ is related to better hedge market risks. In that context, the treatment of degenerate case appears to be important. Though the covariance matrix of a group of asset returns is usually non-degenerate, it may have some very small eigenvalues [2]. The reason is that typically a large portfolio of financial assets is driven by a smaller number of significant factors, while the other degrees of freedom represent low-variance noise. Thus the inversion of the covariance matrix is often seen as undesirable by practitioners, if no robustness analysis is provided. Our study of the degenerate case justifies in a way the robustness of the optimal discretization algorithm when the diffusion coefficient is degenerate or close to being degenerate.

(b) Some important examples of diffusion models with degenerate diffusion coefficient come as well from random mechanics, see [13] for an overview. Typically, a body is modeled by its position X and its velocity V: it is subjected to random forces, so that due to the second Newton law of motion, its dynamics writes

$$\begin{cases} X_t = X_0 + \int_0^t V_s \,\mathrm{d}s, \\ V_t = V_0 + \int_0^t \phi(X_s, V_s) \,\mathrm{d}s + \int_0^t \psi(X_s, V_s) \,\mathrm{d}W_s. \end{cases} \tag{1.5}$$

In [15], these equations describe the response of structural systems subjected to severe environmental loads (like earthquakes, strong winds, recurrent waves...). The authors study examples like seismic-excited ten-storey building (see [15, Section 5]) where they propose to optimally control the structure by activating tendons, in order to compensate external forces. They derive a continuous-time optimal control, but in practice, only discrete-time controls can be applied. Our study gives a theoretical framework to determine when to apply the controls in order to minimize the deviation from optimally-controlled building.

In [18], the author studies the approximation of stochastic Hamiltonian systems of the form (1.5). The author emphasizes the technical difficulty of the analysis coming from the polynomial growth of the coefficients and the degeneracy of the infinitesimal generators. In our context of optimal discretization problem, our a.s. analysis allows for arbitrary growth conditions on the coefficients.

- Third, we provide a strategy $\mathcal{T}^n$ attaining the lower bound, while in [6], only a μ-optimal strategy (with μ small) is designed. Informally, the natural candidate for optimality is a sequence of hitting times by S of random ellipsoids which characteristics depend on $D_x v$ and S. However, in general and in particular because of the degenerate setting on σ_t and $D_x v(t, S_t)$, this strategy is not admissible (ellipsoids may be flat or infinite). Alternatively, we prove that a suitable perturbation makes the strategy admissible, without altering its asymptotic optimality.

Our main result (Theorem 5.1) states that an optimal strategy is of the form

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n := \inf\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^\mathsf{T} \Lambda_{\tau_{i-1}^n}^{(n)} (S_t - S_{\tau_{i-1}^n}) \geq \widetilde{\varepsilon}_n\} \wedge T, \end{cases}$$

for a sequence $\widetilde{\varepsilon}_n \to 0$, where $\Lambda_t^{(n)}$ is a suitable perturbation of $\Lambda_t := (\sigma_t^\dagger)^\mathsf{T} X_t \sigma_t^\dagger$ (where $\mathcal{M}^\dagger$ is the pseudo-inverse matrix of $\mathcal{M}$), and X_t is the symmetric non-negative definite matrix solution to the equation

$$2\,\mathrm{Tr}(X_t) X_t + 4X_t^2 = \sigma_t^\mathsf{T} \big(D_x v(t, S_t)\big)^\mathsf{T} \sigma_t \sigma_t^\mathsf{T} D_x v(t, S_t) \sigma_t.$$

Additionally the asymptotic lower bound to (1.2) is $(\int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t)^2$.

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, and designing an optimal strategy is made in Section 5. We also present a few examples and a numerical experiment in Section 5.3. Technical results are postponed to the 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 $\mathrm{Tr}(A)$ 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 $d \times d$ matrices with real coefficients.
- For $A \in \mathcal{S}^d(\mathbb{R})$ we denote $\Lambda(A) = (\lambda_1(A), \ldots, \lambda_d(A))$ 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 $\mathrm{Diag}(a_1, \ldots, a_d)$ the square matrix of size d with diagonal entries $a_1, \ldots, a_d$.
- For the partial derivatives of a function $f(t, x)$ we write

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

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

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

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

$$S_t = A_t + M_t, \quad 0 \le t \le T, \tag{2.2}$$

where the processes A and M satisfy the following hypotheses.

($\mathbf{H}_A$) The process A is continuous, adapted and of finite variation, and satisfies

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

($\mathbf{H}_M$) The process M is a continuous local martingale of the form

$$M_t = \int_0^t \sigma_s \, \mathrm{d}B_s, \quad 0 \le t \le T, \tag{$\mathbf{H}_M$}$$

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

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

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 [6], we shall allow its derivatives in uniform norm (in space) to explode as $t \to T$, whilst remaining bounded a.s. in an infinitesimal tube centered at $(t, S_t)_{0\le t<T}$. This is stated precisely in what follows.

($\mathbf{H}_v$) Let $\mathcal{D} \in \{D_{x_j}, D^2_{x_j x_k}, D_t : 1 \le j, k \le d\}$, then

$$\mathbb{P}\Big(\lim_{\delta\to 0} \sup_{0\le t<T} \sup_{|x-S_t|\le\delta} \big|\mathcal{D}v(t, x)\big| < +\infty\Big) = 1. \tag{$\mathbf{H}_v$}$$

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 $n \to +\infty$, a strategy is naturally indexed by $n \in \mathbb{N}$: a *strategy* is a finite sequence of increasing stopping times

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

We now define the appropriate asymptotic framework. Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence of positive deterministic real numbers such that

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

In the following, all convergences are taken as $n \to +\infty$. The above summability enables to derive a.s. convergence results: alternatively, had we assumed only $\varepsilon_n \to 0$, using a subsequence-based argument (see [7, 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:

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

Here the lower argument in the assumption ($\mathbf{A}_{\cdot}^{\text{osc.}}$) 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:

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

In the above, ρ_N is a given parameter satisfying

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

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 and the parameters $(\varepsilon_n)_{n\in\mathbb{N}}$ and ρ_N 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.

Remark 1. The notion of admissible sequence is quite general, in particular, it includes the following two wide families of random grids.

(i) Let $\rho\in(0,1)$ and let $(\varepsilon_n)_{n\geq 0}$ be a deterministic sequence such that $\sum_{n\geq 0}\varepsilon_n^2<+\infty$. Consider $\mathcal{T}=\{\mathcal{T}^n\}_{n\geq 0}$ where each $\mathcal{T}^n=(\tau_i^n)_{0\leq i\leq N_T^n}$ is a sequence of stopping times (with N_T^n possibly random) and such that

$$C^{-1}\varepsilon_n^{\frac{2}{(1-\rho)}}\leq\min_{1\leq i\leq N_T^n}\Delta\tau_i^n\leq\max_{1\leq i\leq N_T^n}\Delta\tau_i^n\leq C\varepsilon_n^{\frac{2}{(1-\rho)}},\quad n\geq 0,\ \text{a.s.},$$

for an a.s. finite positive random variable $C>0$. This example contains in particular the sequences of deterministic grids for which the time steps are controlled from below and from above (like those of [10] used for building $\hat{X}_h^{**}$ mentioned in introduction), and for which the step size tends to zero fast enough.

Let us check ($\mathbf{A}_S^{\text{osc.}}$)and ($\mathbf{A}_N$). First, note that S is a.s. Hölder continuous on $[0,T]$ with exponent $\frac{1-\rho}{2}$: this is a consequence of ($\mathbf{H}_A$) for the finite-variation component A and of [1, Theorem 5.1] for the martingale component M under the assumption ($\mathbf{H}_M$). Therefore, a.s. for each $n\geq 0$

$$\sup_{1\leq i\leq N_T^n}\sup_{t\in[\tau_{i-1}^n,\tau_i^n]}|S_t-S_{\tau_{i-1}^n}|\leq C_S\Big[\max_{1\leq i\leq N_T^n}\Delta\tau_i^n\Big]^{\frac{1-\rho}{2}}\leq C_S C^{\frac{1-\rho}{2}}\varepsilon_n.$$

Furthermore,

$$N_T^n\leq\frac{T}{\min_{1\leq i\leq N_T^n}\Delta\tau_i^n}\leq TC\varepsilon_n^{-\frac{2}{(1-\rho)}}$$

so that ($\mathbf{A}_N$) is verified with $2\rho_N=2(1-\rho)$ provided that we take ρ small enough to satisfy the upper bound (2.3). Thus the sequence of strategies $\mathcal{T}$ is admissible for $(\varepsilon_n)_{n\geq 0}$ and ρ_N given above.

(ii) Consider a sequence of adapted random processes $\{D_t^n : 0 \le t \le T\}$ where each D_t^n is an open set such that

$$B(0, C_1\varepsilon_n) \subset D_t^n \subset B(0, C_2\varepsilon_n)$$

for some a.s. finite positive random variables C_1, C_2, here $B(0,r)$ denotes the ball centered at 0 with radius r. Here again the deterministic sequence $(\varepsilon_n)_{n\ge 0}$ is such that $\sum_{n\ge 0}\varepsilon_n^2 < +\infty$. Define the sequence of strategies $\mathcal{T} = \{\mathcal{T}^n\}_{n\ge 0}$ with $\mathcal{T}^n = (\tau_i^n)_{0\le i\le N_T^n}$ as follows: $\tau_0^n = 0$ and for $i \ge 1$

$$\tau_i^n = \inf\bigl\{t > \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n}) \notin D^n_{\tau_{i-1}^n}\bigr\} \wedge T.$$

In other words, we consider exit times of random sets of size ε_n. The assumption $(\mathbf{A}_S^{\text{osc.}})$ follows from the definition of $\mathcal{T}^n$:

$$\sup_{1\le i\le N_T^n}\sup_{t\in[\tau_{i-1}^n,\tau_i^n]} |S_t - S_{\tau_{i-1}^n}| \le C_2\varepsilon_n.$$

Further to check $(\mathbf{A}_N)$, we write (using Proposition 3.8)

$$C_1^2\varepsilon_n^2 N_T^n \le C_1^2\varepsilon_n^2 + \sum_{\tau_{i-1}^n<T} |\Delta S_{\tau_i^n}|^2 \underset{n\to+\infty}{\to} \operatorname{Tr}\bigl(\langle S\rangle_T\bigr) < +\infty \quad \text{a.s.}$$

This proves the admissibility of $\mathcal{T}$. A particular case is the ellipsoid exit times, see [6, Proposition 2.4].

3. General results for admissible strategies

This section gathers preliminary results, needed to establish the subsequent main results. In Section 3.1, we recall without proof some estimates about the mesh size $\sup_{1\le i\le N_T^n}\Delta\tau_i^n$ 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 Section 3.2 where we establish an important result: in our setting, admissible sequences of strategies for S and M are the same. Last in Section 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 ([6, Corollary 2.1]). *Let $p > 0$, and let $\{(K_t^n)_{0\le t\le T} : n \ge 0\}$ be a sequence of continuous scalar local martingales vanishing at zero. Then*

$$\sum_{n\ge 0}\langle K^n\rangle_T^{p/2} < +\infty \quad \text{a.s.} \quad \Longleftrightarrow \quad \sum_{n\ge 0}\sup_{0\le t\le T}\bigl|K_t^n\bigr|^p < +\infty \quad \text{a.s.}$$

The useful application is the sense $\Rightarrow$: by controlling the summability of quadratic variations, we obtain the non trivial a.s. convergence of $\sup_{0\le t\le T}|K_t^n|$ 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 $\mathcal{T} = \{\mathcal{T}^n : n \ge 0\}$ 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 ([6, Corollary 2.2]). *Assume $(\mathbf{H}_M)$ and let $\mathcal{T} = \{\mathcal{T}^n : n \ge 0\}$ be a sequence of strategies. Let $\rho > 0$, then the following hold*:

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

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

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

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

Lemma 3.3 ([6, 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 $\langle K^n\rangle_t=\int_0^t\kappa_r^n\,\mathrm{d}r$ 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*

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

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

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

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

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

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

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 it is admissible for M with the same parameter ρ_N: in other words, if $\mathcal{T}$ satisfies* $(\mathbf{A}_N)$,

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

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

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

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 $(\mathbf{A}_M^{\mathrm{osc.}})$ 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

$$|\Delta M|_\infty\leq|\Delta S|_\infty+C_0\left|\Delta\tau^n\right|_\infty^\alpha\leq C_0\big(\varepsilon_n+\left|\Delta\tau^n\right|_\infty^\alpha\big).\tag{3.1}$$

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

$$0 \le t - s \le C_E^{-1} \int_s^t \mathrm{Tr}\big(\sigma_r \sigma_r^{\mathsf{T}}\big)\,\mathrm{d}r = C_E^{-1} \sum_{j=1}^d \big(\langle S^j\rangle_t - \langle S^j\rangle_s\big) \tag{3.2}$$

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

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

$$\Delta t \le C_E^{-1}\left(C_0 \varepsilon_n^2 + 2\sum_{j=1}^d \left|\int_{\phi(t)}^t \Delta S_r^j\,\mathrm{d}A_r^j\right| + 2\sum_{j=1}^d \left|\int_{\phi(t)}^t \Delta S_r^j\,\mathrm{d}M_r^j\right|\right). \tag{3.4}$$

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

$$\sum_{j=1}^d \left|\int_{\phi(t)}^t \Delta S_r^j\,\mathrm{d}A_r^j\right| \le C_0 \varepsilon_n. \tag{3.5}$$

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

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

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

$$\sup_{0\le t\le T} \left|\int_0^t \Delta S_r^j\,\mathrm{d}M_r^j\right| \le C_0 \varepsilon_n^{1-\frac{2}{p}} = C_0 \varepsilon_n^{1-\delta}, \tag{3.6}$$

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

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

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

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

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

$\triangleright$ *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 $\beta > 0$*

$$\sup_{n\ge 0}\Big(\varepsilon_n^{-\beta} \sup_{1\le i\le N_T^n} \sup_{t\in(\tau_{i-1}^n, \tau_i^n]} |\Delta M_t|^2\Big) < +\infty \quad \textit{a.s.} \tag{3.9}$$

Then for any $\rho > 0$

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

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 $(\mathbf{A}_N)$ 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

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

Then it follows that

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

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

▷ *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

$$d_{m+1}-d_m\geq\alpha\rho_0, \tag{3.10}$$

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

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

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

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

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

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

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

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

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

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

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

Remark 2.

- 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 $(\mathbf{A}_\cdot^{\text{osc.}})$ 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 [6]): in particular, for any admissible sequences of strategies (for M or S), we have $\sup_{1\le i\le N_T^n} |\Delta\tau_i^n| \le C_0\varepsilon_n^{2-\gamma}$ 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 $\rho > 0$,*

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

3.3. *Convergence results for quadratic variation*

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

Proposition 3.7 ([6, Proposition 2.3]). *Assume $(\mathbf{H}_M)$ and let $\mathcal{T}$ be a sequence of strategies satisfying $(\mathbf{A}_M^{\text{osc.}})$. Let $(H_t)_{0\le t<T}$ be a continuous adapted $d\times d$-matrix process such that $\sup_{t\in[0,T)}|H_t| < +\infty$ a.s., and let $(K_t)_{0\le t\le T}$ be a $\mathbb{R}^d$-valued continuous local martingale such that $\langle K\rangle_t = \int_0^t \kappa_r \,\mathrm{d}r$ with $\sup_{t\in[0,T]}|\kappa_t| < +\infty$ a.s. Then*

$$\sum_{\tau_{i-1}^n < T} \Delta K_{\tau_i^n}^{\mathsf{T}} H_{\tau_{i-1}^n} \Delta K_{\tau_i^n} \overset{a.s.}{\to} \int_0^T \operatorname{Tr}\big(H_t \,\mathrm{d}\langle K\rangle_t\big).$$

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 $(H_t)_{0\le t<T}$ be as in Proposition* 3.7. *Then*

$$\sum_{\tau_{i-1}^n<T} \Delta S_{\tau_i^n}^{\mathsf{T}} H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \overset{a.s.}{\to} \int_0^T \operatorname{Tr}\bigl(H_t \,\mathrm{d}\langle M\rangle_t\bigr).$$

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

$$\int_0^T \Delta S_t^{\mathsf{T}}\bigl(H_{\varphi(t)} + H_{\varphi(t)}^{\mathsf{T}}\bigr)\,\mathrm{d}S_t + \int_0^T \operatorname{Tr}\bigl([H_{\varphi(t)} - H_t]\,\mathrm{d}\langle M\rangle_t\bigr). \tag{3.12}$$

Due to ($\mathbf{H}_M$), the second term is bounded by $C_0 \int_0^T |H_{\varphi(t)} - H_t|\mathrm{d}t$: 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

$$\left|\int_0^T \Delta S_t^{\mathsf{T}}\bigl(H_{\varphi(t)} + H_{\varphi(t)}^{\mathsf{T}}\bigr)\,\mathrm{d}A_t\right| \le C_0 \sup_{1\le i\le N_T^n} \sup_{\tau_{i-1}^n\le t\le \tau_i^n} |\Delta S_t| \sup_{t\in[0,T)} |H_t| \overset{\text{a.s.}}{\to} 0 \tag{3.13}$$

in view of ($\mathbf{A}_S^{\text{osc.}}$). On the other hand, $\int_0^T \Delta S_t^{\mathsf{T}}(H_{\varphi(t)} + H_{\varphi(t)}^{\mathsf{T}})\,\mathrm{d}M_t \overset{\text{a.s.}}{\to} 0$ by proceeding very similarly to the proof of (3.6). □

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 $(H_t)_{0\le t<T}$ (or a perturbation of it). This extends [6, 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 $(H_t)_{0\le t<T}$ be a continuous adapted symmetric non-negative definite $d\times d$ matrix process, such that a.s.*

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

The strategy $\mathcal{T}^n$ given by

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

defines a admissible sequence of strategies.

The proof is given later. The condition $\sup_{0\le t<T}\lambda_{\max}(H_t) < +\infty$ ensures that none of the corresponding ellipsoids $\mathcal{E}_t := \{x^{\mathsf{T}} H_t x \le c\}$ 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

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

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 $(H_t)_{0\le t<T}$ 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*

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

(*notice that H is not necessarily continuous*).

Define a sequence of processes $H^{(n)}$ by

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

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

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

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

(ii)

$$\sum_{\tau_{i-1}^n<T} \Delta S_{\tau_i^n}^{\mathsf{T}} H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \overset{a.s.}{\to} \int_0^T \mathrm{Tr}\big(H_t\, \mathrm{d}\langle M\rangle_t\big),$$

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

$$\lambda_{\max}\big(H_t^{(n)}\big) \le C_H + \sup_{n\ge0} \varepsilon_n^\delta < +\infty, \quad \forall t \in [0, T) \text{ a.s.}$$

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

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

we have

$$\begin{aligned} 0 < \varepsilon_n^{2+\delta} &= (S_{\tau_i^n} - S_{\tau_{i-1}^n})^{\mathsf{T}} H_{\tau_{i-1}^n}^{(n)} (S_{\tau_i^n} - S_{\tau_{i-1}^n}) \\ &\le \Big(C_H + \sup_{n\ge0} \varepsilon_n^\delta\Big)|S_{\tau_i^n} - S_{\tau_{i-1}^n}|^2 \xrightarrow{i\to+\infty} 0, \end{aligned}$$

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

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

Thus

$$\begin{aligned}&\varepsilon_n^{-2}\sup_{1\le i\le N_T^n}\sup_{\tau_{i-1}^n\le t\le\tau_i^n}|\Delta S_t|^2\\&\quad\le\Bigl(\inf_{t\in[0,T)}\lambda_{\min}\bigl(H_t^{(n)}\bigr)\Bigr)^{-1}\varepsilon_n^{-2}\sup_{1\le i\le N_T^n}\sup_{\tau_{i-1}^n\le t\le\tau_i^n}\bigl(\Delta S_t^{\mathsf T}H^{(n)}_{\tau_{i-1}^n}\Delta S_t\bigr)\le 2\varepsilon_n^{-\delta}\varepsilon_n^{-2}\varepsilon_n^{2+\delta}=2,\end{aligned}$$

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

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 $N_T^n=1+\sum_{1\le i\le N_T^n-1}1$ and using $2+\delta\le 2\rho_N$, we observe that (for n large enough so that $\varepsilon_n\le 1$)

$$\varepsilon_n^{2\rho_N}N_T^n\le\varepsilon_n^{2+\delta}N_T^n\le\varepsilon_n^{2+\delta}+\sum_{\tau_i^n<T}\Delta S_{\tau_i^n}^{\mathsf T}H_{\tau_{i-1}^n}\Delta S_{\tau_i^n}+\sum_{\tau_i^n<T}\Delta S_{\tau_i^n}^{\mathsf T}\bigl(H^{(n)}_{\tau_{i-1}^n}-H_{\tau_{i-1}^n}\bigr)\Delta S_{\tau_i^n}. \tag{3.16}$$

Now by (ii) we have

$$\sum_{\tau_i^n<T}\Delta S_{\tau_i^n}^{\mathsf T}H_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\overset{\text{a.s.}}{\to}\int_0^T\mathrm{Tr}\bigl(H_t\,\mathrm{d}\langle M\rangle_t\bigr)\overset{\text{a.s.}}{<}+\infty$$

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

$$\Bigl|\sum_{\tau_i^n<T}\Delta S_{\tau_i^n}^{\mathsf T}\bigl(H^{(n)}_{\tau_{i-1}^n}-H_{\tau_{i-1}^n}\bigr)\Delta S_{\tau_i^n}\Bigr|\le\varepsilon_n^{\delta}\sum_{\tau_i^n<T}|\Delta S_{\tau_i^n}|^2\overset{\text{a.s.}}{\to}0, \tag{3.17}$$

using $\delta>0$ and Proposition 3.8 (valid since ($\mathbf{A}_S^{\mathrm{osc.}}$) 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 ($\mathbf{A}_N$). □

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 ($\mathbf{A}_S^{\mathrm{osc.}}$) stems from

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

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

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

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

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 [6, 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:

$$Z_s^n=\int_0^s\bigl(v(t,S_t)-v\bigl(\phi(t),S_{\phi(t)}\bigr)\bigr)\cdot\mathrm{d}M_t+\int_0^s\bigl(v(t,S_t)-v\bigl(\phi(t),S_{\phi(t)}\bigr)\bigr)\cdot\mathrm{d}A_t.$$

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*

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

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 [6, 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 $c=\sigma^\top(D_xv)^\top\sigma$, i.e.*

$$X_t:=x\big(\sigma_t^\top(D_xv_t)^\top\sigma_t\big),\quad \text{for } 0\le t<T. \tag{4.2}$$

Then we have

$$\liminf_{n\to+\infty}N_T^n\langle Z^n\rangle_T\ge\left(\int_0^T\operatorname{Tr}(X_t)\,\mathrm{d}t\right)^2\quad a.s.$$

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

$$\int_0^s\big(v(t,S_t)-v\big(\phi(t),S_{\phi(t)}\big)\big)\cdot\mathrm{d}M_t=:\int_0^s(D_xv_{\phi(t)}\Delta S_t)\,\mathrm{d}M_t+R_s^n. \tag{4.3}$$

Therefore the quadratic variation of Z^n is given by

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

where

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

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 $\mathrm{d}\langle M\rangle_t$ by $\sigma_{\phi(t)}\sigma_{\phi(t)}^\top\,\mathrm{d}t$, which leads to

$$\langle Z^n\rangle_T=\int_0^T\Delta B_t^\top\big(\sigma_{\phi(t)}^\top(D_xv_{\phi(t)})^\top\sigma_{\phi(t)}\sigma_{\phi(t)}^\top D_xv_{\phi(t)}\sigma_{\phi(t)}\big)\Delta B_t\,\mathrm{d}t+e_{0,T}^n+e_{1,T}^n+e_{2,T}^n,$$

where

$$e_{2,T}^n := \int_0^T \Delta M_t^{\mathsf{T}} (D_x v_{\phi(t)})^{\mathsf{T}} \Delta\big(\sigma_t\sigma_t^{\mathsf{T}}\big) D_x v_{\phi(t)} \Delta M_t \,\mathrm{d}t$$
$$+ \int_0^T (\Delta M_t + \sigma_{\phi(t)}\Delta B_t)^{\mathsf{T}} (D_x v_{\phi(t)})^{\mathsf{T}} \sigma_{\phi(t)}\sigma_{\phi(t)}^{\mathsf{T}} D_x v_{\phi(t)} (\Delta M_t - \sigma_{\phi(t)}\Delta B_t)\,\mathrm{d}t.$$

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

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

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

$$\sum_{\tau_{i-1}^n<T} \big(\Delta B_{\tau_i^n}^{\mathsf{T}} X_{\tau_{i-1}^n} \Delta B_{\tau_i^n}\big)^2 = \int_0^T \Delta B_t^{\mathsf{T}}\big(2\operatorname{Tr}(X_{\phi(t)})X_{\phi(t)} + 4X_{\phi(t)}^2\big)\Delta B_t\,\mathrm{d}t$$
$$+ 4\int_0^T \Delta B_t^{\mathsf{T}} X_{\phi(t)} \Delta B_t \Delta B_t^{\mathsf{T}} X_{\phi(t)}\,\mathrm{d}B_t.$$

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 $\sigma^{\mathsf{T}}(D_x v)^{\mathsf{T}}\sigma$ is. Then a simplified representation of $\langle Z^n\rangle_T$ readily follows:

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

where

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

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

$$N_T^n \sum_{\tau_{i-1}^n<T} \big(\Delta B_{\tau_i^n}^{\mathsf{T}} X_{\tau_{i-1}^n} \Delta B_{\tau_i^n}\big)^2 \ge \Big(\sum_{\tau_{i-1}^n<T} \Delta B_{\tau_i^n}^{\mathsf{T}} X_{\tau_{i-1}^n} \Delta B_{\tau_i^n}\Big)^2.$$

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

$$\sum_{\tau_i^n<T} \Delta B_{\tau_i^n}^{\mathsf{T}} X_{\tau_{i-1}^n} \Delta B_{\tau_i^n} \xrightarrow{\text{a.s.}} \int_0^T \operatorname{Tr}(X_t)\,\mathrm{d}t.$$

To summarize we have obtained that

$$\liminf_{n\to+\infty}\big(N_T^n\langle Z^n\rangle_T - N_T^n\big(e_{0,T}^n + e_{1,T}^n + e_{2,T}^n + e_{3,T}^n\big)\big) \ge \Big(\int_0^T \operatorname{Tr}(X_t)\,\mathrm{d}t\Big)^2 \quad \text{a.s.}$$

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

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

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

$$\left|e^n_{0,T}\right| \le C_0 \int_0^T |\Delta A_t|\big(|\Delta S_t| + |\Delta M_t|\big)\,\mathrm{d}t \le C_0 \varepsilon_n^{1+2\alpha-\rho},$$

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

Contribution $e^n_{1,T}$. To handle it, we need the following lemma; its proof follows that of [6, Lemma 3.2], with minor adaptations (see Appendix A.1).

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

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

$$\varepsilon_n^{-2\rho_N}\left|e^n_{1,T}\right| \le \varepsilon_n^{-2\rho_N}\big(\langle R^n\rangle_T\big) + 2C_0\varepsilon_n\big(\langle R^n\rangle_T\big)^{1/2} = o\big(\varepsilon_n^{2\rho_N-2}\big) + o(1) \overset{\text{a.s.}}{\to} 0.$$

Contributions $e^n_{2,T}$ *and* $e^n_{3,T}$. The proof is similar to that of [6, Theorem 3.1], we skip the details. □

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

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

Notice that an existence result is proved in [6, Theorem 3.3], only under the conditions that σ is invertible, that $v(t,x) = \nabla_x u(t,x)$ with

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

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 $\mathcal{M}$

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

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

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

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

$$\begin{cases} \mathcal{M}\mathcal{M}^\dagger\mathcal{M} = \mathcal{M}, \qquad \mathcal{M}^\dagger\mathcal{M}\mathcal{M}^\dagger = \mathcal{M}^\dagger, \\ \text{the matrices } \mathcal{M}\mathcal{M}^\dagger \text{ and } \mathcal{M}^\dagger\mathcal{M} \text{ are symmetric.}\end{cases} \tag{5.1}$$

5.2. *Main result*

We wish to design optimal stopping times in terms of the process S to allow better tractability. Inspired by [6], a good candidate is then the sequence $\{\mathcal{T}^n : n \geq 0\}$ where $\mathcal{T}^n$ is defined as:

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

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 [6, 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 ($\mathbf{A}^{\mathrm{osc.}}$)). To overcome these problems we use $\sigma_t^\dagger$ 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.

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

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

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

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

Note that $\sigma^\dagger$ may be discontinuous, so Λ may be too. Recall (see (3.14)) that $\chi(\cdot)$ stands for a continuous 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* ($\mathbf{H}_A$), ($\mathbf{H}_M$), ($\mathbf{H}_v$), ($\mathbf{H}_\Lambda$) *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*

$$\Lambda_t^{(n)} = \Lambda_t + \varepsilon_n^\delta \chi_{\varepsilon_n^\delta}\left(\lambda_{\min}(\Lambda_t)\right)\mathrm{Id}_d,$$

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

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

Then the sequence of strategies $\mathcal{T} = \{\mathcal{T}_{\varepsilon_n^\delta}^n : n \geq 0\}$ *is admissible for the parameter* ρ_N (*in the sense of Definition* 1 *and Theorem* 3.4) *and is asymptotically optimal*, *i.e.*

$$\lim_{n \to +\infty} N_T^n \langle Z^n \rangle_T = \left(\int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t\right)^2 \quad a.s.$$

To conclude this subsection, we provide a condition simpler than ($\mathbf{H}_\Lambda$), 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*

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

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

5.3. *Examples*

5.3.1. *About the assumptions* ($\mathbf{H}_\Lambda$) *and* ($\mathbf{H}_C$)

Recall that under our assumptions, X is a.s. uniformly bounded on $[0, T)$. Thus in order to satisfy ($\mathbf{H}_\Lambda$), it is enough to have $\sigma^\dagger$ 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 $\mathrm{rank}(\sigma_t) < d$), then $\sigma_t^\dagger$ is also constant in time (and thus bounded).
(c) The previous principle can be generalized to the time-dependent case $\sigma_t = \left(\begin{smallmatrix}\Sigma_t & 0\\ 0 & 0\end{smallmatrix}\right)$ where Σ_t is a square matrix, a.s. invertible at any time: indeed $\sigma_t^\dagger = \left(\begin{smallmatrix}\Sigma_t^{-1} & 0\\ 0 & 0\end{smallmatrix}\right)$ is bounded on $[0, T]$.

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

$$S_t^i = S_0^i + \int_0^t \left(\delta_i - \lambda_i S_s^i\right) \mathrm{d}s + 2\int_0^t \sqrt{S_s^i}\, \mathrm{d}B_s^i,$$

where $S_0^i > 0$, $\delta_i \geq 0$, $\lambda_i \in \mathbb{R}$ (see [8]). The matrix σ_t is diagonal and its ith 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 $\sigma_t^\dagger$ is diagonal with ith 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, $\sigma^\dagger$ is unbounded on $[0, T]$ and it is not clear anymore to check directly ($\mathbf{H}_\Lambda$). 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 $\mathrm{Tr}(C_t) = \mathrm{Tr}(D_x v_t \sigma_t \sigma_t^\mathsf{T}) > 0$ since $\sigma_t\sigma_t^\mathsf{T} \neq 0$ and $D_x v_t$ is invertible.

5.3.2. *A numerical example*

We consider a two-dimensional example, defined by

$$S_t = \begin{pmatrix} B_t^1 + 0.3B_t^2 \\ \int_0^t B_s^1\, \mathrm{d}s \end{pmatrix}.$$

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, ($\mathbf{H}_\Lambda$) 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

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

From Theorem 4.2 we must have $\liminf_{n\to+\infty} \alpha_n \geq 1$ 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.

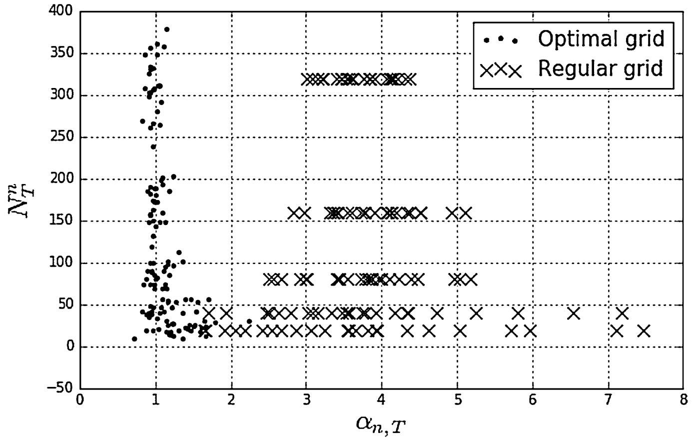

Fig. 1. The values $\alpha_{n,\text{opt}}$ and $\alpha_{n,\text{det}}$ with respect to N_T^n.

To simulate the process S on [0, 1] we use a thin uniform time mesh with $\bar{n} = 10{,}000$ points. The same mesh is later used to calculate the true value of the stochastic integral and the optimal lower bound equal to $(\int_0^T \operatorname{Tr}(X_t)\,\mathrm{d}t)^2$. The hitting times are calculated as well on this mesh. Using this thin grid induces a discrete-time sampling error but by taking $\bar{n}$ quite large as we do, we guess that this error can be neglected in our subsequent results.

We simulate 25 trajectories of the process S on [0, 1]. Further we test the optimal discretization strategy and the regular deterministic discretization on these trajectories, for different discretization parameters ε_n.

(a) To test the performance of the optimal discretization we take 5 different values of ε_n, namely 0.2, 0.14, 0.1, 0.07, 0.05, and apply the strategy given in Theorem 5.1.
(b) Further we test the performance of the deterministic discretization strategy with N_T^n equidistant times, for $N_T^n = 20, 40, 80, 160, 320$ (the values of N_T^n are empirically chosen as approximately equal to the average number of discretization times in the optimal algorithm for the values of ε_n given above).

We denote $(\alpha_{n,\text{opt}}, \beta_{n,\text{opt}})$ and $(\alpha_{n,\text{det}}, \beta_{n,\text{det}})$ the pairs (α_n, β_n) respectively for the optimal and the regular deterministic strategy.

Regarding further details of implementation, we refer to [6, Proof of Lemma 3.1] for the detailed construction of the solution to the matrix equation (4.1). For the computation of the pseudo-inverse matrix in ($\mathbf{H}_\Lambda$), this is straightforward since σ_t is constant. For the perturbation procedure appearing in (5.4), we take $\delta = 0.6 \le 2(\rho_N - 1) < \frac{2}{3}$ and the function $\chi(x) = \sin(\pi(x \vee 1/2) \wedge 1)$.

Figure 1 shows the values of $\alpha_{n,\text{opt}}$ and $\alpha_{n,\text{det}}$ with respect to the number of the discretization times N_T^n for the optimal and the regular discretization in all the tests belonging to 5 different groups. We observe that the values $\alpha_{n,\text{opt}}$ become less and less dispersed and converges to 1 as N_T^n increases ($\varepsilon_n \to 0$), which confirms the theoretical results. In particular, from $N_T^n = 80$ the quality of the algorithm is already good and it largely outperforms the regular discretization.

Figure 2 illustrates the pairs (α_n, β_n) for the same 25 simulations, where $\varepsilon_n = 0.05$ was used for the optimal discretization and $N_T^n = 320$ was used for the regular deterministic strategy (i.e. the last group of the tests). As expected from Theorems 4.2 and 5.1, we observe the inequality $\alpha_{n,\text{opt}} < \alpha_{n,\text{det}}$ and the limit $\alpha_{n,\text{opt}} \approx 1$. Moreover, the inequality $|\beta_{n,\text{opt}}| < |\beta_{n,\text{det}}|$ holds as well for 21 of the 25 simulations. The empirical variances of the values of $\beta_{n,\text{opt}}$ and $\beta_{n,\text{det}}$ are equal to 1.07 and 3.52 respectively, which is nearly the same ratio as for the corresponding values of α_n: this observation is coherent with the possible property of Central Limit Theorem for β_n, where the limiting distribution would be a mixture of Gaussian distributions with variance roughly equal to α_n. This latter property is just a conjecture which is delicate to prove and left for further research. Anyway, this observation confirms that the almost sure minimization of the limit of α_n helps to reduce the variance of β_n as expected.

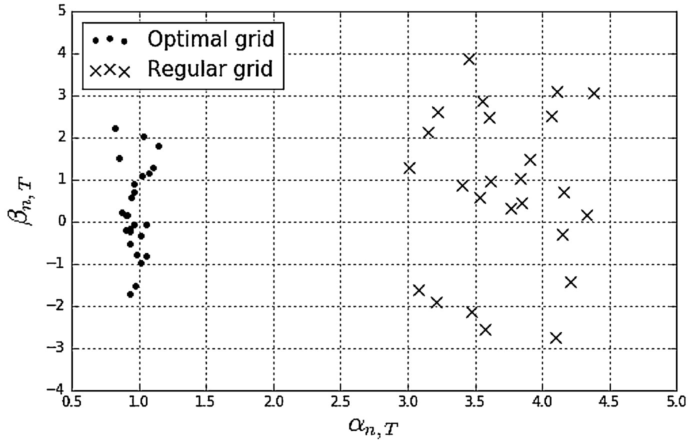

Fig. 2. The pairs $(\alpha_{n,\mathrm{det}}, \beta_{n,\mathrm{det}})$ and $(\alpha_{n,\mathrm{opt}}, \beta_{n,\mathrm{opt}})$ are represented by crosses and points respectively.

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*

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

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 ($\mathbf{H}_\Lambda$).

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

$$\sigma_t \sigma_t^\dagger \sigma_t = \sigma_t \tag{5.6}$$

and the matrix $\sigma_t^\dagger \sigma_t$ is symmetric. We choose an orthonormal basis $(e_i)_{1 \le i \le d}$ 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 $(e_i)_{1 \le i \le d}$), then from (5.6) we get

$$\left(\alpha_1 \sigma_t^1, \dots, \alpha_d \sigma_t^d\right) = \left(\sigma_t^1, \dots, \sigma_t^d\right). \tag{5.7}$$

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

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

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

$$X_t = \big(\sigma_t^{\dagger}\sigma_t\big)^{\mathsf{T}} X_t \big(\sigma_t^{\dagger}\sigma_t\big) = \big(\sigma_t^{\dagger}\sigma_t\big) X_t \big(\sigma_t^{\dagger}\sigma_t\big). \tag{5.9}$$

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:

$$e_i^{\mathsf{T}} X_t e_i = 0 \quad \text{for } i = k+1, \dots, d, \tag{5.10}$$

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

$$2\,\mathrm{Tr}(X_t) X_t + 4X_t^2 = \sigma_t^{\mathsf{T}} \tilde{C}_t \sigma_t, \tag{5.11}$$

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

$$2\,\mathrm{Tr}(X_t) e_i^{\mathsf{T}} X_t e_i + 4 e_i^{\mathsf{T}} X_t^2 e_i = 0.$$

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 $\mathrm{Tr}(X_t) = 0$ (implying $X_t = 0$ and (5.10)), or $\mathrm{Tr}(X_t) > 0$ and $e_i^{\mathsf{T}} X_t e_i = 0$. In any case, (5.10) holds and we are done. □

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

The stopping times (5.4) define a sequence of strategies 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*

$$\sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \overset{a.s.}{\to} \int_0^T \mathrm{Tr}\big(\Lambda_t \,\mathrm{d}\langle M\rangle_t\big) = \int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t.$$

Then, in view of Theorem 3.10(ii), we conclude that the sequence of strategies $\mathcal{T} = \{\mathcal{T}_{\varepsilon_n^\delta}^n : n \ge 0\}$ 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 (3.8) 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

$$\sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} = 2\int_0^T \Delta S_t^{\mathsf{T}} \Lambda_{\phi(t)}\,\mathrm{d}S_t + \int_0^T \mathrm{Tr}\big(\Lambda_{\phi(t)}\,\mathrm{d}\langle M\rangle_t\big). \tag{5.12}$$

Then

$$\begin{aligned}\int_0^T \mathrm{Tr}\big(\Lambda_{\phi(t)}\,\mathrm{d}\langle M\rangle_t\big) &= \int_0^T \mathrm{Tr}\big(\sigma_t^{\mathsf{T}} \Lambda_{\phi(t)} \sigma_t\big)\,\mathrm{d}t \\ &= \int_0^T \mathrm{Tr}\big(\sigma_{\phi(t)}^{\mathsf{T}} \Lambda_{\phi(t)} \sigma_{\phi(t)}\big)\,\mathrm{d}t + \int_0^T \mathrm{Tr}\big((\sigma_t - \sigma_{\phi(t)})^{\mathsf{T}} \Lambda_{\phi(t)} (\sigma_t + \sigma_{\phi(t)})\big)\,\mathrm{d}t.\end{aligned}$$

Observe that the first term on the r.h.s. above is equal to $\int_0^T \mathrm{Tr}(X_{\phi(t)})\,\mathrm{d}t$ 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 $\int_0^T \mathrm{Tr}(X_{\phi(t)})\,\mathrm{d}t \overset{\text{a.s.}}{\to} \int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t$.

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

$$\int_0^T \mathrm{Tr}\big(\Lambda_{\phi(t)}\,\mathrm{d}\langle M\rangle_t\big) \overset{\text{a.s.}}{\to} \int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t.$$

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

$$\int_0^T \Delta S_t^{\mathsf{T}} \Lambda_{\phi(t)}\,\mathrm{d}S_t \overset{\text{a.s.}}{\to} 0.$$

The a.s.-convergence to 0 of the contribution $\int_0^T \Delta S_t^{\mathsf{T}} \Lambda_{\phi(t)}\,\mathrm{d}A_t$ is proved as for (3.13), using ($\mathbf{A}_S^{\text{osc.}}$) and ($\mathbf{H}_\Lambda$). The second contribution $K_T^n := \int_0^T \Delta S_t^{\mathsf{T}} \Lambda_{\phi(t)}\,\mathrm{d}M_t$ 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 ($\mathbf{H}_\Lambda$)). Consequently, an application of Lemma 3.1 with $p=2$, ensures that $K_T^n \overset{\text{a.s.}}{\to} 0$. We are done. □

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

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

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

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

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 $\int_0^T \mathrm{Tr}(\Lambda_t\,\mathrm{d}\langle M\rangle_t)$ (Proposition 5.4). To summarize, we have justified

$$\lim_{n\to+\infty} \varepsilon_n^{2+\delta} N_T^n = \int_0^T \mathrm{Tr}\big(\Lambda_t\,\mathrm{d}\langle M\rangle_t\big) = \int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t \quad \text{a.s.} \tag{5.14}$$

Thus it remains to show that

$$\lim_{n\to+\infty} \varepsilon_n^{-(2+\delta)} \langle Z^n\rangle_T = \int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t \quad \text{a.s.} \tag{5.15}$$

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

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

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:

$$e_{4,T}^n := \sum_{\tau_{i-1}^n < T} \big(\Delta B_{\tau_i^n}^{\mathsf{T}} X_{\tau_{i-1}^n} \Delta B_{\tau_i^n}\big)^2 - \sum_{\tau_{i-1}^n < T} \big(\Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n}\big)^2,$$

$$e_{5,T}^n := \sum_{\tau_{i-1}^n < T} \big(\Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n}\big)^2 - \sum_{\tau_{i-1}^n < T} \big(\Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}\big)^2.$$

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

$$
\begin{aligned}
&\varepsilon_n^{-(2+\delta)} \sum_{\tau_{i-1}^n < T} \bigl(\Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}\bigr)^2 \\
&\quad = \sum_{\tau_i^n < T} \Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n} + \varepsilon_n^{-(2+\delta)} \bigl(\Delta S_T^{\mathsf{T}} \Lambda_{\tau_{N_T^n - 1}^n}^{(n)} \Delta S_T\bigr)^2 \\
&\quad \overset{\text{a.s.}}{\to} \int_0^T \operatorname{Tr}\bigl(\Lambda_t \, \mathrm{d}\langle M\rangle_t\bigr) = \int_0^T \operatorname{Tr}(X_t) \, \mathrm{d}t,
\end{aligned}
$$

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 $\varepsilon_n \le 1$ and since $2 + \delta \le 2\rho_N$)

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

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

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

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

$$
\begin{aligned}
\bigl|\varepsilon_n^{-(2+\delta)} e_{5,T}^n\bigr| &\le \sum_{\tau_{i-1}^n < T} \bigl(\Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} + \Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}\bigr) \\
&\qquad \times \bigl|\Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} - \Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}\bigr| \varepsilon_n^{-(2+\delta)} \\
&\le \sum_{\tau_{i-1}^n < T} \varepsilon_n^{\delta} \chi_{\varepsilon_n^{\delta}}\bigl(\lambda_{\min}(\Lambda_{\tau_{i-1}^n})\bigr) |\Delta S_{\tau_i^n}|^2 \bigl|2\varepsilon_n^{-2-\delta} \Delta S_{\tau_i^n}^{\mathsf{T}} \Lambda_{\tau_{i-1}^n}^{(n)} \Delta S_{\tau_i^n}\bigr| \\
&\le 2\varepsilon_n^{\delta} \sum_{\tau_{i-1}^n < T} |\Delta S_{\tau_i^n}|^2 \overset{\text{a.s.}}{\to} 0
\end{aligned}
$$

thanks to Proposition 3.8.

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

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

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 $n \in \mathbb{N}$

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

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

Moreover by Corollary 3.6 we have

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

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

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

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 $\varepsilon_n^{-(2+\delta)}e_{4,T}^n\to 0$. The proof is finished.

5.5. *Proof of Proposition* 5.2

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

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

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

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

Recall the inequality $\operatorname{Tr}(\mathcal{S}\mathcal{S}')\le\operatorname{Tr}(\mathcal{S})\operatorname{Tr}(\mathcal{S}')$ for any non-negative definite symmetric matrices $\mathcal{S}$ and $\mathcal{S}'$. Thus, $\operatorname{Tr}((\sigma_t\sigma_t^\dagger)^\top\tilde{C}_t(\sigma_t\sigma_t^\dagger))\le d^2\operatorname{Tr}(\tilde{C}_t)$ where we have used the easy inequality $\operatorname{Tr}(\sigma_t\sigma_t^\dagger)\le d$. 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

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

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 $\inf_{t\in[0,T]}\operatorname{Tr}(X_t)>0$ a.s. and we conclude to ($\mathbf{H}_\Lambda$) thanks to (5.19).

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 $\mathcal{A}\in\{D_{x_j},D^2_{x_jx_k},D_t:1\le j,k\le d\}$,

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

Since $\sup_{1\le i\le N_T^n}\Delta\tau_i^n\overset{\text{a.s.}}{\to}0$ 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

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

Hence for $\omega \in \Omega_{\mathcal{C}} \cap \Omega_{\mathcal{D}}$, for n large enough, by a Taylor formula we obtain (the dependence on ω is further omitted, we assume $\omega \in \Omega_{\mathcal{C}} \cap \Omega_{\mathcal{D}}$)

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

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

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

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

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

We are done.

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 ([6, Lemma 2.2]). *Let $\mathcal{C}_0^+$ be the set of non-negative continuous adapted processes, vanishing at $t = 0$. Let $(U^n)_{n\ge0}$ and $(V^n)_{n\ge0}$ 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 $\sum_{n\ge0} V_T^n$ converges a.s.;*
(iii) *there is a constant $c \ge 0$ such that, for every $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $t \in [0, T]$, we have*

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

with the stopping time $\theta_k := \inf\{s \in [0, T] : \bar{V}_s \ge k\}$[2] setting $\bar{V}_t = \sum_{n\ge0} V_t^n$.

Then for any $t \in [0, T]$, the series $\sum_{n\ge0} U_t^n$ converges a.s. As a consequence, $U_t^n \overset{a.s.}{\to} 0$.

References

[1] M. T. Barlow and M. Yor. Semi-martingale inequalities via the Garsia–Rodemich–Rumsey lemma, and applications to local times. *J. Funct. Anal.* **49** (1982) 198–229. MR0680660

[2] J.-P. Bouchaud and M. Potters Financial applications of random matrix theory: A short review. In *The Oxford Handbook of Random Matrix Theory* 824–850. Oxford Univ. Press, Oxford, 2011. MR2932660

[3] M. Fukasawa. Asymptotically efficient discrete hedging. In *Stochastic Analysis with Financial Applications* 331–346. *Progress in Probability* **65**, 2011. MR3050797

[4] M. Fukasawa. Discretization error of stochastic integrals. *Ann. Appl. Probab.* **21** (2011) 1436–1465. MR2857453

[5] C. Geiss and S. Geiss. On approximation of a class of stochastic integrals and interpolation. *Stoch. Stoch. Rep.* **76** (4) (2004) 339–362. MR2075477

[6] E. Gobet and N. Landon. Almost sure optimal hedging strategy. *Ann. Appl. Probab.* **24** (4) (2014) 1652–1690. MR3211007

[7] E. Gobet and N. Landon. Optimization of joint p-variations of Brownian semimartingales. *Electron. J. Probab.* **19** (2014) 36. MR3225867

[8] A. Göing-Jaeschke and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli* **9** (2) (2003) 313–349. MR1997032

[9] M. Hairer, M. Hutzenthaler and A. Jentzen. Loss of regularity for Kolmogorov equations. *Ann. Probab.* **43** (2) (2015) 468–527. MR3305998

[2]With the usual convention $\inf\varnothing = +\infty$.

[10] N. Hoffman, T. Müller-Gronbach and K. Ritter. The optimal discretization of stochastic differential equations. *J. Complexity* **17** (2001) 117–153. MR1817611

[11] J. Jacod and P. Protter. *Discretization of Processes*. *Stochastic Modelling and Applied Probability* **67**. Springer, Heidelberg, 2012. MR2859096

[12] A. Jentzen, T. Müller-Gronbach and L. Yaroslavtseva. On stochastic differential equations with arbitrary slow convergence rates for strong approximation. *Commun. Math. Sci.* **14** (6) (2016) 1477–1500. MR3538358

[13] P. Kree and C. Soize. *Mathematics of Random Phenomena: Random Vibrations of Mechanical Structures*. Springer Science & Business Media, 2012. MR0873731

[14] T. G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** (3) (1991) 1035–1070. MR1112406

[15] J. Li, Y.-B. Peng and J.-B. Chen. Nonlinear stochastic optimal control strategy of hysteretic structures. *Struct. Eng. Mech.* **38** (1) (2011) 39–63.

[16] T. Müller-Gronbach. Strong approximation of systems of stochastic differential equations. Habilitation thesis, Technical University of Darmstadt, 2002. Available at https://www.researchgate.net/publication/34202229_Strong_approximation_of_systems_of_stochastic_differential_equations.

[17] H. Rootzen. Limit distributions for the error in approximations of stochastic integrals. *Ann. Probab.* **8** (2) (1980) 241–251. MR0566591

[18] D. Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields* **8** (2) (2002) 163–198. MR1924934