The Annals of Applied Probability
2014, Vol. 24, No. 4, 1652–1690
DOI: 10.1214/13-AAP959

ALMOST SURE OPTIMAL HEDGING STRATEGY[1]

BY EMMANUEL GOBET AND NICOLAS LANDON

Ecole Polytechnique and CNRS, and Ecole Polytechnique and CNRS and GDF SUEZ

In this work, we study the optimal discretization error of stochastic integrals, in the context of the hedging error in a multidimensional Itô model when the discrete rebalancing dates are stopping times. We investigate the convergence, in an *almost sure* sense, of the renormalized quadratic variation of the hedging error, for which we exhibit an asymptotic lower bound for a large class of stopping time strategies. Moreover, we make explicit a strategy which asymptotically attains this lower bound a.s. Remarkably, the results hold under great generality on the payoff and the model. Our analysis relies on new results enabling us to control a.s. processes, stochastic integrals and related increments.

1. Introduction. *The problem.* We aim at finding a finite sequence of optimal stopping times $\mathcal{T}^n = \{\tau_0^n = 0 < \tau_1^n < \cdots < \tau_i^n < \cdots < \tau_{N_T^n}^n = T\}$ which minimizes the quadratic variation of the discretization error of the stochastic integral

$$Z_s^n = \int_0^s D_x u(t, S_t) \cdot \mathrm{d}S_t - \sum_{\tau_{i-1}^n \le s} D_x u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}\big) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n}),$$

which interpretation is the hedging error [1] of the discrete Delta-hedging strategy of a European option with underlying asset S (multidimensional Itô process), maturity $T > 0$, price function u (for the ease of presentation, here u depends only on S) and payoff $g(S_T)$. The times $(\tau_i^n)_{1 \le i \le N_T^n}$ read as rebalancing dates (or trading dates), and their number N_T^n is a random variable which is finite a.s. The exponent n refers to a control parameter introduced later on; see Section 2. The a.s. minimization of Z_T^n is hopeless since after a suitable renormalization, it is known that it weakly converges to a mixture of Gaussian random variables (see [1, 13, 18, 19] when trading dates are deterministic and under some mild assumptions on the model and payoff; see [9] for stopping times under stronger assumptions). Hence it is more appropriate to investigate the a.s. minimization of the quadratic variation $\langle Z^n \rangle_T$ which, owing to the Lenglart inequality (resp., the Burkholder–Davis–Gundy inequality), allows the control of the distribution (resp., the L_p-moments,

Received January 2012; revised April 2013.
[1]This research is part of the Chair *Financial Risks* of the *Risk Foundation*, the Chair *Derivatives of the Future* and the Chair *Finance and Sustainable Development.*
MSC2010 subject classifications. 60G40, 60F15, 60H05.
Key words and phrases. Almost sure convergence, discretization of stochastic integrals, option hedging, asymptotic optimality.

$p > 0$) of $\sup_{t\le T}|Z_t^n|$ under martingale measure. To avoid trivial lower bounds by letting $N_T^n \to +\infty$, we reformulate our problem into the a.s. minimization of the product

$$N_T^n \langle Z^n\rangle_T. \tag{1.1}$$

As emphasized in [8], the resolution of this optimization problem allows the asymptotic minimization of more general costs of the form $C(N_T^n, \langle Z^n\rangle_T)$, where the function $C : \mathbb{R}^2 \mapsto \mathbb{R}$ is increasing in both variables. Our Theorem 3.1 states that the renormalized error (1.1) has a.s. an asymptotic lower bound over the class of admissible strategies which consist (roughly speaking[2]) of deterministic times and of hitting times of random ellipsoids of the form

$$\tau_0^n := 0, \qquad \tau_i^n := \inf\{t \ge \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n}) \cdot H_{\tau_{i-1}^n}(S_t - S_{\tau_{i-1}^n}) = 1\} \wedge T, \tag{1.2}$$

where $(H_t)_{0\le t\le T}$ is a measurable adapted positive-definite symmetric matrix process. It includes the Karandikar scheme [23] for discretization of stochastic integrals. In addition, in Theorems 3.2 and 3.3 we show the existence of a strategy of the hitting time form attaining the a.s. lower bound. The derivation of a central limit-type theorem for Z^n is left to further research (see [28]), in particular because the verification of the criteria in [9] is difficult to handle in our general setting.

Literature background. Our work extends the existing literature on discretization errors for stochastic integrals with deterministic time mesh, mainly considered with financial applications. Many works deal with hedging rebalancing at regular intervals of length $\Delta t_i = T/n$. In [37] and [1], the authors show that $\mathbb{E}[\langle Z^n\rangle_T]$ converges to 0 at rate n for payoffs smooth enough [this convergence rate originates to consider the product (1.1) as a minimization criterion]. However, in [18] it is proved that the irregularity of the payoff may deteriorate the convergence rate: it becomes $n^{1/2}$ for digital call option. This phenomenon has been intensely analyzed by Geiss and his co-authors using the concept of fractional smoothness (see [10–12, 15] and references therein): by the choice of rebalancing dates suitably concentrated at maturity, we recover the rate n.

The first attempt to find optimal strategies with nondeterministic times goes back to [30]: the authors allow a fixed number n of random rebalancing dates, which actually solve an optimal multiple-stopping problem. Numerical methods are required to compute the solution. In [8], Fukasawa performs an asymptotic analysis for minimizing the product $\mathbb{E}(N_T^n)\mathbb{E}(\langle Z^n\rangle_T)$ (an extension to jump processes has been recently done in [34]). Under regularity and integrability assumptions (and for a convex payoff on a single asset), Fukasawa derives an asymptotic lower bound and provides an optimal strategy. His contribution is the closest to our current work. But there are major differences:

[2]A precise definition is given in Section 2.

(1) We focus on a.s. results, which is probably more meaningful for hedging issues. We are not aware of similar works in this direction.

(2) We allow a quite general model for the asset. It can be a multidimensional diffusion process (local volatility model); see the discussion in Section A.6. As a comparison, in [8] the analysis is carried out for a one-dimensional model (mainly Black–Scholes model).

(3) We also allow a great generality on the payoff. In particular, the payoff can be discontinuous, and the option can be exotic (Asian, lookback, ...) (see Section A.6 for examples): for mathematical reasons, this is a major difference in comparison with [8]. Indeed, in the latter reference, the payoff convexity is needed to ensure the positivity of the option Gamma (second derivative of price), which is a crucial property in the analysis. Also, for discontinuous payoff the L_p integrability of the sensitivities (Greeks) up to maturity may be not satisfied (see [16]); thus, some quantities in the analysis (e.g., the integral of the second moment of the Gamma of digital call option) may become infinite. In our setting, we circumvent these issues by only requiring the sensitivities to be finite a.s. up to maturity: actually, this property is systematically satisfied by payoffs for which the discontinuity set has a zero-measure (see Section A.6), which includes all the usual situations to our knowledge.

To achieve such a level of generality and an a.s. analysis, we design efficient tools to analyze the a.s. control and a.s. convergence of local martingales, of their increments and so forth. All these results represent another important theoretical contribution of this work. Other applications of these techniques are in preparation. At last, although the distribution of hitting time of random ellipsoid of the form (1.2) is not explicit, quite surprisingly we obtain tight estimates on the maximal increments of $\sup_{i \le N_T^n}(\tau_i^n - \tau_{i-1}^n)$, which may have applications in other areas (like stochastic simulation).

Outline of the paper. In the following, we present some notation and assumptions that will be used throughout the paper. Section 2 is aimed at defining our class of stopping time strategies and deriving some general theoretical properties in this class. For that, we establish new key results about a.s. convergence, which fit well our framework. All these results are not specifically related to financial applications. The main results about hedging error are stated and proved in Section 3. Numerical experiments are presented in Section 4, with a practical description of the algorithm to build the optimal sequence of stopping times (actually hitting times) and a numerical illustration regarding the exchange binary option (in dimension 2).

Notation used throughout the paper.

- We denote by $x \cdot y$ the scalar product between two vectors x and y, and by $|x| = (x \cdot x)^{1/2}$ the Euclidean norm of x; the induced norm of a $m \times d$-matrix A is denoted by $|A| := \sup_{x \in \mathbb{R}^d : |x|=1} |Ax|$.
- A^* stands for the transposition of the matrix A; I_d stands for the identity matrix of size d; the trace of a square matrix A is denoted by $\mathrm{Tr}(A)$.

- $\mathcal{S}^d(\mathbb{R})$, $\mathcal{S}^d_+(\mathbb{R})$ and $\mathcal{S}^d_{++}(\mathbb{R})$ are respectively the set of symmetric, symmetric nonnegative-definite and symmetric positive-definite $d \times d$-matrices with coefficients in $\mathbb{R}$: $A \in \mathcal{S}^d_+(\mathbb{R})$ [resp., $\mathcal{S}^d_{++}(\mathbb{R})$] if and only if $x \cdot Ax \geq 0$ (resp., > 0) for any $x \in \mathbb{R}^d \setminus \{0\}$.
- For $A \in \mathcal{S}^d(\mathbb{R})$, $\Lambda(A) := (\lambda_1(A), \dots, \lambda_d(A))$ stands for its spectrum (its $\mathbb{R}$-valued eigenvalues), and we set $\lambda_{\min}(A) := \min_{1\leq i\leq d} \lambda_i(A)$.
- For the partial derivatives of a function $f : (t, x, y) \mapsto f(t, x, y)$, we write $D_t f(t, x, y) = \frac{\partial f}{\partial t}(t, x, y)$, $D_{x_i} f(t, x, y) = \frac{\partial f}{\partial x_i}(t, x, y)$, $D^2_{x_i x_j} f(t, x, y) = \frac{\partial^2 f}{\partial x_i \partial x_j}(t, x, y)$, $D^2_{x_i y_j} f(t, x, y) = \frac{\partial^2 f}{\partial x_i \partial y_j}(t, x, y)$ and so forth.
- When convenient, we adopt the short notation f_t in place of $f(t, S_t, Y_t)$ where f is a given function and $(S_t, Y_t)_{0\leq t\leq T}$ is a continuous time process (introduced below).
- For a $\mathbb{R}^d$-valued continuous semimartingale M, $\langle M\rangle_t$ stands for the matrix of cross-variations $(\langle M^i, M^j\rangle_t)_{1\leq i,j\leq d}$.
- The constants of the multidimensional version of the Burkholder–Davis–Gundy inequalities [25], page 166, are defined as follows: for any $p > 0$ there exists $c_p > 1$ such that for any vector $M = (M^1, \dots, M^d)$ of continuous local martingales with $M_0 = 0$ and any stopping time θ, we have

$$c_p^{-1}\mathbb{E}\left|\sum_{j=1}^{d}\langle M^j\rangle_\theta\right|^p \leq \mathbb{E}\Big(\sup_{t\leq\theta} |M_t|^{2p}\Big) \leq c_p\mathbb{E}\left|\sum_{j=1}^{d}\langle M^j\rangle_\theta\right|^p. \tag{1.3}$$

- For a given sequence of stopping times $\mathcal{T}^n$, the last time before $t \leq T$ is defined by $\varphi(t) = \max\{\tau^n_j; \tau^n_j \leq t\}$: although dependent on n, we omit to indicate this dependency to alleviate notation. Furthermore, for a process $(f_t)_{0\leq t\leq T}$, we write $\Delta f_t := f_t - f_{\varphi(t-)}$ (omitting again the index n for simplicity); in particular, we have $\Delta f_{\tau^n_i} = f_{\tau^n_i} - f_{\tau^n_{i-1}}$. Besides we set $\Delta_t = t - \varphi(t-)$ and $\Delta\tau^n_i := \tau^n_i - \tau^n_{i-1}$.
- We shortly write $X^n \xrightarrow{a.s.}$ if the random variables $(X^n)_{n\geq 0}$ converge almost surely as $n \to \infty$. We write $X^n \xrightarrow{a.s.} X^\infty$ to additionally indicate that the almost sure limit is equal to X^∞. We shall say that the sequence $(X^n)_{n\geq 0}$ is bounded if $\sup_{n\geq 0} |X^n| < +\infty$, a.s.
- C_0 is a a.s. finite nonnegative random variable, which may change from line to line.

Model. Let $T > 0$ be a given terminal time (maturity), 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 $\mathcal{F} = \mathcal{F}_T$. This stochastic basis serves as a modeling of the evolution of d tradable risky assets without dividends, which price processes are denoted by $S = (S^i)_{1\leq i\leq d}$. Their dynamics are given by an Itô continuous semimartingale which solves

$$S_t = S_0 + \int_0^t b_s\,\mathrm{d}s + \int_0^t \sigma_s\,\mathrm{d}B_s \tag{1.4}$$

with measurable and adapted coefficients b and σ. This is the usual framework of complete market; see [31]. Assumptions on σ are given below. Furthermore, for the sake of simplicity we directly assume that the return of the money market account $(r_t)_t$ is zero and that $b \equiv 0$. This simplification is not really a restriction (see [31] for details): indeed, first we can still re-express prices in the money market account numéraire; second, because we deal with a.s. results, we can consider dynamics under any equivalent probability measure, and we choose the martingale measure.

From now on, S is a *continuous local martingale*, and σ satisfies the following assumption.

$(\mathbf{A}_\sigma)$ a.s. for any $t \in [0, T]$ σ_t is nonzero; moreover σ satisfies the continuity condition: there exist a parameter $\theta_\sigma \in (0, 1]$ and a nonnegative a.s. finite random variable C_0 such that

$$|\sigma_t - \sigma_s| \le C_0\big(|S_t - S_s|^{\theta_\sigma} + |t - s|^{\theta_\sigma/2}\big) \qquad \forall 0 \le s, t \le T \text{ a.s.}$$

The above continuity condition is satisfied if $\sigma_t := \sigma(t, S_t)$ for a function $\sigma(\cdot)$ which is θ_σ-Hölder continuous w.r.t. the parabolic distance. For some of our results, the above assumption is strengthened into the following:

$(\mathbf{A}_\sigma^{\text{Ellip.}})$ Assume $(\mathbf{A}_\sigma)$ and that σ_t is elliptic in the sense

$$0 < \lambda_{\min}\big(\sigma_t \sigma_t^*\big) \qquad \forall 0 \le t \le T \text{ a.s.}$$

The assumption $(\mathbf{A}_\sigma^{\text{Ellip.}})$ is undemanding, since we do not suppose any uniform (in ω) lower bound.

We consider an exotic option written on S with payoff $g(S_T, Y_T)$ where Y_T is a functional of $(S_t)_{0 \le t \le T}$. In the subsequent asymptotic analysis, we assume that $Y = (Y^i)_{1 \le i \le d'}$ is a vector of adapted continuous nondecreasing processes. Examples of such an option are given below: this illustrates that the current setting covers numerous relevant situations beyond the case of simple vanilla options [with payoff of form $g(S_T)$].

EXAMPLE 1.1. (1) Asian options: $Y_t^j := \int_0^t S_s^j \,\mathrm{d}s$ and $g(x, y) := (\sum_{1 \le j \le d} \pi_j y^j - K)_+$, for some weights π_j and a given $K \in \mathbb{R}$.

(2) Lookback options: $Y_t^j := \max_{0 \le s \le t} S_s^j$ and $g(x, y) := \sum_{1 \le j \le d} (\pi_j y^j - \pi'_j x^j)$.

Furthermore, we assume that the price at time t of such an option is given by $u(t, S_t, Y_t)$ where u is a $\mathcal{C}^{1,3,1}([0, T[\times \mathbb{R}^d \times \mathbb{R}^{d'})$ function verifying

$$\begin{aligned} &u(T, S_T, Y_T) = g(S_T, Y_T) \quad \text{and} \\ &u(t, S_t, Y_t) = u(0, S_0, Y_0) + \int_0^t D_x u(s, S_s, Y_s) \cdot \mathrm{d}S_s \end{aligned} \tag{1.5}$$

for any $t \in [0, T]$. The above set of conditions is related to probabilistic and analytical properties. First, although not strictly equivalent, it essentially means that the pair (S, Y) forms a Markov process and this originates why the randomness of the fair price $\mathbb{E}(g(S_T, Y_T)|\mathcal{F}_t)$ at time t only comes from (S_t, Y_t). Observe that this Markovian assumption about (S, Y) is satisfied in the above examples. Second, the regularity of the price function u is usually obtained by applying PDE results thanks to Feynman–Kac representations: it is known that the expected regularity can be achieved under different assumptions on the smoothness of the coefficients of S and Y, of the payoff g, combined with some appropriate nondegeneracy conditions on (S, Y). The pictures are multiple, and it is not our current aim to list all the known related results; we refer to [36] for various Feynman–Kac representations related to exotic options, and to [32] for regularity results and references therein. See Section A.6 for extra regularity results. Besides, we assume

($\mathbf{A}_u$) Let $\mathcal{A} \in \mathcal{D} := \{D^2_{x_j x_k}, D^3_{x_j x_k x_l}, D^2_{t x_j}, D^2_{x_j y_m} : 1 \le j, k, l \le d, 1 \le m \le d'\}$,

$$\mathbb{P}\Big(\lim_{\delta\to 0} \sup_{0\le t<T} \sup_{|x-S_t|\le\delta, |y-Y_t|\le\delta} \big|\mathcal{A}u(t, x, y)\big| < +\infty\Big) = 1.$$

Observe that the above assumption is really weak: this is a pathwise result, and we do not require any L_p-integrability of the derivatives of u. In Section A.6, we provide an extended list of payoffs (continuous or not) of options (vanilla, Asian, lookback) in log-normal or local volatility models, for which ($\mathbf{A}_u$) holds. Even for the simple option payoff $g(S_T)$ in the simple log-normal model, we have not been able to exhibit a payoff function g for which ($\mathbf{A}_u$) is not satisfied.

2. Class $\mathcal{T}^{\mathrm{adm.}}$ of strategies and convergence results. In this section, we define the class of strategies under consideration, and establish some preliminary almost sure convergence results in connection with this class.

A strategy is a finite sequence of increasing stopping times $\{\tau_0 = 0 < \tau_1 < \cdots < \tau_i < \cdots < \tau_{N_T} = T\}$ (with $N_T < +\infty$ a.s.) which stand for the rebalancing dates. Furthermore, the number of risky assets held on each interval $[\tau_i, \tau_{i+1})$ follows the usual Delta-neutral rule $D_x u(\tau_i, S_{\tau_i}, Y_{\tau_i})$.

2.1. *Assumptions.* Now to derive *asymptotically* optimal results, we consider a sequence of strategies indexed by the integers $n = 0, 1, \ldots,$ that is, writing

$$\mathcal{T}^n := \{\tau_0^n = 0 < \tau_1^n < \cdots < \tau_i^n < \cdots < \tau^n_{N^n_T}\} \qquad \text{for } n = 0, 1, \ldots,$$

and we define an appropriate *asymptotic framework*, as the convergence parameter n goes to infinity. Let $(\varepsilon_n)_{n\ge 0}$ be a sequence of positive deterministic real numbers converging to 0 as $n \to \infty$; assume that it is a square-summable sequence

$$\sum_{n\ge 0} \varepsilon_n^2 < +\infty. \tag{2.1}$$

On the one hand, the parameter ${\varepsilon_n}^{-2\rho_N}$ (for some $\rho_N \geq 1$) upper bounds (up to a constant) the number of rebalancing dates of the strategy $\mathcal{T}^n$, that is:

($\mathbf{A}_N$) The following nonnegative random variable is a.s. finite:

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

for a parameter ρ_N satisfying $1 \leq \rho_N < (1 + \frac{\theta_\sigma}{2}) \wedge \frac{4}{3}$.

On the other hand, the parameter ε_n controls the size of variations of S between two stopping times in $\mathcal{T}^n$.

($\mathbf{A}_S$) The following nonnegative random variable is a.s. finite:

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

Observe that assumptions ($\mathbf{A}_N$) and ($\mathbf{A}_S$) play complementary (and not equivalent) roles. We are now ready to define the class of sequence of strategies in which we are seeking the optimal element.

DEFINITION 2.1. A sequence of strategies $\mathcal{T} := \{\mathcal{T}^n : n \geq 0\}$ is *admissible* if it fulfills the hypotheses ($\mathbf{A}_N$) and ($\mathbf{A}_S$). The set of admissible sequences $\mathcal{T}$ is denoted by $\mathcal{T}^{\mathrm{adm.}}$.

The above definition depends on the sequence $(\varepsilon_n)_{n\geq 0}$, which is fixed from now on.

REMARK 2.1.

- The larger ρ_N, the wider the class of strategies under consideration. The choice $\rho_N = 1$ is allowed, but seemingly it rules out deterministic strategies; see the next remark.
- If $\rho_N > 1$, a strategy $\mathcal{T}^n$ consisting of $N_T^n = 1 + \lfloor {\varepsilon_n}^{-2\rho_N} \rfloor$ deterministic times with mesh size $\sup_{1\leq i\leq N_T^n} \Delta\tau_i^n \leq C\varepsilon_n^{2\rho_N}$ (this includes the cases of uniform and some nonuniform time grids) forms an admissible sequence of strategies, thanks to the $\frac{1}{2}^-$-Hölder property of the Dambis–Dubins–Schwarz Brownian motion of S^j ($1 \leq j \leq d$) (under the additional assumption that σ is uniformly bounded to safely maintain the time-changes into a fixed compact interval).
- Our setting allows us to consider stopping times satisfying the *strong predictability condition* (i.e., τ_i^n is $\mathcal{F}_{\tau_{i-1}^n}$-measurable); see [21], Chapter 14.
- We show in Proposition 2.4 that the strategy $\mathcal{T}^n$ of successive hitting times of ellipsoid of size ε_n forms a sequence in $\mathcal{T}^{\mathrm{adm.}}$.

- In Sections 2.3–2.4, we investigate properties of admissible sequences of strategies. Among others, we show that the mesh size of $\mathcal{T}^n$ shrinks a.s. to 0, and we establish tight a.s. upper bounds (see Corollary 2.2): namely for any $\rho \in (0, 2]$, there is a a.s. finite random variable C_ρ such that $\sup_{1\le i\le N_T^n} \Delta\tau_i^n \le C_\rho \varepsilon_n^{2-\rho}$ for any $n \ge 0$.

We require an extra technical condition on the nondecreasing process Y which is fulfilled in practical cases for an admissible sequence of strategies.

$(\mathbf{A}_Y)$ The following nonnegative random variable is a.s. finite: for some $\rho_Y > 4(\rho_N - 1)$

$$\sup_{n\ge 0}\Big(\varepsilon_n^{-\rho_Y} \sup_{1\le i\le N_T^n} |\Delta Y_{\tau_i^n}|\Big) < +\infty.$$

EXAMPLE 2.1. Let $\mathcal{T} := \{\mathcal{T}^n : n \ge 0\}$ satisfy $(\mathbf{A}_S)$–$(\mathbf{A}_N)$.

(1) Asian options: applying Corollary 2.2 [item (ii)] with $\rho = \frac{2}{3}$ and taking $\rho_Y = \frac{4}{3} > 4(\rho_N - 1)$ (since $\rho_N < \frac{4}{3}$) gives

$$\sup_{n\ge 0}\Big(\varepsilon_n^{-\rho_Y} \sup_{1\le i\le N_T^n} |\Delta Y_{\tau_i^n}|\Big) \le \sup_{0\le t\le T} |S_t| \sup_{n\ge 0}\Big(\varepsilon_n^{\rho-2} \sup_{1\le i\le N_T^n} \Delta\tau_i^n\Big) < +\infty \qquad \text{a.s.}$$

(2) Lookback options: clearly, we have

$$\sup_{n\ge 0}\Big(\varepsilon_n^{-1} \sup_{1\le i\le N_T^n} |\Delta Y_{\tau_i^n}|\Big) \le \sup_{n\ge 0}\Big(\varepsilon_n^{-1} \sup_{0\le t\le T} |\Delta S_t|\Big) < +\infty \qquad \text{a.s.;}$$

thus $(\mathbf{A}_Y)$ is satisfied with $\rho_Y = 1$ provided that $\rho_N < 5/4$.

2.2. *Fundamental lemmas about almost sure convergence.* This subsection is devoted to the main ingredient (Lemmas 2.1 and 2.2) about almost sure convergence, which is involved in the subsequent asymptotic analysis.

We first recall some usual approaches to establish that a sequence $(U_T^n)_{n\ge 0}$ converges to 0 in probability or almost surely, as $n \to \infty$: it serves as a preparation for the comparative discussion we will have regarding our almost sure convergence results.

- *Convergence in probability.* It can be handled, for instance, by using the Markov inequality and showing that the L_p-moment (for some $p > 0$) of U_T^n converges to 0: for $p = 1$ and $\delta > 0$, it writes $\mathbb{P}(|U_T^n| \ge \delta) \le \frac{\mathbb{E}|U_T^n|}{\delta} \to_{n\to\infty} 0$. Observe that this approach requires a bit of integrability of the random variable U_T^n.

  To achieve the uniform convergence in probability of $(U_t^n)_{0\le t\le T}$ to 0, Lenglart [29] introduced an extra condition: the relation of domination. Namely, assume that $(U_t^n)_{0\le t\le T}$ is a nonnegative continuous adapted process and that it is dominated by a nondecreasing continuous adapted process $(V_t^n)_{0\le t\le T}$ (with

$V_0^n = 0$) in the sense $\mathbb{E}(U_\theta^n) \leq \mathbb{E}(V_\theta^n)$ for any stopping time $\theta \in [0, T]$. Then, for any $c_1, c_2 > 0$, we have

$$\mathbb{P}\Big(\sup_{t \leq T} U_t^n \geq c_1\Big) \leq \frac{1}{c_1}\mathbb{E}\big(V_T^n \wedge c_2\big) + \mathbb{P}\big(V_T^n \geq c_2\big).$$

A standard application consists in taking U^n as the square of a continuous local martingales M^n; then, the convergence in probability of $\langle M^n, M^n\rangle_T$ to 0 implies the uniform convergence in probability of $(M_t^n)_{0\leq t\leq T}$ to 0. The converse is also true, the relation of domination deriving from BDG inequalities. This kind of result leads to useful tools for establishing the convergence in probability of triangular arrays of random variables: for instance, see [14], Lemma 9, in the context of parametric estimation of stochastic processes.

- *Almost sure convergence.* We may use a Borel–Cantelli type argument, assuming that $\sum_{n\geq 0} \mathbb{E}|U_T^n| < +\infty$. Fubini–Tonelli's theorem yields that the series $\sum_{n\geq 0} |U_T^n|$ converges a.s., and in particular $U_T^n \xrightarrow{a.s.} 0$. Here again, the integrability of U_T^n is required.

  Bichteler and Karandikar leveraged this type of series argument to establish the a.s. convergence of stochastic integrals under various assumptions, with in view either approximation issues or pathwise stochastic integration; see [2, 22–24] and references therein.

Our result below (Lemma 2.1) is inspired by the above references, but its conditions of applicability are less stringent, and it allows more flexibility in our framework. We assume a relation of domination, but:

(1) not for all stopping times (as in Lenglart domination);
(2) the processes $(U_t^n)_{0\leq t\leq T}$ are not assumed to be continuous [nor $(\sum_{n\geq 0} U_t^n)_{0\leq t\leq T}$];
(3) the dominating process V^n is not assumed to be nondecreasing.

Thus, our assumptions are less demanding, but on the other hand, we do not obtain any uniform convergence result. Moreover, we emphasize that we do not assume any integrability on U_T^n. This is crucial, because the typical applications of Lemma 2.1 are related to U_T^n defined as a (possibly stochastic) integral of the derivatives of u evaluated along the path $(S_t, Y_t)_{0\leq t\leq T}$: since usual payoff functions are irregular, it is known that the L_p-moments of related derivatives blow up as time goes to maturity, and it is hopeless to obtain the required integrability on U_T^n assuming only $(\mathbf{A}_u)$.

We are now ready for the statement of our a.s. convergence result.

LEMMA 2.1. *Let $\mathcal{M}_0^+$ be the set of nonnegative measurable processes vanishing at $t = 0$. Let $(U^n)_{n\geq 0}$ and $(V^n)_{n\geq 0}$ be two sequences of processes in $\mathcal{M}_0^+$. Assume that*:

(i) *the series* $\sum_{n\geq 0} V_t^n$ *converges for all* $t \in [0, T]$, *almost surely*;

(ii) *the above limit is upper bounded by a process* $\bar{V} \in \mathcal{M}_0^+$ *and that* $\bar{V}$ *is continuous 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*

$$\mathbb{E}[U^n_{t\wedge\theta_k}] \leq c\mathbb{E}[V^n_{t\wedge\theta_k}]$$

with the random time $\theta_k := \inf\{s \in [0, T] : \bar{V}_s \geq k\}$.[3]

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

PROOF. First, observe that $(\theta_k)_{k\geq 0}$ defines well random times since $\bar{V}$ is continuous.

Denote by $\mathcal{N}_V$ the $\mathbb{P}$-negligible set on which the series $(\sum_{n\geq 0} V_t^n)_{0\leq t\leq T}$ do not converge, and on which $\bar{V}$ and then $(\theta_k)_{k\geq 0}$ are not defined; observe that for $\omega \notin \mathcal{N}_V$, we have $\bar{V}_{t\wedge\theta_k}(\omega) \leq k$ for any $t \in [0, T]$ and $k \in \mathbb{N}$. Set $\bar{V}^p := \sum_{n=0}^p V^n$: we have $\bar{V}^p \leq \bar{V}$ on $\mathcal{N}_V^c$; thus, the localization of $\bar{V}$ entails that of $\bar{V}^p$ and we have $\bar{V}^p_{t\wedge\theta_k} \leq k$ for any k, p and t (on $\mathcal{N}_V^c$).

Moreover, for any n and k, the relation of domination writes

$$\mathbb{E}\left[\sum_{n=0}^p U^n_{t\wedge\theta_k}\right] \leq c\mathbb{E}\left[\sum_{n=0}^p V^n_{t\wedge\theta_k}\right] = c\mathbb{E}[\bar{V}^p_{t\wedge\theta_k}] \leq ck. \tag{2.2}$$

From Fatou's lemma, we get $\mathbb{E}[\sum_{n\geq 0} U^n_{t\wedge\theta_k}] < +\infty$: in particular, for any $k \in \mathbb{N}$, there is a $\mathbb{P}$-negligible set $\mathcal{N}_{k,t}$, such that $\sum_{n\geq 0} U^n_{t\wedge\theta_k}(\omega)$ converges for all $\omega \notin \mathcal{N}_{k,t}$. The set $\mathcal{N}_t = \bigcup_{k\in\mathbb{N}} \mathcal{N}_{k,t} \cup \mathcal{N}_V$ is $\mathbb{P}$-negligible, and it follows that for $\omega \notin \mathcal{N}_t$, the series $\sum_{n\geq 0} U^n_{t\wedge\theta_k}(\omega)$ converges for all $k \in \mathbb{N}$. For $\omega \notin \mathcal{N}_t$, we have $\theta_k(\omega) = +\infty$ as soon as $k > \bar{V}_T(\omega)$; thus by taking such k, we complete the convergence of $\sum_{n\geq 0} U_t^n$ on $\mathcal{N}_t^c$. $\square$

Observe that in our argumentation, we do not assume that the nonnegative random variables U_t^n and V_t^n have a finite expectation (and in some examples, it is false, especially at $t = T$). However, note that in (2.2) we prove that $U^n_{t\wedge\theta_k}$ and $V^n_{t\wedge\theta_k}$ have a finite expectation: in other words, $(\theta_k)_{k\geq 0}$ serves as a common localization for U^n and V^n. In addition, Lemma 2.1 is general and thorough since we do not assume any adaptedness or regularity properties of the processes U^n and V^n. We provide a simpler version that can be customized for our further applications:

LEMMA 2.2. *Let* $\mathcal{C}_0^+$ *be the set of nonnegative 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* $\mathcal{C}_0^+$. *Replace the two first items of Lemma* 2.1 *by*:

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

(i′) *$t \mapsto V_t^n$ is a nondecreasing function on* $[0, T]$, *almost surely*;

(ii′) *the series* $\sum_{n\geq 0} V_T^n$ *converges almost surely*;

(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*

$$\mathbb{E}[U_{t\wedge\theta_k}^n] \leq c\mathbb{E}[V_{t\wedge\theta_k}^n] \tag{2.3}$$

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

Then, *the conclusion of Lemma* 2.1 *still holds*.

PROOF. We just have to prove that items (i′) + (ii′) entails items (i) + (ii) of Lemma 2.1 for U^n and V^n in $\mathcal{C}_0^+ \subset \mathcal{M}_0^+$. Since V^n is nondecreasing, the a.s. convergence of $\sum_{n\geq 0} V_T^n$ implies that of $\sum_{n\geq 0} V_t^n$. Moreover $\sum_{n\geq 0} \sup_{0\leq t\leq T} V_t^n = \sum_{n\geq 0} V_T^n < +\infty$ a.s. Therefore, a.s. the series associated with V^n is normally convergent on $[0, T]$ and $\bar{V} := \sum_{n\geq 0} V^n \in \mathcal{C}_0^+$: items (i) + (ii) are satisfied. Observe θ_k is a *stopping* time since $\bar{V}$ is continuous and adapted. □

We apply Lemma 2.2 to derive a simple criterion for the convergence of continuous local martingales.

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

$$\sum_{n\geq 0} \langle M^n\rangle_T^{p/2} \xrightarrow{a.s.} \quad \Longleftrightarrow \quad \sum_{n\geq 0} \sup_{0\leq t\leq T} |M_t^n|^p \xrightarrow{a.s.}.$$

PROOF. We first prove the implication $\Rightarrow$. Set $U_t^n := \sup_{0\leq s\leq t} |M_s^n|^p$ and $V_t^n := \langle M^n\rangle_t^{p/2}$, and let us check the conditions of Lemma 2.2: (i′) V^n is nondecreasing and (ii′) $\sum_{n\geq 0} V_T^n$ converges a.s. The relation of domination (2.3) follows from the BDG inequalities [see the RHS of (1.3)] and we are done. The implication $\Leftarrow$ is proved similarly, using the LHS of (1.3) regarding the BDG inequalities. □

2.3. *Controls of* $\Delta\tau^n$ *and of the martingales increments*. Being inspired by the scaling property of Brownian motion, we might intuitively guess that a sequence of strategy $(\mathcal{T}^n)_{n\geq 0}$ satisfying $(\mathbf{A}_S)$ yields stopping times increments of magnitude equal roughly to ε_n^2. Actually, thorough estimates are difficult to derive: for instance, the exit times of balls by a Brownian motion define unbounded random variables.

To address these issues, we take advantage of Lemma 2.2 to establish estimates on the sequence $(\Delta\tau_i^n := \tau_i^n - \tau_{i-1}^n)_{1\leq i\leq N_T^n}$, which show that we almost recover the familiar scaling ε_n^2.

PROPOSITION 2.1. *Assume* $(\mathbf{A}_\sigma)$. *Let* $\mathcal{T}$ *be a sequence of strategies satisfying* $(\mathbf{A}_S)$ *and let* $p \geq 0$. *Then*:

(i) *The series* $\sum_{n\ge 0}\varepsilon_n^{-(p-2)}\sup_{1\le i\le N_T^n}(\Delta\tau_i^n)^p \xrightarrow{a.s.}$.

(ii) *Assume moreover that* $\mathcal{T}\in\mathcal{T}^{\mathrm{adm.}}$: *the series* $\sum_{n\ge 0}\varepsilon_n^{-2(p-1)+2\rho_N}\times\sum_{\tau_{i-1}^n<T}(\Delta\tau_i^n)^p \xrightarrow{a.s.}$.

The proof is postponed to Appendix A.1. As a consequence of Proposition 2.1, the mesh size of $\mathcal{T}^n$, that is, $\sup_{1\le i\le N_T^n}\Delta\tau_i^n$, converges a.s. to 0 as $n\to\infty$, with some explicit rates of convergence: this is the statement below.

COROLLARY 2.2. *With the same assumptions and notation as Proposition* 2.1, *we have the following estimates*, *for any* $\rho>0$:

(i) *Under* $(\mathbf{A}_S)$, $\sup_{n\ge 0}(\varepsilon_n^{\rho-1}\sup_{1\le i\le N_T^n}\Delta\tau_i^n)<+\infty$ *a.s.*

(ii) *Under* $(\mathbf{A}_S)$–$(\boldsymbol{A}_N)$, $\sup_{n\ge 0}(\varepsilon_n^{\rho-2}\sup_{1\le i\le N_T^n}\Delta\tau_i^n)<+\infty$ *a.s.*

PROOF. Item (i). Clearly, from Proposition 2.1(i), we obtain $\sup_{n\ge 0}(\varepsilon_n^{-(p-2)}\times\sup_{1\le i\le N_T^n}(\Delta\tau_i^n)^p)<+\infty$ a.s. for any $p\ge 0$ and the result follows by taking $p=2/\rho$.

Item (ii). We proceed similarly by observing that Proposition 2.1(ii) gives

$$\begin{aligned}\sup_{n\ge 0}\Big(\varepsilon_n^{-2(p-1-\rho_N)}\sup_{1\le i\le N_T^n}(\Delta\tau_i^n)^p\Big)&\le\sup_{n\ge 0}\Big(\varepsilon_n^{-2(p-1-\rho_N)}\sum_{\tau_{i-1}^n<T}(\Delta\tau_i^n)^p\Big)\\&<+\infty\qquad\text{a.s.}\end{aligned}$$

□

We are now in a position to control the a.s. convergence of some stochastic integrals appearing in our further optimality analysis. The following proposition and corollary will play a crucial role in the estimations of the error terms appearing in the main theorems; see Section 3.

PROPOSITION 2.2. *Assume* $(\mathbf{A}_\sigma)$. *Let* $\mathcal{T}=(\mathcal{T}^n)_{n\ge 0}$ *be a sequence of strategies*, $((M_t^n)_{0\le t\le T})_{n\ge 0}$ *be a sequence of* $\mathbb{R}$-*valued continuous local martingales such that* $\langle M^n\rangle_t=\int_0^t\alpha_r^n\,\mathrm{d}r$ *for a nonnegative measurable adapted* α^n *satisfying the following inequality*: *there exists a nonnegative a.s. finite random variable* C_α *and a parameter* $\theta\ge 0$ *such that*

$$0\le\alpha_r^n\le C_\alpha\big(|\Delta S_r|^{2\theta}+|\Delta r|^{\theta}\big)\qquad\forall 0\le r<T,\forall n\ge 0,\ a.s.$$

Then, *the following convergences hold*:

(i) *Assume* $\mathcal{T}$ *satisfies* $(\mathbf{A}_S)$ *and let* $p\ge 2$

$$\sum_{n\ge 0}\Big(\varepsilon_n^{3-((1+\theta)/2)p}\sum_{\tau_{i-1}^n<T}\sup_{\tau_{i-1}^n\le t\le\tau_i^n}|\Delta M_t^n|^p\Big)<+\infty\qquad a.s.$$

(ii) *Assume furthermore that $\mathcal{T}$ satisfies* $(\mathbf{A}_N)$ (*i.e.*, $\mathcal{T} \in \mathcal{T}^{\mathrm{adm.}}$), *and let* $p > 0$

$$\sum_{n \geq 0} \left(\varepsilon_n^{2-(1+\theta)p+2\rho_N} \sum_{\tau_{i-1}^n < T} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta M_t^n|^p \right) < +\infty \qquad a.s.$$

The proof is postponed in Appendix A.2. A straightforward consequence of the aforementioned proposition is given by the following corollary, which proof is left to the reader.

COROLLARY 2.3. *Using the assumptions and notation of Proposition* 2.2, *we have the following estimates*, *for any* $\rho > 0$:

(i) *Under* $(\mathbf{A}_S)$, $\sup_{n \geq 0}(\varepsilon_n^{\rho-(1+\theta)/2} \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta M_t^n|) < +\infty$, *a.s*.

(ii) *Under* $(\mathbf{A}_S)$–$(\mathbf{A}_N)$, $\sup_{n \geq 0}(\varepsilon_n^{\rho-(1+\theta)} \sup_{1 \leq i \leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |\Delta M_t^n|) < +\infty$, *a.s*.

REMARK 2.2. Observe that in the proofs of the Section 2.3 results, we have not used the knowledge of the upper bound on ρ_N [stated in $(\mathbf{A}_N)$]: it means that all the related results are true for any admissible sequence of strategies assuming only $\rho_N \geq 1$.

2.4. *Almost sure convergence of weighted discrete quadratic variation.*

PROPOSITION 2.3. *Assume* $(\mathbf{A}_\sigma)$ *and let* $\mathcal{T}$ *be a sequence of strategies satisfying* $(\mathbf{A}_S)$. *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* $(M_t)_{0 \leq t \leq T}$ *be a* $\mathbb{R}^d$-*valued continuous local martingale such that* $\langle M \rangle_t = \int_0^t \alpha_r \,\mathrm{d}r$ *with* $\sup_{0 \leq t \leq T} |\alpha_t| < +\infty$ *a.s. Then*

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

PROOF. From Itô's lemma, $\sum_{\tau_{i-1}^n < T} \Delta M_{\tau_i^n}^* H_{\tau_{i-1}^n} \Delta M_{\tau_i^n}$ is equal to

$$\begin{aligned}
&\sum_{k,l=1}^d \sum_{\tau_{i-1}^n < T} \Delta M_{\tau_i^n}^k H_{\tau_{i-1}^n}^{k,l} \Delta M_{\tau_i^n}^l \\
&\qquad = \sum_{k,l=1}^d \int_0^T H_{\varphi(t)}^{k,l} \big(\Delta M_t^k \,\mathrm{d}M_t^l + \Delta M_t^l \,\mathrm{d}M_t^k + \mathrm{d}\langle M^k, M^l \rangle_t\big) \\
&\qquad = \int_0^T \Delta M_t^* \big(H_{\varphi(t)} + H_{\varphi(t)}^*\big) \,\mathrm{d}M_t + \int_0^T \mathrm{Tr}\big(H_{\varphi(t)} \,\mathrm{d}\langle M \rangle_t\big).
\end{aligned}$$

The second term in the above RHS converges a.s. to $\int_0^T \mathrm{Tr}(H_t\,\mathrm{d}\langle M\rangle_t)$: indeed, the difference is bounded by $C_0\int_0^T |H_t - H_{\varphi(t)}|\,\mathrm{d}t$, and we conclude by an application of the dominated convergence theorem, invoking the continuity and boundedness of H and the convergence to 0 of the mesh size of $\mathcal{T}^n$; see Corollary 2.2.

Thus it remains to show that the stochastic integral w.r.t. $\mathrm{d}M_t$ converges a.s. to 0. Owing to Corollary 2.1, it is enough to study the series of quadratic variations, that is, to show that $\sum_{n\geq 0}[\int_0^T(\Delta M_t^*(H_{\varphi(t)} + H^*_{\varphi(t)})\,\mathrm{d}\langle M\rangle_t(H_{\varphi(t)} + H^*_{\varphi(t)})\Delta M_t)]^3 \xrightarrow{a.s.}$, and since α and H are a.s. bounded on $[0, T)$, it is sufficient to show

$$\sum_{n\geq 0}\biggl[\int_0^T |\Delta M_t|^2\,\mathrm{d}t\biggr]^3 \xrightarrow{a.s.} . \tag{2.4}$$

Clearly $[\int_0^T |\Delta M_t|^2\,\mathrm{d}t]^3$ is bounded by

$$d^3T^3 \sup_{1\leq j\leq d}\ \sup_{1\leq i\leq N_T^n}\ \sup_{\tau_{i-1}^n\leq t\leq \tau_i^n} |\Delta M_t^j|^6 \leq C_0\varepsilon_n^2$$

owing to Corollary 2.3 [item (i)] for $\theta = 0$ and $\rho = \frac{1}{6}$. The convergence (2.4) is proved, and we are done. □

2.5. *Verification of the hypothesis on a special family of hitting times.* One of the more appealing results of the paper is that a very large family of hitting times fulfills the assumptions $(\mathbf{A}_N)$ and $(\mathbf{A}_S)$ with a threshold depending of ε_n.

PROPOSITION 2.4. *Assume* $(\mathbf{A}_\sigma)$. *Let* $(H_t)_{0\leq t<T}$ *be a continuous adapted nonnegative-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*

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

defines a sequence of strategies satisfying assumptions $(\mathbf{A}_N)$ [*with* $\sup_{n\geq 0}(\varepsilon_n^2 N_T^n) < +\infty$ *a.s.*] *and* $(\mathbf{A}_S)$, *that is* $\{\mathcal{T}^n : n\geq 0\}\in \mathcal{T}^{\mathrm{adm.}}$.

The proof is postponed to Appendix A.3. Observe that the above sequence of strategies is admissible even in the most constrained case $\rho_N = 1$. As we shall see later on, the optimal stopping times are given by the hitting times by the process S of an ellipsoid (corresponding to the case H symmetric).

3. Main results.

3.1. *Statements.* We now go back to the hedging issue: at time $s \in [0, T]$, the fair value of the option is $u(s, S_s)$, and the hedging portfolio with discrete rebalancing dates $\mathcal{T}^n$ is $u(0, S_0) + \sum_{\tau_{i-1}^n \le s} D_x u(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n})$, which yields an hedging error equal to

$$
\begin{aligned}
Z_s^n &:= u(s, S_s) - \Big(u(0, S_0) + \sum_{\tau_{i-1}^n \le s} D_x u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}\big) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n}) \Big) \\
&= \int_0^s (D_x u_t - D_x u_{\varphi(t)}) \cdot \mathrm{d}S_t
\end{aligned}
\tag{3.1}
$$

using (1.5), where the integrand appears as the difference of Delta between τ_{i-1}^n and $t \in]\tau_{i-1}^n, \tau_i^n]$ for each $0 \le i \le N_T^n$.

One main result of the paper is a lower bound of the renormalized quadratic variation of the hedging error Z^n: it is partly derived from a *smart* representation of

$$
\langle Z^n \rangle_T = \int_0^T (D_x u_t - D_x u_{\varphi(t)})^* \,\mathrm{d}\langle S \rangle_t (D_x u_t - D_x u_{\varphi(t)}) \tag{3.2}
$$

as a sum of squared random variables and an application of the Cauchy–Schwarz inequality. To derive this suitable representation, we apply the Itô formula and identify the bounded variation term; it is straightforward in dimension one, much more intricate in a multidimensional setting and this is equivalent to solving the following matrix equation.

LEMMA 3.1. *Let $c \in \mathcal{S}^d(\mathbb{R})$. Then the equation*

$$
2\,\mathrm{Tr}(x)x + 4x^2 = c^2 \tag{3.3}
$$

admits exactly one solution $x(c) \in \mathcal{S}_+^d(\mathbb{R})$. In addition, $x(c)$ is positive-definite if and only if c^2 is positive-definite. Last, the mapping $c \mapsto x(c)$ is continuous.

The proof is given in Section A.4. We are now in a position to give an explicit asymptotic lower bound for $N_T^n \langle Z^n \rangle_T$: this is the contents of the following theorem.

THEOREM 3.1. *Assume assumptions $(\mathbf{A}_\sigma)$, $(\mathbf{A}_u)$, $(\mathbf{A}_S)$, $(\mathbf{A}_N)$ and $(\mathbf{A}_Y)$ are in force. Let X be the solution of (3.3) with $c := \sigma^* D_{xx}^2 u \sigma$. Then*

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

Let us comment a bit on the above lower bound:

- First, it is a.s. finite: indeed, $\sup_{t<T} |\sigma_t^* D_{xx}^2 u_t \sigma_t| < +\infty$ a.s., and the continuity of $c \mapsto x(c)$ imply $\sup_{t<T} |X_t| < +\infty$ a.s.
- Second, observe that a.s.

$$\left\{\int_0^T \operatorname{Tr}(X_t)\,\mathrm{d}t = 0\right\} = \{\forall t < T : \sigma_t^* D_{xx}^2 u_t \sigma_t = 0\}$$

$$\overset{\text{under } (\mathbf{A}_\sigma^{\text{Ellip.}})}{=} \{\forall t < T : D_{xx}^2 u_t = 0\}$$

using at the first equality that $\operatorname{Tr}(x(c)) > 0 \Leftrightarrow x(c) \neq 0 \Leftrightarrow c \neq 0$. Then we obtain that except in degenerate situations [where the Gamma matrix $D_{xx}^2 u_t$ is zero at any time, assuming $(\mathbf{A}_\sigma^{\text{Ellip.}})$], the lower bound in Theorem 3.1 is nonzero.
- As a consequence, we immediately obtain a lower bound for the L_p-criterion: indeed, using the Fatou lemma and the Cauchy–Schwarz inequality, we derive (for any $p > 0$)

$$\begin{aligned}\left[\mathbb{E}\left(\int_0^T \operatorname{Tr}(X_t)\,\mathrm{d}t\right)^p\right]^2 &\le \left[\mathbb{E}\left(\liminf_{n\to+\infty}\left(N_T^n \langle Z^n\rangle_T\right)^{p/2}\right)\right]^2 \\ &\le \liminf_{n\to+\infty}\left[\mathbb{E}\left(N_T^n \langle Z^n\rangle_T\right)^{p/2}\right]^2 \\ &\le \liminf_{n\to+\infty} \mathbb{E}\left((N_T^n)^p\right)\mathbb{E}\left(\langle Z^n\rangle_T^p\right).\end{aligned}$$

For $p = 1$ we recover the Fukasawa approach [8].

The next theorem tells us that along a suitable sequence $\mathcal{T}^n$ (the hitting times of some random ellipsoids) the lower bound of Theorem 3.1 is reached. Let $\chi(\cdot)$ be a smooth function such that $\mathbf{1}_{]-\infty,1/2]} \le \chi(\cdot) \le \mathbf{1}_{]-\infty,1]}$ and for $\mu > 0$, set $\chi_\mu(x) = \chi(x/\mu)$.

THEOREM 3.2. *Assume assumptions* $(\mathbf{A}_\sigma^{\text{Ellip.}})$, $(\mathbf{A}_u)$, $(\mathbf{A}_S)$, $(\mathbf{A}_N)$ *and* $(\mathbf{A}_Y)$ *are in force. Let* $\mu > 0$, *for* $t \ge 0$ *set* $\Lambda_t := (\sigma_t^{-1})^* X_t \sigma_t^{-1}$ *and* $\Lambda_t^\mu := \Lambda_t + \mu\chi_\mu(\lambda_{\min}(\Lambda_t)) I_d$.

For a given $n \in \mathbb{N}$, *define the strategy* $\mathcal{T}_\mu^n$ *by*

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n = \inf\{t \ge \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^* \Lambda_{\tau_{i-1}^n}^\mu (S_t - S_{\tau_{i-1}^n}) > \varepsilon_n^2\} \wedge T. \end{cases} \tag{3.4}$$

Then, the sequence of strategies $\mathcal{T}_\mu = \{\mathcal{T}_\mu^n : n \ge 0\}$ *is admissible, and it is* μ-*asymptotically optimal in the following sense*:

$$\limsup_{n\to+\infty}\left|N_T^n \langle Z^n\rangle_T - \left(\int_0^T \operatorname{Tr}(X_t)\,\mathrm{d}t\right)^2\right| \le C_\mu \mu \int_0^T \chi_\mu(\lambda_{\min}(\Lambda_t)) \operatorname{Tr}(\sigma_t\sigma_t^*)\,\mathrm{d}t,$$

where the random variable $C_\mu := \int_0^T (4\operatorname{Tr}(X_t) + 3\mu\chi_\mu(\lambda_{\min}(\Lambda_t))\operatorname{Tr}(\sigma_t\sigma_t^*))\,\mathrm{d}t$ *is a.s. finite* (*locally uniformly w.r.t.* $\mu \ge 0$).

In particular, on the event $\{\forall t \in [0,T] : \lambda_{\min}(\Lambda_t) \geq \mu\}$, $N_T^n \langle Z^n \rangle_T$ *converges a.s. to* $(\int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t)^2$.

Observe that we require the ellipticity condition to hold. The proof is given in Section 3.3.

We can strengthen the above theorem by allowing $\mu = 0$ under stronger assumptions.

THEOREM 3.3. *Assume the assumptions of Theorem* 3.2 *and additionally that*

$$\mathbb{P}\Big(\inf_{t\in[0,T[}\lambda_{\min}(D^2_{xx}u_t) > 0\Big) = 1. \tag{3.5}$$

Then, the sequence of strategies $\mathcal{T}_0 = \{\mathcal{T}^n(0) : n \geq 0\}$ *defined in* (3.4) *with* $\mu = 0$ *is admissible and asymptotically optimal*,

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

For the proof, see Section 3.4. The extra assumption (3.5) is satisfied in dimension one for call/put option in Black–Scholes model only if the hedging time horizon is strictly smaller than the option maturity. But it is not satisfied in digital call/put option. This discussion can be extented to higher multidimensional situations.

REMARK 3.1. In the one dimensional case, we have

$$X_t = \frac{1}{\sqrt{6}}\sigma_t^2 |D^2_{xx}u_t|, \qquad \Lambda_t = \frac{1}{\sqrt{6}}|D^2_{xx}u_t|,$$

and the μ-optimal stopping times read

$$\tau_i^n = \inf\left\{t \geq \tau_{i-1}^n : |S_t - S_{\tau_{i-1}^n}| > \frac{\varepsilon_n}{\sqrt{|D^2_{xx}u_{\tau_{i-1}^n}|/\sqrt{6} + \mu\chi_\mu(|D^2_{xx}u_{\tau_{i-1}^n}|/\sqrt{6})}}\right\} \wedge T.$$

For $|D^2_{xx}u_t|$ bounded from below, we can take $\mu = 0$ and the optimal strategy coincides with that of [8], Theorem C.

The threshold $\mu \neq 0$ ensures that the hedging rebalancing occurs often enough, even if $\Lambda_t \neq 0$ for some time t: this interpretation is also valid in the multidimensional case.

3.2. *Proof of Theorem* 3.1. It is split into several steps.

Step 1: *Quadratic variation decomposition*. We start from the hedging error (3.1). A natural idea consists in writing a Taylor expansion (regarding the S

variable only) and showing that the residual terms converge to 0 fast enough as we could expect,

$$Z_s^n = \int_0^s \big(D_{xx}^2 u_{\varphi(t)} \Delta S_t\big) \cdot \mathrm{d}S_t + R_s^n, \tag{3.6}$$

where

$$R_s^n := \int_0^s \big(D_x u_t - D_x u_{\varphi(t)} - D_{xx}^2 u_{\varphi(t)} \Delta S_t\big) \cdot \mathrm{d}S_t, \qquad s \le T. \tag{3.7}$$

Then passing to quadratic variation, we obtain

$$\langle Z^n \rangle_T = \int_0^T \Delta S_t^* D_{xx}^2 u_{\varphi(t)} \, \mathrm{d}\langle S \rangle_t D_{xx}^2 u_{\varphi(t)} \Delta S_t + e_{1,T}^n,$$

where

$$e_{1,T}^n := \langle R^n \rangle_T + 2\left\langle \int_0^{\cdot} \big(D_{xx}^2 u_{\varphi(t)} \Delta S_t\big) \cdot \mathrm{d}S_t, R_\cdot^n \right\rangle_T. \tag{3.8}$$

Now, we wish an expression involving only the Brownian motion for ease of mathematical analysis: hence we replace ΔS_t by $\sigma_{\varphi(t)} \Delta B_t$ and $\mathrm{d}\langle S \rangle_t$ by $\sigma_{\varphi(t)} \sigma_{\varphi(t)}^* \, \mathrm{d}t$, leading to

$$\begin{aligned} \langle Z^n \rangle_T &= \int_0^T \Delta B_t^* \big(\sigma_{\varphi(t)}^* D_{xx}^2 u_{\varphi(t)} \sigma_{\varphi(t)}\big)^2 \Delta B_t \, \mathrm{d}t + e_{1,T}^n + e_{2,T}^n, \\ e_{2,T}^n &:= \int_0^T \Delta S_t^* D_{xx}^2 u_{\varphi(t)} \Delta\big(\sigma_t \sigma_t^*\big) D_{xx}^2 u_{\varphi(t)} \Delta S_t \, \mathrm{d}t \\ &\quad + \int_0^T (\Delta S_t + \sigma_{\varphi(t)} \Delta B_t)^* \\ &\qquad \times D_{xx}^2 u_{\varphi(t)} \sigma_{\varphi(t)} \sigma_{\varphi(t)}^* D_{xx}^2 u_{\varphi(t)} \left(\int_{\varphi(t)}^t \Delta \sigma_r \, \mathrm{d}B_r \right) \mathrm{d}t. \end{aligned} \tag{3.9}$$

As mentioned before, we seek a *smart* representation of the main term of $\langle Z^n \rangle_T$ in the form $\sum_{\tau_{i-1}^n < T} (\Delta B_{\tau_i^n}^* X_{\tau_{i-1}^n} \Delta B_{\tau_i^n})^2$ plus a stochastic integral, where X is a measurable adapted $d \times d$-matrix process which has to be defined. Instead of directly giving the solution, let us discuss a bit on the expected properties of X. Applying Itô's formula on each interval $[\tau_{i-1}^n, \tau_i^n]$, we obtain

$$\begin{aligned} &\sum_{\tau_{i-1}^n < T} \big(\Delta B_{\tau_i^n}^* X_{\tau_{i-1}^n} \Delta B_{\tau_i^n}\big)^2 \\ &\quad = \int_0^T \Delta B_t^* \big(2 \operatorname{Tr}(X_{\varphi(t)}) X_{\varphi(t)} + \big(X_{\varphi(t)} + X_{\varphi(t)}^*\big)^2\big) \Delta B_t \, \mathrm{d}t \\ &\qquad + 2 \int_0^T \Delta B_t^* X_{\varphi(t)} \Delta B_t \Delta B_t^* \big(X_{\varphi(t)} + X_{\varphi(t)}^*\big) \, \mathrm{d}B_t, \end{aligned}$$

with the tentative identification

$$2\,\mathrm{Tr}(X_{\varphi(t)})X_{\varphi(t)} + \big(X_{\varphi(t)} + X^*_{\varphi(t)}\big)^2 = \big(\sigma^*_{\varphi(t)} D^2_{xx} u_{\varphi(t)} \sigma_{\varphi(t)}\big)^2. \tag{3.10}$$

Mainly, two reasons prompt us to impose $X_{\varphi(t)} \in \mathcal{S}^d_+(\mathbb{R})$.

- Gathering the previous identities and anticipating a little bit on the following, the main contribution in $N^n_T \langle Z^n\rangle_T$ is

$$N^n_T \sum_{\tau^n_{i-1}<T} \big(\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big)^2 \ge \bigg(\sum_{\tau^n_{i-1}<T} \big|\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big|\bigg)^2$$

  using the Cauchy–Schwarz inequality. In general the limit of the above lower bound is not easy to handle because of the absolute values, but if the matrix $X_{\varphi(t)}$ is nonnegative-definite, we can remove them and conclude using a convergence result about discrete quadratic variations (Proposition 2.3).
- Once that we have restricted to nonnegative-definite matrices, let us prove that the solution to (3.10) (whenever it exists) is symmetric. If $\mathrm{Tr}(X_{\varphi(t)}) = 0$, then $X_{\varphi(t)} = 0$ (thus symmetric): indeed, $X_{\varphi(t)} + X^*_{\varphi(t)}$ is symmetric nonnegative-definite and has a null trace, thus it is the zero-matrix and consequently $X_{\varphi(t)} = -X^*_{\varphi(t)} = 0$ (since both $X_{\varphi(t)}$ and $X^*_{\varphi(t)}$ are nonnegative-definite). If $\mathrm{Tr}(X_{\varphi(t)}) > 0$, then taking the transposition of (3.10) readily gives $X_{\varphi(t)} = X^*_{\varphi(t)}$.

From Lemma 3.1, there exists exactly one adapted process X with values in $\mathcal{S}^d_+(\mathbb{R})$, solution of the equation $2\,\mathrm{Tr}(X)X + 4X^2 = (\sigma^* D^2_{xx} u\sigma)^2$. In addition, this solution is continuous a.s. because $C := \sigma^* D^2_{xx} u\sigma$ is continuous a.s., and the solution X is continuous as a function of C on $\mathcal{S}^d$. Gathering the previous identities, we have established a nice decomposition of the quadratic variation of the hedging error

$$\langle Z^n\rangle_T = \sum_{\tau^n_{i-1}<T} \big(\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big)^2 + e^n_{1,T} + e^n_{2,T} + e^n_{3,T}, \tag{3.11}$$

$$e^n_{3,T} := -4 \int_0^T \Delta B^*_t X_{\varphi(t)} \Delta B_t \Delta B^*_t X_{\varphi(t)}\, \mathrm{d}B_t. \tag{3.12}$$

Step 2: *Lower bound for the renormalized quadratic variation*. The Cauchy–Schwarz inequality yields that $N^n_T \sum_{\tau^n_{i-1}<T} (\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i})^2$ is bounded from below by

$$\begin{aligned}\bigg(\sum_{\tau^n_{i-1}<T} \big|\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big|\bigg)^2 &= \bigg(\sum_{\tau^n_{i-1}<T} \Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\bigg)^2 \\ &\xrightarrow{a.s.} \bigg(\int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t\bigg)^2,\end{aligned}$$

using that X is a nonnegative-definite matrix process and applying Proposition 2.3.

Step 3: *The renormalized errors* $\varepsilon_n^{-2\rho_N} e^n_{1,T}$, $\varepsilon_n^{-2\rho_N} e^n_{2,T}$ *and* $\varepsilon_n^{-2\rho_N} e^n_{3,T}$ *converge to 0* a.s. Observe that once these convergences are established, in view of (3.11) and $(\mathbf{A}_N)$ we easily complete the proof of Theorem 3.1.

Proof of $\varepsilon_n^{-2\rho_N} e^n_{1,T} \xrightarrow{a.s.} 0$. We first state an intermediate result which is proved in Appendix (Section A.5).

LEMMA 3.2. *Assume hypotheses* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_u)$, $(\mathbf{A}_S)$, $(\mathbf{A}_N)$ *and* $(\mathbf{A}_Y)$ *are in force*. *Then* $\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T \xrightarrow{a.s.} 0$ *where* R^n *is defined in* (3.7).

Then, starting from (3.8), applying the Cauchy–Schwarz inequality to the cross-variation and using $(\mathbf{A}_\sigma)$–$(\mathbf{A}_u)$–$(\mathbf{A}_S)$, we derive

$$
\begin{aligned}
&\varepsilon_n^{-2\rho_N}|e^n_{1,T}|\\
&\quad\le \varepsilon_n^{-2\rho_N}\langle R^n\rangle_T\\
&\qquad + 2\biggl(\varepsilon_n^{-2}\int_0^T \Delta S_t^* D^2_{xx}u_{\varphi(t)}\,\mathrm{d}\langle S\rangle_t D^2_{xx}u_{\varphi(t)}\Delta S_t\biggr)^{1/2}\bigl(\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T\bigr)^{1/2}\\
&\quad\le \varepsilon_n^{2(\rho_N-1)}\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T + 2C_0\bigl(\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T\bigr)^{1/2} \xrightarrow{a.s.} 0.
\end{aligned}
$$

Proof of $\varepsilon_n^{-2\rho_N} e^n_{2,T} \xrightarrow{a.s.} 0$. We analyze separately the two contributions in (3.9).

(1) First, simple computations using $(\mathbf{A}_\sigma)$–$(\mathbf{A}_u)$–$(\mathbf{A}_S)$ and Corollary 2.2 directly give (for any given $\rho > 0$)

$$
\begin{aligned}
\varepsilon_n^{-2\rho_N}&\biggl|\int_0^T \Delta S_t^* D^2_{xx}u_{\varphi(t)}\Delta\bigl(\sigma_t\sigma_t^*\bigr)D^2_{xx}u_{\varphi(t)}\Delta S_t\,\mathrm{d}t\biggr|\\
&\le C_0\varepsilon_n^{-2\rho_N+2}\bigl(\varepsilon_n^{\theta_\sigma}+\varepsilon_n^{(\theta_\sigma/2)(2-\rho)}\bigr).
\end{aligned}
$$

Since $\rho_N < 1+\theta_\sigma/2$ and ρ can be taken arbitrarily small, we obtain that the above upper bound converges a.s. to 0.

(2) Second, we apply twice Corollary 2.3(ii), first taking $\theta = 0$ and second taking $\theta = \theta_\sigma$, so that we obtain, for any given $\rho > 0$, a.s. for any $n \ge 0$,

$$
\sup_{1\le i\le N_T^n}\ \sup_{\tau_{i-1}^n\le t\le\tau_i^n}\bigl|\Delta S_t+\sigma_{\varphi(t)}\Delta B_t\bigr| \le C_0\varepsilon_n^{1-\rho}, \tag{3.13}
$$

$$
\sup_{1\le i\le N_T^n}\ \sup_{\tau_{i-1}^n\le t\le\tau_i^n}\biggl|\int_{\varphi(t)}^t \Delta\sigma_r\,\mathrm{d}B_r\biggr| \le C_0\varepsilon_n^{1+\theta_\sigma-\rho}, \tag{3.14}
$$

$$
\begin{aligned}
\varepsilon_n^{-2\rho_N}&\biggl|\int_0^T (\Delta S_t+\sigma_{\varphi(t)}\Delta B_t)^*\\
&\qquad\times D^2_{xx}u_{\varphi(t)}\sigma_{\varphi(t)}\sigma^*_{\varphi(t)}D^2_{xx}u_{\varphi(t)}\biggl(\int_{\varphi(t)}^t\Delta\sigma_r\,\mathrm{d}B_r\biggr)\mathrm{d}t\biggr|\\
&\le C_0\varepsilon_n^{2+\theta_\sigma-2\rho_N-2\rho}.
\end{aligned}
$$

Owing to $\rho_N < 1 + \theta_\sigma/2$, taking ρ small enough implies the a.s. convergence of the latter upper bound to 0. As a result, $\varepsilon_n^{-2\rho_N} e_{2,T}^n \xrightarrow{a.s.} 0$.

Proof of $\varepsilon_n^{-2\rho_N} e_{3,T}^n \xrightarrow{a.s.} 0$. It is a direct consequence of the following lemma.

LEMMA 3.3. *Assume* $(\mathbf{A}_\sigma)$. *Let* $\mathcal{T} = (\mathcal{T}^n)_{n\geq 0}$ *be an admissible sequence of strategies, and 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. Then for any* $p > \frac{2}{3-2\rho_N}$, *the series* $\sum_{n\geq 0} |\varepsilon_n^{-2\rho_N} \int_0^T \Delta B_t^* H_{\varphi(t)} \Delta B_t \Delta B_t^* H_{\varphi(t)}\, \mathrm{d}B_t|^p$ *converges almost surely.*

PROOF. Set $\alpha_t^n := \Delta B_t^* H_{\varphi(t)} \Delta B_t \Delta B_t^* H_{\varphi(t)}$ and define the scalar continuous local martingale $M_t^n := \varepsilon_n^{-2\rho_N} \int_0^t \alpha_s^n \,\mathrm{d}B_s$. In view of Corollary 2.1, it is enough to check that $(\langle M^n\rangle_T^{p/2})_{n\geq 0}$ defines the terms of an a.s. convergent series. An application of Corollary 2.3(ii) with $\rho = \frac{(3-2\rho_N)p-2}{3p} > 0$ and $\theta = 0$ gives $\sup_{1\leq i\leq N_T^n} \sup_{\tau_{i-1}^n \leq t\leq \tau_i^n} |\Delta B_t| < C_0 \varepsilon_n^{1-\rho}$ and therefore

$$
\begin{aligned}
\langle M^n\rangle_T^{p/2} &= \varepsilon_n^{-2p\rho_N} \left(\int_0^T |\alpha_t^n|^2 \,\mathrm{d}t \right)^{p/2} \\
&\leq C_0 \varepsilon_n^{-2p\rho_N} \sup_{1\leq i\leq N_T^n} \sup_{\tau_{i-1}^n \leq t\leq \tau_i^n} |\Delta B_t|^{3p} \leq C_0 \varepsilon_n^2 \qquad \text{a.s.}
\end{aligned}
$$

We are finished. □

3.3. *Proof of Theorem* 3.2. We first check the admissibility of $\mathcal{T}_\mu$, by applying Proposition 2.4. Indeed, owing to $(\mathbf{A}_u)$ and $(\mathbf{A}_\sigma^{\text{Ellip.}})$, $(\Lambda_t)_{0\leq t<T}$ is a continuous adapted nonnegative-definite $d \times d$-matrix process with $\sup_{0\leq t<T} |\Lambda_t| < +\infty$ a.s. The same properties clearly hold for $(\Lambda_t^\mu)_{0\leq t<T}$. In addition, $\lambda_{\min}(\Lambda_t^\mu) \geq \mu/2 > 0$ and $\sup_{0\leq t<T} \lambda_{\max}(\Lambda_t^\mu) \leq \mu + \sup_{0\leq t<T} \lambda_{\max}(\Lambda_t) < +\infty$ a.s. Therefore, $\mathcal{T}_\mu$ is admissible and in addition $\sup_{n\geq 0} \varepsilon_n^2 N_T^n < +\infty$ a.s. Hence it allows us to re-use the computations of the proof of Theorem 3.1 in the case $\rho_N = 1$.

Now let us show the μ-optimality. Writing $N_T^n = 1 + \sum_{1\leq i\leq N_T^n -1} 1$, we point out

$$
\begin{aligned}
\varepsilon_n^2 N_T^n &= \varepsilon_n^2 + \sum_{1\leq i\leq N_T^n-1} \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n}^\mu \Delta S_{\tau_i^n} \\
&= \varepsilon_n^2 - \Delta S_T^* \Lambda_{\tau_{N_T^n-1}^n}^\mu \Delta S_T + \sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n}^\mu \Delta S_{\tau_i^n} \qquad (3.15) \\
&\xrightarrow{a.s.} \int_0^T \operatorname{Tr}(\Lambda_t^\mu \sigma_t \sigma_t^*)\,\mathrm{d}t
\end{aligned}
$$

using the convergence of Proposition 2.3. On the other hand, starting from the decomposition (3.11) of the hedging error quadratic variation, we write

$$
\begin{aligned}
\langle Z^n\rangle_T &= \sum_{1\le i\le N_T^n-1} \bigl(\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\bigr)^2 + e^n_{1,T} + e^n_{2,T} + e^n_{3,T} \\
&\qquad + e^n_{4,T} + e^n_{5,T} + e^n_{6,T},
\end{aligned}
$$

$$
\begin{aligned}
e^n_{4,T} &:= \sum_{\tau_{i-1}^n<T} \bigl(\Delta B^*_{\tau_i^n} X_{\tau_{i-1}^n}\Delta B_{\tau_i^n}\bigr)^2 - \bigl(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\bigr)^2, \\
e^n_{5,T} &:= \sum_{\tau_{i-1}^n<T} \bigl(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\bigr)^2 - \bigl(\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\bigr)^2, \\
e^n_{6,T} &:= \bigl(\Delta S^*_T \Lambda^\mu_{\tau^n_{N_T^n-1}}\Delta S_T\bigr)^2.
\end{aligned}
\tag{3.16}
$$

In view of the definition of the strategy $\mathcal{T}^n_\mu$, (3.16) becomes

$$
\varepsilon_n^{-2}\langle Z^n\rangle_T = \sum_{1\le i\le N_T^n-1} \Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n} + \varepsilon_n^{-2}\sum_{j=1}^6 e^n_{j,T}. \tag{3.17}
$$

Similarly to (3.15), we show that $\sum_{1\le i\le N_T^n-1}\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}(\Lambda^\mu_t\sigma_t\sigma_t^*)\,\mathrm{d}t$. Furthermore we have already established (see step 3 of proof of Theorem 3.1) that $\varepsilon_n^{-2}e^n_{j,T} \xrightarrow{a.s.} 0$ for $j=1,2,3$ (remind that we can take $\rho_N = 1$); the case $j=6$ is also fulfilled because $0\le e^n_{6,T}\le \varepsilon_n^4$.

To analyze $e^n_{4,T}$, set $D_{B,i} := \sigma_{\tau_{i-1}^n}\Delta B_{\tau_i^n}$ and $D_{S,i} := \Delta S_{\tau_i^n}$, write $X_{\tau_{i-1}^n} = \sigma^*_{\tau_{i-1}^n}\Lambda_{\tau_{i-1}^n}\sigma_{\tau_{i-1}^n}$ and

$$
\begin{aligned}
&\bigl(\Delta B^*_{\tau_i^n} X_{\tau_{i-1}^n}\Delta B_{\tau_i^n}\bigr)^2 - \bigl(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\bigr)^2 \\
&\quad = \bigl(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i}\bigr)^2 - \bigl(D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\bigr)^2 \\
&\quad = \bigl(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i} - D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\bigr)\bigl(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i} + D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\bigr) \\
&\quad = (D_{B,i}+D_{S,i})^*\Lambda_{\tau_{i-1}^n}(D_{B,i}-D_{S,i})\bigl(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i} + D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\bigr).
\end{aligned}
$$

Then we deduce that $\varepsilon_n^{-2}|e^n_{4,T}|$ is bounded by

$$
\begin{aligned}
&\varepsilon_n^{-2}N_T^n \sup_{1\le i\le N_T^n}\ \sup_{\tau_{i-1}^n\le t\le\tau_i^n]} |\Lambda_{\varphi(t)}|^2|\Delta S_t + \sigma_{\varphi(t)}\Delta B_t|\left|\int_{\varphi(t)}^t \Delta\sigma_s\,\mathrm{d}B_s\right| \\
&\qquad\qquad \times \bigl(|\Delta S_t|^2 + |\sigma_{\varphi(t)}\Delta B_t|^2\bigr) \\
&\quad \le C_0\varepsilon_n^{-2}\varepsilon_n^{-2}\varepsilon_n^{1-\rho}\varepsilon_n^{(1+\theta_\sigma-\rho)}\varepsilon_n^{2(1-\rho)} = C_0\varepsilon_n^{\theta_\sigma/5} \xrightarrow{a.s.} 0,
\end{aligned}
$$

where we have used ($\mathbf{A}_N$) (with $\rho_N = 1$), and estimates (3.13)–(3.14) with $\rho = \theta_\sigma/5$ (which are available for any sequence of admissible strategies). This proves $\varepsilon_n^{-2} e_{4,T}^n \xrightarrow{a.s.} 0$.

Finally regarding $e_{5,T}^n$, recalling that the matrix $\Lambda_{\tau_{i-1}^n}$ is nonnegative-definite, we obtain that $|\varepsilon_n^{-2} e_{5,T}^n|$ is bounded by

$$\begin{aligned}
\varepsilon_n^{-2} & \sum_{\tau_{i-1}^n < T} \big|\Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} - \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n}^\mu \Delta S_{\tau_i^n}\big| \\
& \qquad \times \big(\Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} + \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n}^\mu \Delta S_{\tau_i^n}\big) \\
& \le \sum_{\tau_{i-1}^n < T} \mu \chi_\mu\big(\lambda_{\min}(\Lambda_{\tau_{i-1}^n})\big) |\Delta S_{\tau_i^n}|^2 \big[2\varepsilon_n^{-2} \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n}^\mu \Delta S_{\tau_i^n}\big] \\
& \le 2\mu \sum_{\tau_{i-1}^n < T} \chi_\mu\big(\lambda_{\min}(\Lambda_{\tau_{i-1}^n})\big) |\Delta S_{\tau_i^n}|^2,
\end{aligned}$$

where we have used the definition of $\mathcal{T}_\mu$ at the last inequality. Thus Proposition 2.3 yields

$$\limsup_{n\to+\infty} \big|\varepsilon_n^{-2} e_{5,T}^n\big| \le 2\mu \int_0^T \chi_\mu\big(\lambda_{\min}(\Lambda_t)\big) \operatorname{Tr}\big(\sigma_t \sigma_t^*\big) \,\mathrm{d}t \qquad \text{a.s.}$$

Let us summarize: setting $L_T := \int_0^T \operatorname{Tr}(\Lambda_t \sigma_t \sigma_t^*) \,\mathrm{d}t = \int_0^T \operatorname{Tr}(X_t) \,\mathrm{d}t$ and $L_T^\mu := \int_0^T \chi_\mu(\lambda_{\min}(\Lambda_t)) \operatorname{Tr}(\sigma_t \sigma_t^*) \,\mathrm{d}t$ so that $\int_0^T \operatorname{Tr}(\Lambda_t^\mu \sigma_t \sigma_t^*) \,\mathrm{d}t = L_T + \mu L_T^\mu$, we have shown

$$\begin{aligned}
& \varepsilon_n^2 N_T^n \xrightarrow{a.s.} L_T + \mu L_T^\mu, \qquad \limsup_{n\to+\infty} \big|\varepsilon_n^{-2} \langle Z^n \rangle_T - \big(L_T + \mu L_T^\mu\big)\big| \le 2\mu L_T^\mu \qquad \text{a.s.}, \\
& \limsup_{n\to+\infty} \big|N_T^n \langle Z^n \rangle_T - (L_T)^2\big| \\
& \quad \le \limsup_{n\to+\infty} \big|\varepsilon_n^{-2} \langle Z^n \rangle_T - L_T\big| \limsup_{n\to+\infty} \varepsilon_n^2 N_T^n + L_T \limsup_{n\to+\infty} \big|\varepsilon_n^2 N_T^n - L_T\big| \\
& \quad \le 3\mu L_T^\mu \big(L_T + \mu L_T^\mu\big) + L_T \mu L_T^\mu = \mu L_T^\mu \big(4 L_T + 3\mu L_T^\mu\big) \qquad \text{a.s.}
\end{aligned}$$

Theorem 3.2 is proved.

3.4. *Proof of Theorem* 3.3. Here, arguments are simpler in all steps of the proof of Section 3.3, so we shall skip details; the admissibility of the strategy comes readily from the ad hoc assumption (3.5) and Proposition 2.4; the optimality follows as before from

$$\varepsilon_n^2 N_T^n = \varepsilon_n^2 + \sum_{1 \le i \le N_T^n - 1} \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \operatorname{Tr}(X_t) \,\mathrm{d}t,$$

and from [setting $\bar{e}^n_{6,T} := (\Delta S^*_T \Lambda_{\tau^n_{N^n_T-1}} \Delta S_T)^2$]

$$\begin{aligned}\varepsilon_n^{-2}\langle Z^n\rangle_T &= \varepsilon_n^{-2} \sum_{1\le i\le N^n_T-1} (\Delta S^*_{\tau^n_i} \Lambda_{\tau^n_{i-1}} \Delta S_{\tau^n_i})^2 + \varepsilon_n^{-2}\sum_{j=1}^4 e^n_{j,T} + \varepsilon_n^{-2}\bar{e}^n_{6,T}\\ &\xrightarrow{a.s.} \int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t\end{aligned}$$

with the help of the convergence results already obtained. Theorem 3.3 is proved.

4. Numerical experiments.

4.1. *Algorithm for the optimal stopping times.* From the previous section (Theorem 3.2), the μ-optimal stopping times ($\mu > 0$) are iteratively given by $\tau^n_0 := 0$ and

$$\tau^n_i := \inf\{t \ge \tau^n_{i-1} : (S_t - S_{\tau^n_{i-1}})^* \Lambda^\mu_{\tau^n_{i-1}} (S_t - S_{\tau^n_{i-1}}) \ge \varepsilon_n^2\} \wedge T,$$

where for any t, $\Lambda^\mu_t := \Lambda_t + \mu\chi_\mu(\lambda_{\min}(\Lambda_t)) I_d$, $\Lambda_t := (\sigma_t^{-1})^* X_t \sigma_t^{-1}$ and X_t solves (3.3) with $c_t = \sigma^*_t D^2_{xx} u_t \sigma_t$. Thus, τ^n_i is the first hitting time of an ellipsoid centered at $S_{\tau^n_{i-1}}$ with principal axes equal to the orthogonal eigenvectors of the symmetric positive-definite matrix $\Lambda^\mu_{\tau^n_{i-1}}$ (or equivalently those of $\Lambda_{\tau^n_{i-1}}$). We briefly recall (see Section A.4) the main steps to compute the matrix $X_{\tau^n_{i-1}}$ ($i \ge 1$) from which we derive $\Lambda_{\tau^n_{i-1}}$ and $\Lambda^\mu_{\tau^n_{i-1}}$:

(1) Diagonalize the symmetric matrix $c_{\tau^n_{i-1}} = \sigma^*_{\tau^n_{i-1}} D^2_{xx} u_{\tau^n_{i-1}} \sigma_{\tau^n_{i-1}} := P_{\tau^n_{i-1}} \times \mathrm{Diag}(\lambda_j(c_{\tau^n_{i-1}}) : 1 \le j \le d) P^*_{\tau^n_{i-1}}$, where $P_{\tau^n_{i-1}}$ is an orthogonal matrix.

(2) Find the zero $y_{\tau^n_{i-1}} \in \mathbb{R}^+$ of the increasing function $y \mapsto (4+d)y - \sum_{j=1}^d \sqrt{y^2 + 4\lambda_j^2(c_{\tau^n_{i-1}})}$. This root lies in the interval $[0, d|\lambda(c_{\tau^n_{i-1}})|/\sqrt{4+2d}]$; see the proof of Lemma 3.1.

(3) From (A.7), we obtain

$$X_{\tau^n_{i-1}} = P_{\tau^n_{i-1}} \mathrm{Diag}\biggl(\frac{-y_{\tau^n_{i-1}} + \sqrt{y^2_{\tau^n_{i-1}} + 4\lambda_j^2(c_{\tau^n_{i-1}})}}{4} : 1 \le j \le d\biggr) P^*_{\tau^n_{i-1}}.$$

Last, we mention that even if $\Lambda^\mu_{\tau^n_{i-1}}$ is tractable, the exact simulation of τ^n_i is in generally impossible, and approximations are required; see [17] and references therein.

4.2. *Numerical tests.* This section is dedicated to an application of Theorem 3.2 to the case of an exchange binary option $g(S_T) = \mathbf{1}_{S^1_T \ge S^2_T}$. This example is relevant in our study (and improves the setting of [8]) because this is a simple *bi-dimensional nonconvex* function, for which the value function u and its sensitivities are available in the Black–Scholes model

$$\mathrm{d}\begin{pmatrix} S_t^1 \\ S_t^2 \end{pmatrix} = \begin{pmatrix} \sigma_1 S_t^1 & 0 \\ \rho\sigma_2 S_t^2 & \sqrt{1-\rho^2}\sigma_2 S_t^2 \end{pmatrix} \mathrm{d}\begin{pmatrix} B_t^1 \\ B_t^2 \end{pmatrix},$$

where (B^1, B^2) are two independent Brownian motions. The model parameters are set to $S_0^1 = 100$, $S_0^2 = 100$, $\sigma_1 = 0.3$, $\sigma_2 = 0.4$, $\rho = 0.5$ and $T = 1$.

We take $\varepsilon_n = 0.05$. In our different tests, we have not observed a significant difference by taking $\mu = 0$ or μ small; hence, we only report the values for $\mu = 0$. We generate 1000 experiments ω, independently. To compute the hitting times for each ω, we use a thin uniform time mesh $\pi_{\bar{n}} = (iT/\bar{n})_{0\le i\le \bar{n}}$ ($\bar{n} = 50{,}000$ in our tests): we draw $S^1(\omega)$ and $S^2(\omega)$ along $\pi_{\bar{n}}$ and compute (with the help of the previous algorithm) the hitting times $\tau_i^n(\omega) = \inf\{t \in \pi_{\bar{n}} \cap]\tau_{i-1}^n(\omega), T] : [(S_t - S_{\tau_{i-1}^n})^* \Lambda_{\tau_{i-1}^n}^{\mu} (S_t - S_{\tau_{i-1}^n})](\omega) \ge \varepsilon_n^2\} \wedge T$; at the end of the process, we get the number $N_T^n(\omega)$ of discrete times. The mesh $\pi_{\bar{n}}$ is also used to compute subsequent quadratic variations and time integrals.

We compare ω by ω the above strategy with that based on the *uniform* mesh $\pi_{N_T^n(\omega)}$ and with that based on the so-called *fractional* mesh[4] $(T[1 - (1 - i/N_T^n(\omega))^2])_{1\le i\le N_T^n(\omega)}$: this comparison looks quite fair from a practitioner point of view since he is allowed to rebalance the hedging portfolio the same number of times. The use of the optimal stochastic grid is slightly more demanding since it requires the computations of more Greeks than only the Delta (because of the matrix Λ^{μ}); however, these sensitivities are widely available in any trading system, which makes this higher complexity likely negligible in view of the benefit of optimal times.

We define $\beta_{\mathrm{stochastic}}(\omega)$, $\beta_{\mathrm{uniform}}(\omega)$, $\beta_{\mathrm{fractional}}(\omega)$ where we compute $\beta_{\cdot}(\omega) := \frac{N_T^n \langle Z^n\rangle_T}{(\int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t)^2}(\omega)$ according to each of these three strategies: in view of Theorem 3.2, this ratio is asymptotically greater than 1 and adimensional; moreover, the closer to 1 the ratio, the better the strategy.

Results. Figure 1 displays, for each ω, the couples

$$\bigl(\beta_{\mathrm{stochastic}}(\omega), \beta_{\mathrm{uniform}}(\omega)\bigr) \quad \text{and} \quad \bigl(\beta_{\mathrm{stochastic}}(\omega), \beta_{\mathrm{fractional}}(\omega)\bigr).$$

Most of the times, the points are above the diagonal, showing that the μ-optimal strategy lessens the quadratic variation ω-wise (remind that the strategies have got the same number of discrete times N_T^n), compared to the quadratic variation worked out over the deterministic time mesh. In addition, $\beta_{\mathrm{stochastic}}$ is concen-

[4]According to [12], the fractional smoothness of $g(S_T)$ is $\frac{1}{2}$; thus, when $N_T^n(\omega)$ is deterministic, this choice of fractional mesh yields that $\mathbb{E}(\langle Z^n\rangle_T)$ is of order 1 w.r.t. the inverse of the number of times, instead of order $\frac{1}{2}$ with the uniform mesh.

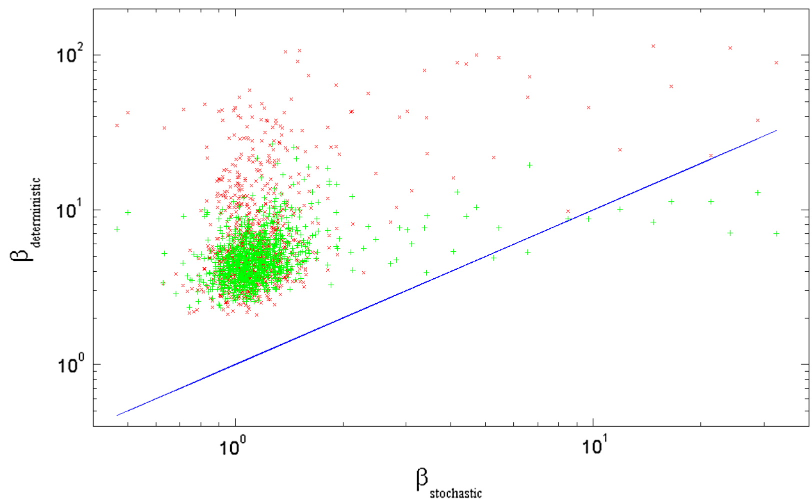

FIG. 1. *“×,” “+” and the blue line correspond respectively to “$(\beta_{\text{stochastic}}, \beta_{\text{uniform}})$,” “$(\beta_{\text{stochastic}}, \beta_{\text{fractional}})$” and the identity function.*

trated around 1, which means a convergence of $N_T^n \langle Z^n \rangle_T$ toward the lower bound $(\int_0^T \mathrm{Tr}(X_t)\,\mathrm{d}t)^2$.

Figure 2 displays $\langle Z^n \rangle_T$ as a function of N_T^n for the three strategies and for different ω: here again, we observe that the μ-optimal strategy outperforms deterministic strategies.

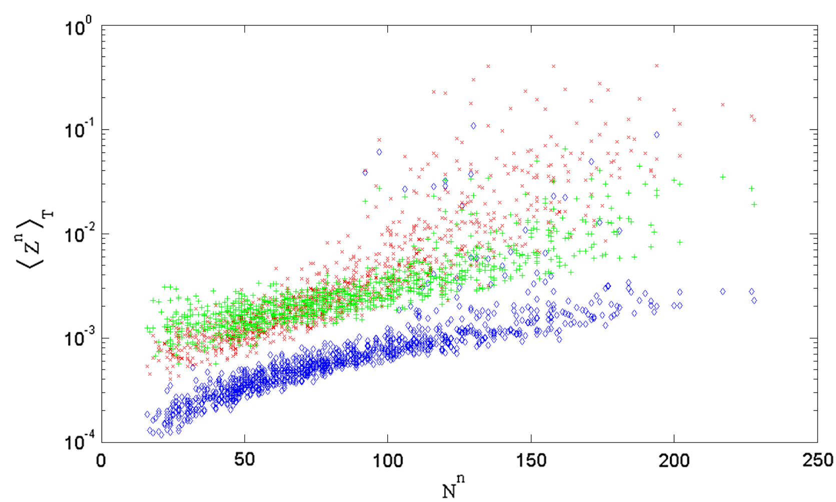

FIG. 2. *“×,” “+” and correspond respectively to “$\langle Z^n \rangle_{T,\text{uniform}}$,” “$\langle Z^n \rangle_{T,\text{fractional}}$” and “$\langle Z^n \rangle_{T,\text{stochastic}}$.”*

APPENDIX

A.1. Proof of Proposition 2.1. Let us prove (i), assuming only $(\mathbf{A}_S)$. For $p=0$, this is trivial.

Now consider the case $p>0$. Since σ_t is nonzero for any t and continuous, $C_E := \inf_{t\in[0,T]}(\sum_{j=1}^d e_j\cdot\sigma_t\sigma_t^* e_j)>0$ a.s., where e_j is the jth element of the canonical basis in $\mathbb{R}^d$. Therefore, a.s. for any $0\le s\le t\le T$ we have

$$\begin{aligned}0\le t-s &\le C_E^{-1}\int_s^t\left(\sum_{j=1}^d e_j\cdot\sigma_r\sigma_r^* e_j\right)\mathrm{d}r = C_E^{-1}\sum_{j=1}^d[\langle S^j\rangle_t-\langle S^j\rangle_s]\\ &= C_E^{-1}\sum_{j=1}^d\left[(S_t^j-S_s^j)^2-2\int_s^t(S_r^j-S_s^j)\,\mathrm{d}S_r^j\right],\end{aligned}\tag{A.1}$$

applying the Itô formula at the last equality. Take $s=\tau_{i-1}^n$, $t=\tau_i^n$ and use $(\mathbf{A}_S)$

$$\begin{aligned}\Delta\tau_i^n &\le C_E^{-1}\left(C_0\varepsilon_n^2+2\sum_{j=1}^d\left|\int_{\tau_{i-1}^n}^{\tau_i^n}\Delta S_r^j\,\mathrm{d}S_r^j\right|\right)\\ &\le C_E^{-1}\left(C_0\varepsilon_n^2+4\sum_{j=1}^d\sup_{0\le t\le T}\left|\int_0^t\Delta S_r^j\,\mathrm{d}S_r^j\right|\right).\end{aligned}\tag{A.2}$$

Now for $j=1,\dots,d$, set $M_t^{j,n}:=\varepsilon_n^{2/p-1}\int_0^t\Delta S_r^j\,\mathrm{d}S_r^j$ (recalling that $p>0$). Then

$$\sum_{n\ge0}\langle M^{j,n}\rangle_T^{p/2}=\sum_{n\ge0}\varepsilon_n^{2-p}\left(\int_0^T|\Delta S_t^j|^2\,\mathrm{d}\langle S^j\rangle_t\right)^{p/2}\le C_0\sum_{n\ge0}\varepsilon_n^2<+\infty\qquad\text{a.s.}$$

Thus owing to Corollary 2.1 the terms $(\sup_{0\le t\le T}|M_t^{j,n}|^p)_{n\ge0}$ define an a.s. convergent series. Combining this with (A.2), we finally derive

$$\begin{aligned}&\sum_{n\ge0}\left[\varepsilon_n^{2/p-1}\sup_{1\le i\le N_T^n}|\Delta\tau_i^n|\right]^p\\ &\qquad\le C_0\left(\sum_{n\ge0}[\varepsilon_n^{2/p-1}\varepsilon_n^2]^p+\sum_{j=1}^d\sum_{n\ge0}\sup_{0\le t\le T}|M_t^{j,n}|^p\right)<+\infty\qquad\text{a.s.}\end{aligned}$$

It remains to justify (ii). For $p=0$, the result directly follows from $(\mathbf{A}_N)$ and the inequality (2.1). Now take $p>0$, and set

$$\begin{aligned}U_t^n &:= \varepsilon_n^{-2(p-1)+2\rho_N}\sum_{\tau_{i-1}^n<t}\left|\sum_{j=1}^d\Delta\langle S^j\rangle_{\tau_i^n\wedge t}\right|^p,\\ V_t^n &:= \varepsilon_n^{-2(p-1)+2\rho_N}\sum_{\tau_{i-1}^n<t}\sup_{s\in(\tau_{i-1}^n,\tau_i^n\wedge t]}|\Delta S_s|^{2p}.\end{aligned}$$

If $\sum_{n\geq 0} U_T^n \xrightarrow{a.s.}$, (A.1) immediately yields that $\sum_{n\geq 0}\varepsilon_n^{-2(p-1)+2\rho_N} \times \sum_{\tau_{i-1}^n<T}(\Delta\tau_i^n)^p \xrightarrow{a.s.}$. Thus, it is sufficient to show $\sum_{n\geq 0} U_t^n \xrightarrow{a.s.}$, for any $t\in[0,T]$, and this is achieved by an application of Lemma 2.2. The sequences of processes $(U^n)_{n\geq 0}$ and $(V^n)_{n\geq 0}$ are in $\mathcal{C}_0^+$. Then V^n is nondecreasing, and using $(\mathbf{A}_S)$–$(\mathbf{A}_N)$

$$\sum_{n\geq 0} V_T^n \leq C_0 \sum_{n\geq 0}\varepsilon_n^{-2(p-1)+2\rho_N}\varepsilon_n^{2p} N_T^n \leq C_0\sum_{n\geq 0}\varepsilon_n^2 < +\infty \qquad \text{a.s.}$$

Then we deduce that items (i′) and (ii′) of Lemma 2.2 are fulfilled. It remains to check the relation of domination [item (iii′)]. Let $k\in\mathbb{N}$. On the set $\{\tau_{i-1}^n < t\wedge\theta_k\}$, from the multidimensional BDG inequality in a conditional version, we have

$$\mathbb{E}\left(\left|\sum_{j=1}^d \Delta\langle S^j\rangle_{\tau_i^n\wedge t\wedge\theta_k}\right|^p \Big|\mathcal{F}_{\tau_{i-1}^n}\right) \leq c_p\mathbb{E}\Big(\sup_{\tau_{i-1}^n<s\leq\tau_i^n\wedge t\wedge\theta_k}|\Delta S_s|^{2p}\big|\mathcal{F}_{\tau_{i-1}^n}\Big). \tag{A.3}$$

Then, it follows

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

The proof is complete.

A.2. Proof of Proposition 2.2. Let $p>0$. Let δ be the parameter standing for $\frac{1}{2}$ under $(\mathbf{A}_S)$ and 1 under $(\mathbf{A}_S)$–$(\mathbf{A}_N)$. Set

$$U_t^n := \varepsilon_n^{-2\delta((p(\theta+1)/2)-2(1-\delta))+2+2\rho_N(2\delta-1)}\sum_{\tau_{i-1}^n<t}\sup_{\tau_{i-1}^n\leq s\leq\tau_i^n\wedge t}|\Delta M_t^n|^p,$$

$$V_t^n := \varepsilon_n^{-2\delta((p(\theta+1)/2)-2(1-\delta))+2+2\rho_N(2\delta-1)}\sum_{\tau_{i-1}^n<t}\left|\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\alpha_r^n\,\mathrm{d}r\right|^{p/2}.$$

Observe that the announced result reads as $\sum_{n\geq 0} U_T^n \xrightarrow{a.s.}$. To prove this convergence, it is enough to establish that $\sum_{n\geq 0} V_T^n \xrightarrow{a.s.}$. Indeed, following the arguments of the proof of Proposition 2.1(ii), we can apply Lemma 2.2 since $(U^n)_{n\geq 0}$ and $(V^n)_{n\geq 0}$ are two sequences of continuous adapted processes and:

(i′) V^n is nondecreasing on $[0,T]$ a.s.;
(iii′) the domination is satisfied thanks to the BDG inequalities, similarly to (A.3).

Now to prove (ii′), that is, $\sum_{n\geq 0} V_T^n \xrightarrow{a.s.}$, write

$$\sum_{n\geq 0} V_T^n \leq \sum_{n\geq 0} \varepsilon_n^{-2\delta((p(\theta+1)/2)-2(1-\delta))+2+2\rho_N(2\delta-1)} \times \sum_{\tau_{i-1}^n<T} \left|C_0\left(\varepsilon_n^{2\theta}+\left(\Delta\tau_i^n\right)^\theta\right)\Delta\tau_i^n\right|^{p/2} \qquad \text{a.s.}$$

First, consider the case $(\mathbf{A}_S)$ and set $D_n^{(q)} := \sup_{1\leq i\leq N_T^n}(\Delta\tau_i^n)^q$ for $q\geq 0$: Proposition 2.1(i) yields $\mathcal{D}^{(q)} := \sum_{n\geq 0}\varepsilon_n^{-(q-2)}D_n^{(q)} < +\infty$ a.s. Using $p\geq 2$, it readily follows that

$$\begin{aligned}
\sum_{n\geq 0} V_T^n &\leq \sum_{n\geq 0}\varepsilon_n^{-(p(\theta+1)/2-3)}C_0^{p/2}\sum_{\tau_{i-1}^n<T}\left(\varepsilon_n^{2\theta}+\left(\Delta\tau_i^n\right)^\theta\right)^{p/2}\left(\Delta\tau_i^n\right)^{p/2-1}\Delta\tau_i^n \\
&\leq \sum_{n\geq 0}\varepsilon_n^{-(p(\theta+1)/2-3)}C_0^{p/2}2^{p/2-1}T\left(\varepsilon_n^{p\theta}D_n^{(p/2-1)}+D_n^{((\theta+1)p/2-1)}\right) \\
&\leq C_0^{p/2}2^{p/2-1}T\left(\left(\sup_{n\geq 0}\varepsilon_n\right)^{p\theta/2}\mathcal{D}^{(p/2-1)}+\mathcal{D}^{((\theta+1)p/2-1)}\right) < +\infty \qquad \text{a.s.}
\end{aligned}$$

Second for the case $(\mathbf{A}_S)$–$(\mathbf{A}_N)$, setting $D_n^{(q)} := \sum_{\tau_{i-1}^n<T}(\Delta\tau_i^n)^q$ for $q\geq 0$, we have $\mathcal{D}^{(q)} := \sum_{n\geq 0}\varepsilon_n^{-2(q-1)+2\rho_N}D_n^{(q)} < +\infty$ a.s., thanks to Proposition 2.1(ii). Then we easily deduce (for any $p>0$)

$$\begin{aligned}
\sum_{n\geq 0} V_T^n &\leq C_0^{p/2}2^{(p/2-1)_+}\sum_{n\geq 0}\varepsilon_n^{-2(p(\theta+1)/2-1)+2\rho_N} \times \sum_{\tau_{i-1}^n<T}\left(\varepsilon_n^{p\theta}\left(\Delta\tau_i^n\right)^{p/2}+\left(\Delta\tau_i^n\right)^{(\theta+1)p/2}\right) \\
&= C_0^{p/2}2^{(p/2-1)_+}\left(\mathcal{D}^{(p/2)}+\mathcal{D}^{((\theta+1)p/2)}\right) < +\infty \qquad \text{a.s.}
\end{aligned}$$

A.3. Proof of Proposition 2.4. It is standard to check that $\mathcal{T}^n$ is a sequence of increasing stopping times; we skip details. Let us justify that the size of $\mathcal{T}^n$ is a.s. finite, for any $n\geq 0$. For a given $n\geq 0$, define the event $\mathcal{N}^n := \{N_T^n = +\infty\}$. For $\omega\in\mathcal{N}^n$, the infinite sequence $(\tau_i^n(\omega))_{i\geq 0}$ converges, because increasing and bounded by T. Thus, on $\mathcal{N}^n\cap E_S$ with $E_S = \{(S_t)_{0\leq t\leq T}$ continuous and $\sup_{0\leq t<T}\lambda_{\max}(H_t) < +\infty\}$, we have

$$\begin{aligned}
0 < \varepsilon_n &= (S_{\tau_i^n}-S_{\tau_{i-1}^n})^* H_{\tau_{i-1}^n}(S_{\tau_i^n}-S_{\tau_{i-1}^n}) \\
&\leq \sup_{0\leq t<T}\lambda_{\max}(H_t)|S_{\tau_i^n}-S_{\tau_{i-1}^n}|^2 \underset{i\to+\infty}{\to} 0,
\end{aligned}$$

which is impossible. Thus, $\mathcal{N}^n\subset E_S^c$ and $\mathbb{P}(\mathcal{N}^n)=0$ since S is a.s. continuous and $\sup_{0\leq t<T}\lambda_{\max}(H_t)$ is a.s. finite.

Besides, we have $C_H := \inf_{0\le t<T} \lambda_{\min}(H_t) > 0$ a.s., and we immediately get

$$\varepsilon_n^{-2} \sup_{1\le i\le N_T^n} \sup_{t\in(\tau_{i-1}^n,\tau_i^n]} |\Delta S_t|^2 \le C_H^{-1}\varepsilon_n^{-2} \sup_{1\le i\le N_T^n} \sup_{t\in(\tau_{i-1}^n,\tau_i^n]} \big(\Delta S_t^* H_{\tau_{i-1}^n} \Delta S_t\big) \le C_H^{-1},$$

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

Then, writing $N_T^n = 1 + \sum_{1\le i\le N_T^n-1} 1$, we point out (for n large enough so that $\varepsilon_n \le 1$)

$$\begin{aligned}\varepsilon_n^{2\rho_N} N_T^n &\le \varepsilon_n^2 N_T^n\\ &\le \varepsilon_n^2 + \sum_{1\le i\le N_T^n-1} \Delta S_{\tau_i^n}^* H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \le \varepsilon_n^2 + \sum_{\tau_{i-1}^n<T} \Delta S_{\tau_i^n}^* H_{\tau_{i-1}^n} \Delta S_{\tau_i^n};\end{aligned}$$

using moreover from Proposition 2.3, we know that under the assumption $(\mathbf{A}_S)$ only,

$$\sum_{\tau_{i-1}^n<T} \Delta S_{\tau_i^n}^* H_{\tau_{i-1}^n} \Delta S_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}\big(H_t \,\mathrm{d}\langle S\rangle_t\big) < +\infty.$$

This validates the assumption $(\mathbf{A}_N)$.

REMARK A.1. The structure of hitting times of ellipsoids with size ε_n has a specific feature compared to general admissible strategies: the assumption $(\mathbf{A}_S)$ entails the assumption $(\mathbf{A}_N)$.

A.4. Proof of Lemma 3.1. We split the proof into several steps. Let

$$h: \begin{cases} \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R},\\ (\lambda, y) \mapsto (4+d)y - \displaystyle\sum_{i=1}^d \sqrt{y^2 + 4\lambda_i^2}.\end{cases}$$

Assume for a while that:

($\star$) (a) for any $\lambda \in \mathbb{R}^d$, there exists a unique nonnegative root y_λ satisfying $h(\lambda, y_\lambda) = 0$;
(b) $y_0 = 0$; $\lambda \neq 0 \Rightarrow y_\lambda > 0$;
(c) the mapping $\lambda \mapsto y_\lambda$ is continuous.

Necessary conditions on the spectrum of $x(c)$. Let Diag denote the set of $d \times d$ diagonal matrices. Take $c \in \mathcal{S}^d(\mathbb{R})$ and let $x(c) \in \mathcal{S}_+^d(\mathbb{R})$ be a solution (whenever it exists) to (3.3). Then by the spectral theorem, $x(c)$ is diagonalizable: there exists an orthogonal matrix $p_{x(c)}$ such that $p_{x(c)}^* x(c) p_{x(c)} \in$ Diag. Equation (3.3) is stable by unitary transformation

$$\begin{aligned}&2\,\mathrm{Tr}\big(p_{x(c)}^* x(c) p_{x(c)}\big) p_{x(c)}^* x(c) p_{x(c)} + 4\big(p_{x(c)}^* x(c) p_{x(c)}\big)^2\\ &\quad = p_{x(c)}^* c^2 p_{x(c)} \in \mathrm{Diag}.\end{aligned} \tag{A.4}$$

The diagonal elements of $p^*_{x(c)} c^2 p_{x(c)}$ must be the eigenvalues of c^2, that is the square of the eigenvalues of c [which is in $\mathcal{S}^d(\mathbb{R})$]. Identifying the diagonal elements from (A.4) gives a relation between the spectra of c and $x(c)$,

$$2\,\mathrm{Tr}\big(x(c)\big)\lambda_i\big(x(c)\big) + 4\lambda_i\big(x(c)\big)^2 = \lambda_i(c)^2, \qquad 1 \le i \le d.$$

Thus, the nonnegative eigenvalues of $x(c)$ must satisfy $\lambda_i(x(c)) = (-\mathrm{Tr}(x(c)) + \sqrt{\mathrm{Tr}(x(c))^2 + 4\lambda_i(c)^2})/4$. By summing over $i = 1, \dots, d$, we obtain an implicit equation for $\mathrm{Tr}(x(c))$, which is $h(\lambda(c), \mathrm{Tr}(x(c))) = 0$. By $(\star)$, there is a unique solution and

$$\mathrm{Tr}\big(x(c)\big) = y_{\lambda(c)}. \tag{A.5}$$

Thus, we have proved that the eigenvalues of $x(c)$ must be

$$\lambda_i\big(x(c)\big) = \frac{-y_{\lambda(c)} + \sqrt{y^2_{\lambda(c)} + 4\lambda_i(c)^2}}{4}. \tag{A.6}$$

Existence/uniqueness of solution to (3.3). Take $c \in \mathcal{S}^d(\mathbb{R})$. Starting from (3.3), owing to (A.5) $x(c)$ must solve

$$\big(2x(c) + \tfrac{1}{2} y_{\lambda(c)} I_d\big)^2 = \tfrac{1}{4} y^2_{\lambda(c)} I_d + c^2.$$

The matrix $c^2 + \frac{1}{4} y^2_{\lambda(c)} I_d$ is symmetric nonnegative-definite, and thus it has a unique square-root (symmetric nonnegative-definite matrix) [20], Theorem 7.2.6, page 405, and we obtain

$$x(c) := -\frac{y_{\lambda(c)}}{4} I_d + \frac{1}{2}\left(\frac{y^2_{\lambda(c)}}{4} I_d + c^2\right)^{1/2}. \tag{A.7}$$

The uniqueness is proved. It is now easy to check that $x(c)$ given in (A.7) solves (3.3), using the implicit equation satisfied by $\mathrm{Tr}(x(c))$. Last, $\lambda_{\min}(c^2) > 0$ if and only if $\lambda_{\min}(x(c)) > 0$ [owing to (A.6)].

Continuity. From Hoffman and Wielandt's theorem [20], page 368, the function $c \mapsto \lambda(c)$ is continuous on $\mathcal{S}^d(\mathbb{R})$ into $\mathbb{R}^d$. Hence, combined with $(\star)(c)$, we obtain the continuity of $c \mapsto y_{\lambda(c)}$ on $\mathcal{S}^d(\mathbb{R})$ into $\mathbb{R}$.

Then, the continuity of $x(\cdot)$ at $c_0 = 0$ easily follows since as $c \to 0$, $y_{\lambda(c)} \to y_0 = 0$ and $\lambda(x(c)) \to 0$ [using (A.6)]: thus $x(c) \to 0 = x_0$. For $c_0 \neq 0$, we invoke the property that $c \mapsto c^{1/2}$ is locally lipschitz (and even analytic) on $\mathcal{S}^d_{++}(\mathbb{R})$ into $\mathcal{S}^d_{++}(\mathbb{R})$ ([35], Lemma 5.2.1 page 131): we use this with $\frac{y^2_{\lambda(c)}}{4} I_d + c^2 \in \mathcal{S}^d_{++}(\mathbb{R})$ for c close enough to c_0 (using $y_{\lambda(c)} > 0$ for $c \neq 0$). In view of (A.7), the continuity of $x(\cdot)$ at $c_0 \neq 0$ follows.

Proof of $(\star)$. h is continuous on $\mathbb{R}^d \times [0, \infty[$ into $\mathbb{R}$. Moreover:

- $h(\lambda, 0) = -2\sum_{i=1}^d |\lambda_i| \le 0$ and $\lim_{y \to +\infty} h(\lambda, y) = +\infty$;
- h is continuously differentiable on $\mathbb{R}^d \times]0, \infty[$;

- $D_y h(\lambda, y) = 4 + d - \sum_{1\le j\le d} \frac{y}{\sqrt{y^2+4\lambda_i^2}} \ge 4$, implying that $y \mapsto h(\lambda, y)$ is (strictly) increasing.

Then, there is a unique $y_\lambda \in \mathbb{R}_+$ such that $h(\lambda, y_\lambda) = 0$. We point out at first glance, $\lambda \neq 0 \Leftrightarrow y_\lambda > 0$. The continuity of $y_\cdot$ is proved on $\mathbb{R}_*^d$ on the one hand, and at 0 on the other hand.

- On $\mathbb{R}_*^d \times]0, +\infty[: D_y h(\lambda, y)$ exists and is nonzero: then by the implicit function theorem, there exists an open set $U \subset \mathbb{R}_*^d$ containing λ and an open set $V \subset]0, +\infty[$ containing y_λ such that y is continuously differentiable from U to V. This proves the continuously differentiability of $y_\cdot$ in $\mathbb{R}_*^d$.
- At $\lambda = 0 : h((|\lambda|)_{1\le i\le d}, y) \le h(\lambda, y)$ and $y \ge \frac{d|\lambda|}{\sqrt{4+2d}} \Leftrightarrow h((|\lambda|)_{1\le i\le d}, y) \ge 0$. It implies $0 \le y_\lambda \le \frac{d|\lambda|}{\sqrt{4+2d}}$ and $\lim_{|\lambda|\to 0} y_\lambda = 0$.

That completes the continuity of $\lambda \mapsto y_\lambda$ on $\mathbb{R}^d$ and by the previous discussion, the proof of the lemma. $\square$

A.5. Proof of Lemma 3.2. We have $\langle R^n\rangle_T = \int_0^T |\sigma_t^*(D_x u_t - D_x u_{\varphi(t)} - D_{xx}^2 u_{\varphi(t)} \Delta S_t)|^2 \,\mathrm{d}t$: to prove the result, we aim at performing a Taylor expansion using $(\mathbf{A}_u)$, that is, derivatives of u are a.s. finite in a small tube around $(t, S_t, Y_t)_{0\le t\le T}$. Because of this local assumption, a careful treatment is required, which we now detail. In view of $(\mathbf{A}_u)$, there exists $\Omega_{\mathcal{D}}$ such that $\mathbb{P}(\Omega_{\mathcal{D}}) = 1$ and for every $\omega \in \Omega_{\mathcal{D}}$ there is $\delta(\omega) > 0$ such that

$$|\mathcal{A}u|_{\delta}(\omega) := \sup_{0\le t<T} \sup_{|x-S_t(\omega)|\le\delta(\omega), |y-Y_t(\omega)|\le\delta(\omega)} \big|\mathcal{A}u(t, x, y)\big| < +\infty$$

for any $\mathcal{A} \in \mathcal{D} := \{D_{x_j x_k}^2, D_{x_j x_k x_l}^3, D_{t x_j}^2, D_{x_j y_m}^2 : 1 \le j, k, l \le d, 1 \le m \le d'\}$.

Since $\sup_{1\le i\le N_T^n} \Delta\tau_i^n \xrightarrow{a.s.} 0$ and $(S_t, Y_t)_{0\le t\le T}$ are a.s. continuous on the compact interval $[0, T]$, there exists $\Omega_{\mathcal{C}}$ with $\mathbb{P}(\Omega_{\mathcal{C}}) = 1$ such that for every $\omega \in \Omega_{\mathcal{C}}$, there is $p(\omega) \in \mathbb{N}$ such that $\forall n \ge p(\omega)$,

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

Hence for $\omega \in \Omega_{\mathcal{D}} \cap \Omega_{\mathcal{C}}$, let $n \ge p(\omega)$, $i \in \{1, \dots, N_T^n\}$ and $t \in [\tau_{i-1}^n, \tau_i^n]$, and write

$$\begin{aligned}
&D_x u(t, S_t, Y_t) - D_x u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}, Y_{\tau_{i-1}^n}\big) - D_{xx}^2 u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}, Y_{\tau_{i-1}^n}\big)\Delta S_t \\
&\quad = \big[D_x u(t, S_t, Y_t) - D_x u\big(\tau_{i-1}^n, S_t, Y_t\big)\big] \\
&\qquad + \big[D_x u\big(\tau_{i-1}^n, S_t, Y_t\big) - D_x u\big(\tau_{i-1}^n, S_t, Y_{\tau_{i-1}^n}\big)\big] \\
&\qquad + \big[D_x u\big(\tau_{i-1}^n, S_t, Y_{\tau_{i-1}^n}\big) - D_x u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}, Y_{\tau_{i-1}^n}\big) \\
&\qquad\qquad - D_{xx}^2 u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}, Y_{\tau_{i-1}^n}\big)\Delta S_t\big].
\end{aligned}$$

Now apply Taylor's theorem to the terms above, by observing that the involved derivatives of u are locally bounded by the (a.s. finite) random variable $C_u := \max_{\mathcal{A}\in\mathcal{D}} |\mathcal{A}u|_\delta$,

$$\begin{aligned}&\big|D_x u(t, S_t, Y_t) - D_x u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}, Y_{\tau_{i-1}^n}\big) - D_{xx}^2 u\big(\tau_{i-1}^n, S_{\tau_{i-1}^n}, Y_{\tau_{i-1}^n}\big)\Delta S_t\big|\\ &\quad\le \sqrt{d} C_u\bigg(\big(t-\tau_i^n\big) + \sqrt{d'}|Y_t - Y_{\tau_{i-1}^n}| + \frac{d}{2}|\Delta S_t|^2\bigg).\end{aligned}$$

Plugging this estimate into $\langle R^n\rangle_T$ and using that Y is nondecreasing, we derive that a.s. for n large enough,

$$\begin{aligned}&\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T\\ &\quad\le 3dC_u^2 \sup_{0\le t\le T}|\sigma_t|^2\varepsilon_n^{2-4\rho_N}\sum_{\tau_{i-1}^n<T}\bigg(\big(\Delta\tau_i^n\big)^3 + d'|\Delta Y_{\tau_i^n}|^2\Delta\tau_i^n\\ &\qquad + \frac{d^2}{4}\Delta\tau_i^n \sup_{\tau_{i-1}^n\le t\le\tau_i^n}|\Delta S_t|^4\bigg).\end{aligned}$$

To prove the a.s. convergence of the upper bound to 0, we separately analyze each of the three contributions:

- $\varepsilon_n^{2-4\rho_N}\sum_{\tau_{i-1}^n<T}(\Delta\tau_i^n)^3 \le \varepsilon_n^{2-4\rho_N} N_T^n \sup_{1\le i\le N_T^n}(\Delta\tau_i^n)^3 \le C_0\varepsilon_n^{4-3\rho_N} \xrightarrow{a.s.} 0$ by Corollary 2.2(ii) with $\rho = \frac{4}{3} - \rho_N > 0$; see ($\mathbf{A}_N$).
- Combining ($\mathbf{A}_Y$) and Corollary 2.2(ii) with $\rho = \frac{\rho_Y}{2} - 2(\rho_N - 1) > 0$, we easily obtain

$$\begin{aligned}\varepsilon_n^{2-4\rho_N}\sum_{\tau_{i-1}^n<T}|\Delta Y_{\tau_i^n}|^2\Delta\tau_i^n &\le \sum_{j=1}^{d'}\big(Y_T^j - Y_0^j\big)\varepsilon_n^{2-4\rho_N}\sup_{1\le i\le N_T^n}\big|\Delta Y_{\tau_i^n}^j\big|\sup_{1\le i\le N_T^n}\Delta\tau_i^n\\ &\le \sqrt{d'}|Y_T - Y_0|C_0\varepsilon_n^{2-4\rho_N}\varepsilon_n^{\rho_Y}\varepsilon_n^{2-\rho}\\ &\le C_0\varepsilon_n^{\rho_Y/2-2(\rho_N-1)} \xrightarrow{a.s.} 0.\end{aligned}$$

- Using ($\mathbf{A}_S$), $\varepsilon_n^{2-4\rho_N}\sum_{\tau_{i-1}^n<T}\Delta\tau_i^n \sup_{\tau_{i-1}^n\le t\le\tau_i^n}|\Delta S_t|^4 \le C_0\varepsilon_n^{6-4\rho_N}T \xrightarrow{a.s.} 0$ since $\rho_N < \frac{3}{2}$.

All these convergences lead to the results.

A.6. Assumption ($\mathbf{A}_u$). We show that assumption ($\mathbf{A}_u$) is satisfied in most usual situations, even if the payoff g is not smooth. Actually, we have not been able to exhibit an example of g for which ($\mathbf{A}_u$) does not hold. The following discussion should convince the reader that finding a counter-example is far from being straightforward, but we conjecture that it is possible.

Vanilla option in Black–Scholes model. For *pedagogic reasons*, we start with the one-dimensional log-normal model $\mathrm{d}S_t = \sigma S_t \,\mathrm{d}B_t$ ($\sigma > 0$). Consider first the Call option with strike $K > 0$: for $t < T$ we have $D_x u(t,x) = \mathcal{N}(\frac{\log(x/K)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}) \in [0,1]$ where $\mathcal{N}(\cdot)$ is the c.d.f. of the standard Gaussian law. The second derivative writes

$$D^2_{xx}u(t,x) = \frac{1}{\sigma x\sqrt{2\pi(T-t)}}\exp\left(-\frac{1}{2}\left[\frac{\log(x/K)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}\right]^2\right);$$

thus bounding the exponential term by 1, we have for any given $t_0 < T$ $\lim_{\delta\to 0}\sup_{0\le t\le t_0}\sup_{|x-S_t|\le\delta}|D^2_{xx}u(t,x)| \le \frac{1}{\sigma\inf_{0\le t\le T} S_t\sqrt{2\pi(T-t_0)}} < +\infty$. It shows that an a.s. finite bound on the second derivative is available provided that the time to maturity does not vanish. For the third derivative, this is similar: indeed using $\sup_{y\in\mathbb{R}} e^{y^2/4}|\partial_y(e^{-y^2/2})| = \sup_{y\in\mathbb{R}}|y|e^{-y^2/4} = \sqrt{2}e^{-1/2} \le 1$, we deduce

$$|D^3_{xxx}u(t,x)| \le \frac{1+\sigma\sqrt{T}}{x^2\sqrt{2\pi}\sigma^2(T-t)}\exp\left(-\frac{1}{4}\left[\frac{\log(x/K)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}\right]^2\right),$$

and as before $\lim_{\delta\to 0}\sup_{0\le t\le t_0}\sup_{|x-S_t|\le\delta}|D^3_{xxx}u(t,x)| < +\infty$ for any given $t_0 < T$.

The next step consists in deriving a.s. upper bounds on derivatives for arbitrary small time to maturity. We take advantage of the property $\mathbb{P}(S_T \ne K) = 1$, which implies (by a.s. continuity of S) that for $\mathbb{P}$-a.e. ω there exists $t_0(\omega) \in [0,T[$ such that $\inf_{t_0(\omega)\le t\le T}|S_t(\omega) - K| \ge |S_T(\omega) - K|/2 := 2\delta_0(\omega) > 0$. Then, for $t \in [t_0(\omega), T]$ and $\delta \le \delta_0 \wedge [2^{-1}\inf_{0\le t\le T} S_t]$, we have $\inf_{|x-S_t|\le\delta}|\log(x/K)| \ge \inf_{u>0:\,|u-1|\ge\delta_0/K}|\log(u)| := c(\omega) > 0$ and $\inf_{|x-S_t|\le\delta} x \ge S_t/2$: therefore using the inequality $-(\alpha+\beta)^2 \le -\frac{\alpha^2}{2} + \beta^2$, we obtain, for $t \in [t_0(\omega), T[$

$$\sup_{|x-S_t|\le\delta}|D^2_{xx}u(t,x)| \le \frac{2}{\sigma S_t\sqrt{2\pi(T-t)}}\exp\left(-\frac{c^2(\omega)}{4\sigma^2(T-t)} + \frac{1}{8}\sigma^2 T\right).$$

Observe that $c(\omega) > 0$ implies that the above upper bound converges to 0 as $t \to T$: thus, we have completed the proof of $\lim_{\delta\to 0}\sup_{0\le t<T}\sup_{|x-S_t|\le\delta}|D^2_{xx}u(t,x)| < +\infty$ a.s. For the third derivative, similarly we obtain for $t \in [t_0(\omega), T[$ and $\delta \le \delta_0(\omega) \wedge [2^{-1}\inf_{0\le t\le T} S_t(\omega)]$

$$\sup_{|x-S_t|\le\delta}|D^3_{xxx}u(t,x)| \le \frac{4(1+\sigma\sqrt{T})}{S_t^2\sqrt{2\pi}\sigma^2(T-t)}\exp\left(-\frac{c^2(\omega)}{8\sigma^2(T-t)} + \frac{1}{16}\sigma^2 T\right),$$

and we conclude as for the second derivative. To derive the property for $D^2_{tx}u$, we use the relation $D^2_{tx}u = -\frac{1}{2}\sigma^2 x^2 D^3_{xxx}u - \sigma^2 x D^2_{xx}u$. Finally, $(\mathbf{A}_u)$ is proved for the call option (and thus for the put option).

The same argumentation can be applied for the digital call option which payoff is of the form $g(x) = \mathbf{1}_{x\ge K}$: indeed, the derivatives of u blow up only at the discontinuity point K which has null probability for the law of S_T. $(\mathbf{A}_u)$ holds for digital options.

Vanilla option in general local volatility model. The previous arguments are based on the explicit Black–Scholes formula for call and digital call options, but we can generalize them to more general models and payoffs and handle derivatives at any order. Denote by $X^j = \log(S^j)$ $(1 \le j \le d)$ the log-asset price in a diffusion model, and assume that $\mathrm{d}X_t = b^X(t, X_t)\,\mathrm{d}t + \sigma^X(t, X_t)\,\mathrm{d}B_t$ for coefficients b^X and σ^X of class $\mathcal{C}_b^\infty([0,T] \times \mathbb{R}^d)$ (bounded with bounded derivatives). The price function in the log-variables is then $v(t,x) := u(t, \exp(x^1), \dots, \exp(x^d)) = \mathbb{E}(g(S_T)|S_t^j = \exp(x^j), 1 \le j \le d) := \mathbb{E}(G(X_T)|X_t = x)$. We first consider the simple case of $\mathcal{C}^\infty$-payoff G with exponentially bounded derivatives: for any $k \ge 0$, there is a constant $C_k^G \ge 0$ such that $|D_x^k G(x)| \le C_k^G \exp(C_k^G|x|)$ for $x \in \mathbb{R}^d$. In this case, a direct differentiation of $\mathbb{E}(G(X_T)|X_t = x)$ using the smooth flow $x \mapsto X_T^{t,x}$ [26] shows the differentiability of v w.r.t. the space variable with derivatives bounded on compact subsets of $[0,T] \times \mathbb{R}^d$; in addition the time smoothness is obtained using Itô's formula; these arguments are standard and we skip details. $(\mathbf{A}_u)$ is proved for these smooth payoffs.

Now we tackle the case of discontinuous payoffs of the form $G(x) = \mathbf{1}_{x \in \mathcal{D}}\varphi(x)$ for a closed set $\mathcal{D} \subset \mathbb{R}^d$ and a $\mathcal{C}^\infty$-function φ with exponentially bounded derivatives: observe that by combining the analysis for smooth payoffs and that for discontinuous ones will allow to cover a quite large class of g satisfying $(\mathbf{A}_u)$ (such as call/put, digital call/put, exchange call, digital exchange call and so on). We assume that a uniform ellipticity assumption is satisfied: $\inf_{0 \le t \le T, x \in \mathbb{R}^d} \inf_{|\xi|=1} \xi \cdot [\sigma^X(\sigma^X)^*](t,x)\xi > 0$. In this setting, $v(t,x) = \int_{\mathbb{R}^d} \mathbf{1}_{z \in \mathcal{D}} p(t,x,T,z)\varphi(z)\,\mathrm{d}z$ where p is the transition density function of X, which is smooth and satisfies to Aronson-type estimates ([7], Theorem 8, page 263): for any $i \ge 0$ and any differentiation index α, there exists a constant $C_{i,\alpha} = C_{i,\alpha}(T, b^X, \sigma^X) > 0$ such that

$$\left|D_{tx}^{i,\alpha} p(t,x,T,z)\right| \le C_{i,\alpha}(T-t)^{-(d+2i+|\alpha|)/2} \exp\bigl(-|x-z|^2/\bigl[C_{i,\alpha}(T-t)\bigr]\bigr)$$

for any $0 \le t < T$, $x \in \mathbb{R}^d$, $z \in \mathbb{R}^d$. From the integral representation of v, it readily follows that

$$\begin{aligned}
&\left|D_{tx}^{i,\alpha} v(t,x)\right| \\
&\quad \le C_{i,\alpha}(T-t)^{-(2i+|\alpha|)/2} \int_{\mathbb{R}^d} C_0^\varphi e^{C_0^\varphi |z|}(T-t)^{-d/2} e^{-|x-z|^2/[C_{i,\alpha}(T-t)]}\,\mathrm{d}z \\
&\quad \le C_{i,\alpha}(T-t)^{-(2i+|\alpha|)/2} C_0^\varphi e^{C_0^\varphi |x|} \int_{\mathbb{R}^d} e^{C_0^\varphi \sqrt{T}|w|} e^{-|w|^2/C_{i,\alpha}}\,\mathrm{d}w,
\end{aligned}$$

which proves locally uniform bounds on derivatives provided that the time to maturity remains bounded away from 0. To handle the case $t \to T$, we additionally assume that *the boundary $\partial\mathcal{D}$ of $\mathcal{D}$ is Lebesgue-negligible* (thus including usual situations but excluding Cantor like sets; see [5], page 114): thus for $\mathbb{P}$-a.e. ω, the distance to the boundary (a closed set) is positive, that is, $\delta_0(\omega) := \frac{1}{4} d(X_T(\omega), \partial\mathcal{D}) >$

0, and there exists $t_0(\omega) \in [0, T[$ such that $\inf_{t_0(\omega)\le t\le T} d(X_t(\omega), \partial\mathcal{D}) \ge 3\delta_0(\omega)$ [we recall that the distance function $x \mapsto d(x, \partial\mathcal{D})$ is Lipschitz continuous]. Now, let ω be given as above; by the smooth version of the Urysohn lemma [6], page 90, Chapter IV, there exists a smooth function ξ (depending on ω) such that $\mathbf{1}_{x\in\mathcal{D},\delta_0\le d(x,\partial\mathcal{D})} \le \xi(x) \le \mathbf{1}_{x\in\mathcal{D}}$. Decompose the price function into two parts $v = v_1 + v_2$ with

$$v_1(t,x) := \int_{\mathbb{R}^d} \mathbf{1}_{z\in\mathcal{D}}\, p(t,x,T,z)\varphi(z)\xi(z)\,\mathrm{d}z,$$

$$v_2(t,x) = \int_{\mathcal{D}} p(t,x,T,z)\varphi(z)\bigl(1-\xi(z)\bigr)\,\mathrm{d}z.$$

We easily handle the derivatives of v_1 using the first case of smooth functions since $\mathbf{1}_{\mathcal{D}}\varphi\xi = \varphi\xi \in \mathcal{C}^\infty$ with exponentially bounded derivatives. Regarding v_2, observe that we integrate over the z such that $z \in \mathcal{D}$ and $d(z, \partial\mathcal{D}) < \delta_0$; for such z, for $t \in [t_0, T[$ and $|x - X_t| \le \delta \le \delta_0$, we have $|x - z| \ge d(X_t, \partial\mathcal{D}) - |x - X_t| - d(z, \partial\mathcal{D}) \ge \delta_0$ and thus

$$\begin{aligned}
&\sup_{|x-X_t|\le\delta} \bigl|D^{i,\alpha}_{tx} v_2(t,x)\bigr| \\
&\quad\le \sup_{|x-X_t|\le\delta} \int_{\mathcal{D}} C_0^{\varphi} e^{C_0^{\varphi}|z|} C_{i,\alpha}(T-t)^{-(d+2i+|\alpha|)/2} e^{-|x-z|^2/[2C_{i,\alpha}(T-t)]} \\
&\qquad\qquad \times e^{-\delta_0^2/[2C_{i,\alpha}(T-t)]}\,\mathrm{d}z \\
&\quad\le C_{i,\alpha}(T-t)^{-(2i+|\alpha|)/2} e^{-\delta_0^2/[2C_{i,\alpha}(T-t)]} C_0^{\varphi} e^{C_0^{\varphi}(|X_t|+\delta_0)} \\
&\qquad \times \int_{\mathbb{R}^d} e^{C_0^{\varphi}\sqrt{T}|w|} e^{-|w|^2/[2C_{i,\alpha}]}\,\mathrm{d}w.
\end{aligned}$$

The above upper bound converges to 0 as $t \to T$, and the proof of $(\mathbf{A}_u)$ is complete.

Interestingly, we can weaken the ellipticity assumption into a hypoellipticity assumption: indeed, our analysis essentially relies on transition density estimates in small time and away from the diagonal. These estimates are available in the hypoelliptic homogeneous diffusion case ([27], Corollary 3.25) and in the inhomogeneous case [3], Assumption (1.10).

Asian option in general local volatility model. The payoff is of the form $g(S_T, I_T)$ where $I_T = \int_0^T S_t\,\mathrm{d}t$ and S is a one-dimensional homogeneous diffusion $\mathrm{d}S_t = \sigma(S_t)\,\mathrm{d}B_t$. The analysis is reduced to the previous case of vanilla option by considering the 2-dimensional diffusion $(S_t, I_t)_{0\le t\le T}$: it is not elliptic but hypoelliptic [27] provided that σ is smooth and that $\sigma(x) > 0$ for $x \in I$ where $I \subset \mathbb{R}$ is given by $\mathbb{P}(\forall t \in [0, T]: X_t \in I) = 1$ (in usual cases, $I =]0, +\infty[$). It includes the Black–Scholes model and any model with local volatility bounded away from 0 and smooth. We skip details.

Lookback option in Black–Scholes model. The payoff is of the form $S_T - m \wedge \min_{0\le t\le T} S_t$ or $M \vee \max_{0\le t\le T} S_t - S_T$ for lookback call or put, $(M \vee$

$\max_{0\le t\le T} S_t - K)_+$ or $(K - m \wedge \min_{0\le t\le T} S_t)_+$ for call on maximum or on minimum, $(S_T - \lambda m \wedge \min_{0\le t\le T} S_t)_+$ (with $\lambda > 1$) or $(\lambda M \vee \max_{0\le t\le T} S_t - S_T)_+$ (with $\lambda < 1$) for partial lookback call or put. In all these cases, Black–Scholes-type formulas are available in closed forms [4]. Then it is straightforward to check that $(\mathbf{A}_u)$ is satisfied, and this is essentially based on the property that under the assumption of nonzero volatility, the joint law $(S_T, \max_{0\le t\le T} S_t, \min_{0\le t\le T} S_t)$ has a density (derived from [33], Exercise 3.15), implying that the events on which the derivatives may blow up (such as $\{S_T = \min_{0\le t\le T} S_t\}, \dots$) have zero probability.

Acknowledgments. The author is grateful to the Association Nationale de la Recherche Technique and GDF SUEZ for their financial support.

REFERENCES

[1] BERTSIMAS, D., KOGAN, L. and LO, A. W. (2000). When is time continuous? *Journal of Financial Economics* **55** 173–204.

[2] BICHTELER, K. (1981). Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.* **9** 49–89. MR0606798

[3] CATTIAUX, P. and MESNAGER, L. (2002). Hypoelliptic non-homogeneous diffusions. *Probab. Theory Related Fields* **123** 453–483. MR1921010

[4] CONZE, A. and VISWANATHAN (1991). Path dependent options: The case of lookback options. *J. Finance* **46** 1893–1907.

[5] DIBENEDETTO, E. (2002). *Real Analysis*. Birkhäuser, Boston, MA. MR1897317

[6] DIEUDONNÉ, J. (1990). *Eléments d'analyse*. Jacques Gabay, Paris.

[7] FRIEDMAN, A. (1964). *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, NJ. MR0181836

[8] FUKASAWA, M. (2011). Asymptotically efficient discrete hedging. In *Stochastic Analysis with Financial Applications*. *Progress in Probability* **65** 331–346. Birkhäuser, Basel. MR3050797

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

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

[11] GEISS, C., GEISS, S. and GOBET, E. (2012). Generalized fractional smoothness and L_p-variation of BSDEs with non-Lipschitz terminal condition. *Stochastic Process. Appl.* **122** 2078–2216. MR2921973

[12] GEISS, S. and GOBET, E. (2011). Fractional smoothness and applications in finance. In *Advanced Mathematical Methods for Finance* (G. Di Nunno and B. Øksendal, eds.) 313–331. Springer, Heidelberg. MR2792085

[13] GEISS, S. and TOIVOLA, A. (2009). Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces. *Bernoulli* **15** 925–954. MR2597578

[14] GENON-CATALOT, V. and JACOD, J. (1993). On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **29** 119–151. MR1204521

[15] GOBET, E. and MAKHLOUF, A. (2010). $\mathbf{L}_2$-time regularity of BSDEs with irregular terminal functions. *Stochastic Process. Appl.* **120** 1105–1132. MR2639740

[16] GOBET, E. and MAKHLOUF, A. (2012). The tracking error rate of the Delta-Gamma hedging strategy. *Math. Finance* **22** 277–309. MR2897386

[17] GOBET, E. and MENOZZI, S. (2010). Stopped diffusion processes: Boundary corrections and overshoot. *Stochastic Process. Appl.* **120** 130–162. MR2576884

[18] GOBET, E. and TEMAM, E. (2001). Discrete time hedging errors for options with irregular payoffs. *Finance Stoch.* **5** 357–367. MR1849426

[19] HAYASHI, T. and MYKLAND, P. A. (2005). Evaluating hedging errors: An asymptotic approach. *Math. Finance* **15** 309–343. MR2132193

[20] HORN, R. A. and JOHNSON, C. R. (1990). *Matrix Analysis*. Cambridge Univ. Press, Cambridge. MR1084815

[21] JACOD, J. and PROTTER, P. (2012). *Discretization of Processes*. *Stochastic Modelling and Applied Probability* **67**. Springer, Heidelberg. MR2859096

[22] KARANDIKAR, R. L. (1989). On Métivier–Pellaumail inequality, Emery topology and pathwise formulae in stochastic calculus. *Sankhyā Ser. A* **51** 121–143. MR1065565

[23] KARANDIKAR, R. L. (1995). On pathwise stochastic integration. *Stochastic Process. Appl.* **57** 11–18. MR1327950

[24] KARANDIKAR, R. L. (2006). On almost sure convergence results in stochastic calculus. In *In Memoriam Paul-André Meyer*: *Séminaire de Probabilités XXXIX* (M. Émery et al., eds.). *Lecture Notes in Math.* **1874** 137–147. Springer, Berlin. MR2276893

[25] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940

[26] KUNITA, H. (1984). Stochastic differential equations and stochastic flows of diffeomorphisms. In *École D'été de Probabilités de Saint-Flour, XII*—1982. *Lecture Notes in Math.* **1097** 143–303. Springer, Berlin. MR0876080

[27] KUSUOKA, S. and STROOCK, D. (1985). Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32** 1–76. MR0783181

[28] LANDON, N. (2013). Almost sure optimal stopping times: Theory and applications. Ph.D. thesis, Ecole Polytechnique. Available at http://pastel.archives-ouvertes.fr/docs/00/78/80/67/PDF/Thesis_Nicolas_Landon.pdf.

[29] LENGLART, E. (1977). Relation de domination entre deux processus. *Ann. Inst. H. Poincaré Sect. B* (*N.S.*) **13** 171–179. MR0471069

[30] MARTINI, C. and PATRY, C. (1999). Variance optimal hedging in the Black–Scholes model for a given number of transactions. INRIA Rapport de Recherche, No. 3767.

[31] MUSIELA, M. and RUTKOWSKI, M. (2005). *Martingale Methods in Financial Modelling*, 2nd ed. *Stochastic Modelling and Applied Probability* **36**. Springer, Berlin. MR2107822

[32] PASCUCCI, A. (2011). *PDE and Martingale Methods in Option Pricing*. *Bocconi & Springer Series* **2**. Springer, Milan. MR2791231

[33] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* **293**. Springer, Berlin. MR1725357

[34] ROSENBAUM, M. and TANKOV, P. (2014). Asymptotically optimal discretization of hedging strategies with jumps. *Ann. Appl. Probab.* To appear.

[35] STROOCK, D. W. and VARADHAN, S. R. S. (2006). *Multidimensional Diffusion Processes*. Springer, Berlin. MR2190038

[36] WILMOTT, P., DEWYNNE, J. and HOWISON, S. (1994). *Option Pricing*: *Mathematical Models and Computation*. Oxford Financial Press, Oxford.

[37] ZHANG, R. (1999). Couverture approchée des options européennes. Ph.D. thesis, Ecole Nationale des Ponts et Chaussées. Available at http://tel.archives-ouvertes.fr/docs/00/04/65/68/PDF/tel-00005623.pdf.

CENTRE DE MATHEMATIQUES APPLIQUÉES
ECOLE POLYTECHNIQUE AND CNRS
91128 PALAISEAU
FRANCE
E-MAIL: emmanuel.gobet@polytechnique.edu

GDF SUEZ
DIRECTION DE LA STRATÉGIE
ET DU DÉVELOPPEMENT DURABLE
22, RUE DU DOCTEUR LANCEREAUX
75008 PARIS
FRANCE
AND
CENTRE DE MATHEMATIQUES APPLIQUÉES
ECOLE POLYTECHNIQUE AND CNRS
91128 PALAISEAU
FRANCE
E-MAIL: landon.nico@gmail.com