The tracking error rate of the Delta-Gamma hedging strategy

Emmanuel Gobet, Azmi Makhlouf

The tracking error rate of the Delta-Gamma hedging strategy ⋆

Emmanuel Gobet [a,*], Azmi Makhlouf [a]

[a] *Laboratoire Jean Kuntzmann, Université de Grenoble and CNRS, BP 53, 38041 Grenoble Cedex 9, FRANCE*

Current version: 2 July 2009

Abstract

We analyse the convergence rate of the quadratic tracking error, when a Delta-Gamma hedging strategy is used at N discrete times. The fractional regularity of the payoff function plays a crucial role in the choice of the trading dates, in order to achieve optimal rates of convergence.

Key words: hedging strategies, fractional regularity, $\mathbf{L}_2$ convergence

1991 MSC: 91Bxx, 60Fxx

Statement of the problem. An investor who sells an option should hedge against the variations of the underlying asset price, either by making a static replication, or

⋆ This work was supported in part by a MSTIC Grant from University Joseph Fourier (project REFINE).

* Corresponding author
Email addresses: emmanuel.gobet@imag.fr (Emmanuel Gobet), azmi.makhlouf@imag.fr (Azmi Makhlouf).

by making a dynamic hedging strategy, i.e. trading continuously in time in the hedging instruments. In practice, this is often done through the so-called delta hedging strategy (DHS in short), which ensures that the investor's portfolio remains delta-neutral (i.e. equalizing the first sensitivity δ of the hedging portfolio and that of the option value). When the market is complete (we assume this in all what follows), in a Markovian setting where the asset is modeled through a stochastic differential equation and the payoff at time T is of the form $g(S_T)$, it is known (cf Karatzas and Shreve (1998)) that the number of assets to hold at time t is given by $\delta_t = \partial_S u(t, S_t)$ where u is the price function (solution of a valuation PDE). Theoretically, it means that the investor must trade, continuously at each time t between 0 and the maturity, in order to obtain zero residual risk.

However, due to practical considerations and to the transaction costs, one can only use discrete-time hedging strategies. Here we do not consider the impact of transaction costs (we refer to Kabanov and Safarian (1997), Pergamenshchikov (2003) and references therein) and we focus only on the impact of discrete-time rebalancing. Thus, the number of assets between two successive rebalancing dates t_i and t_{i+1} is fixed to δ_{t_i}, and the associate hedging portfolio has a terminal value V_T^N which may differ from the required payoff $g(S_T)$. In other words, this induces a tracking error $\mathcal{E}_N^\Delta := V_T^N - g(S_T)$, which has been intensively studied in the literature. For Call/Put option, in Zhang (1999) it is shown that the $\mathbf{L}_2$ error has a convergence rate $N^{1/2}$ for uniform rebalancing dates $t_i = iT/N$. In Gobet and Temam (2001), it is shown that the rate of convergence actually strongly depends on the smoothness of the payoff function g: for instance for digital options, the order of convergence becomes 1/4 instead of 1/2. Geiss and coauthors deeply investigated these features: in a series of papers initiated by Geiss (2002), they paved the way to connect the fractional regularity of the payoff and the rate of convergence of the related discrete time delta-hedging strategy. Moreover, they showed that for a given payoff, a suitable

non-uniform grid with N dates can be choosen to achieve the rate $N^{1/2}$: the more irregular the payoff, the more concentrated near T the points. Recently in Gobet and Makhlouf (2008), these ideas have been extended to BSDEs and to multidimensional diffusion models. One could alternatively measure the error using weak convergence techniques, instead of $\mathbf{L}_2$ norm. Surprisingly, the convergence rate may be $N^{1/2}$ even for non-smooth payoffs, i.e. $N^{1/2}\mathcal{E}_N^{\Delta}$ weakly converges as N goes to infinity to a non trivial random variable, which may not be square-integrable [1] for non-smooth payoffs. This convergence phenomenon has been noticed in Gobet and Temam (2001) for digital options, then more systematically studied in Hayashi and Mykland (2005). The integrability of the weak limit for non-smooth payoffs is investigated in details in Geiss and Toivola (2008).

In all these works, it is pointed out that the weak limit is related to a suitably weighted integral of the so-called Gamma $(\partial_S^2 u(t,S_t))_{0\leq t<T}$ of the option [2]. Thus, in order to reduce the tracking error, one can think of reducing the *portfolio Gamma* by incorporating in the hedging portfolio another instrument, the price of which is $(C(t,S_t))_{0\leq t\leq T}$: in practice, it is a liquid vanilla option (At-The-Money Call/Put). The resulting strategy is called Delta-Gamma hedging strategy (DGHS in short) and is well-known from practionners (see Hull (2009) p.371). It is obtained by equalizing the second sensitivity of the hedging portfolio and that of the option value. It follows that the numbers of options C and assets to hold at time t_i are equal to

$$\delta_{t_i}^C := \frac{\partial_S^2 u(t_i,S_{t_i})}{\partial_S^2 C(t_i,S_{t_i})}, \qquad \delta_{t_i} := \partial_S u(t_i,S_{t_i}) - \frac{\partial_S^2 u(t_i,S_{t_i})}{\partial_S^2 C(t_i,S_{t_i})}\partial_S C(t_i,S_{t_i}). \tag{0.1}$$

This specification of piecewise-constant strategy on each interval $[t_i,t_{i+1}[$ leads to a

[1] This gives a non trivial situation where the rate of convergence depends on the convergence criterion ($\mathbf{L}_2$ convergence or weak convergence).

[2] Similar integrals of the Gamma also appear in the measure of robustness of the Black and Scholes formula, cf El Karoui, Jeanblanc-Picqué and Shreve (1998).

final wealth process V_T^N hopefully closer to $g(S_T)$ compared with that of the simple DHS.

Objectives. The purpose of this work is to quantify how much the resulting tracking error $\mathcal{E}_N^{\Delta\Gamma} := V_T^N - g(S_T)$ is reduced with respect to the number of rebalancing dates. More precisely, we aim at connecting the convergence rate of $\mathcal{E}_N^{\Delta\Gamma}$ to the payoff regularity, regarding the choice of rebalancing dates. These issues can be investigated under several points of view.

- Choice of the convergence criterion. We conjecture that in the usual cases, the rate of weak convergence of $\mathcal{E}_N^{\Delta}$ is N (instead of $N^{1/2}$), extending the results in Hayashi and Mykland (2005), with a possible non square-integrable limit (our numerical experiments support this phenomenon). We leave the proof of this conjecture to further research and in this work, we prefer to investigate the $\mathbf{L}_2$ convergence because it leads to a more intuitive dispersion measure of the tracking error. Our results below state that equidistant trading dates with DGHS do not systematically yield a better convergence order w.r.t. N compared with DHS (see Theorem 6). But an appropriate choice of trading dates leads to a quadratic error equal to a $O(N^{-1})$. The choice of the trading dates is explicit and depends on the fractional regularity of the payoff.
- Specification of the model for S. In this work, we emphasize the tuning of the trading dates according to the payoff regularity. Analogously to the results in Geiss (2002) or Gobet and Makhlouf (2008), we believe that the results on the rate of convergence (and thus the choice of trading dates) hold for a wide class of SDE model on S. To simplify the analysis, we only consider geometric Brownian motions for the modeling of S (see (1.1)). Nevertheless, the analysis remains rather intricate, but we hope that this model simplification will help the reader to focus

on the measure of the payoff regularity, which is driven by the behavior of the expected conditional variance $V_{t,T}(g) = \mathbb{E}|g(S_T) - \mathbb{E}^{\mathcal{F}_t}(g(S_T))|^2$ as t goes to T. On the other hand, we allow the payoff to depend on several assets (basket options for instance). For this, we extend the single-asset formula (0.1) for DGHS to the multi-asset case, involving the *cross-Gammas* of the options.

- Choice of hedging instruments. Obviously, the instruments C used for DGHS have to yield non-zero gammas, in order to well define the strategy (see (0.1) for the single-asset case and Definition 1 for the multi-asset case). To manage the individual gammas, the most natural choice is Call/Put options whose price are convex in the spot variable (see Martini (1999) for general results on the propagation of convexity,). To handle cross gammas, we propose to use exchange options (spread option with strike zero). Since the models for S are correlated geometric Brownian motions, whenever needed we can take advantage of the closed Margrabe formula for such an exchange option (see Appendix A.1).

The closest related work to ours is the one by Brodén and Wiktorsson (2008). In this reference, S is a one-dimensional asset, following a local volatility model. The payoff function g is the Call payoff and uniform trading dates are studied. They state that the $\mathbf{L}_2$ convergence holds at rate $N^{3/4}$. However, there are some gaps in their proofs, in particular regarding some key estimates related to lower/upper bounds for the second derivatives of the price function. The differences between their work and ours are the following: we restrict to log-normal modelling; we extend the study of DGHS to a multidimensional framework; we study the connection between the payoff regularity and the choice of the trading dates in order to achieve optimal convergence rates. Furthermore, our proofs leverage the intrinsic martingale properties of price processes (see Lemma 2), which leads to simplified computations.

Organization of the paper. In the next section, we define the stochastic model for the underlying assets, the payoff option to hegde and the other instruments used in the hedging portfolio. We state some preliminary results related to the Greeks' processes as martingales. We also introduce notations used in all the paper, in particular for the fractional regularity of the payoff function. The section 2 contains our main results: we define the DGHS strategy, show that the $\mathbf{L}_2$ norm of the tracking error is essentially measured by a suitable integral of the third derivatives of the price function, which can be accurately estimated through the payoff regularity. Then, we give sufficient conditions on the N trading dates to yield a tracking error of order $1/N$ as the number N of dates goes to infinity. In section 3, we present numerical results which corroborate these features. Section 4 is devoted to the proofs of the main results.

1 Preliminaries

1.1 Asset stochastic model, payoff, additional hedging instruments

Hereafter, $\hat{W} = (\hat{W}^1, ..., \hat{W}^d)$ is a d-dimensional Brownian motion, defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ is the historical probability and $(\mathcal{F}_t)_{0 \leq t \leq T}$ (T is a fixed terminal time) is the natural filtration of $\hat{W}$, augmented with $\mathbb{P}$-null sets. We assume that the components of $\hat{W}$ are correlated, with $\langle \hat{W}^j, \hat{W}^k \rangle_t = \rho_{j,k} t$, and that the matrix $(\rho_{j,k})_{1 \leq j,k \leq d}$ has a full rank.

Assets. We denote the non-risky asset by S^0 (bank account) and the risky assets by $(S^j)_{1 \leq j \leq d}$. Although not essential for our results, we assume that the risky assets do not pay dividends and that the interest rates are constant and equal to r. Morever,

we assume that the risky assets $(S^j)_{1\leq j\leq d}$ are lognormal processes with constant volatilities $(\sigma_j)_{1\leq j\leq d}$, constant historical drifts $(\mu_j)_{1\leq j\leq d}$ and correlations equal to $(\rho_{j,k})_{j,k}$. It writes for $1 \leq j \leq d$:

$$\text{(1.1)} \qquad \begin{cases} S_0^j &= s_0^j, \\ \mathrm{d}S_t^j &= \mu_j S_t^j \mathrm{d}t + \sigma_j S_t^j \mathrm{d}\hat{W}_t^j, \end{cases}$$

where S_t^j, s_0^j and σ_j are positive. Set $\lambda_j = \frac{\mu_j - r}{\sigma_j}$: then, we can define the so-called risk-neutral measure $\mathbb{Q}$ such that $(W_t^j := \hat{W}_t^j + \lambda_j t)_{1\leq j\leq d}$ is a $\mathbb{Q}$-Brownian motion (the correlations remaining unchanged). Up to the correlation factor, λ is the market price of risk. Under $\mathbb{Q}$, the model writes

$$\begin{cases} S_0^j &= s_0^j, \\ \mathrm{d}S_t^j &= r S_t^j \mathrm{d}t + \sigma_j S_t^j \mathrm{d}W_t^j. \end{cases}$$

By a slight abuse of notation by setting $\sigma_0 := 0$ and $S_0^0 := 1$, we make the above equation also valid for $j = 0$, which might be useful in the following when short notations are needed.

At least, note that $\bar{S}_t^j := e^{-rt} S_t^j$ is a $\mathbb{Q}$-martingale (for any $0 \leq j \leq d$), and $\mathrm{d}\bar{S}_t^j = \sigma_j \bar{S}_t^j \mathrm{d}W_t^j$. More generally, if U is a stochastic process, $\bar{U}$ denotes its discounted value:

$$\bar{U}_t = e^{-rt} U_t.$$

When an expectation is computed under $\mathbb{P}$, we write $\mathbb{E}_{\mathbb{P}}(.)$, while under $\mathbb{Q}$ we write $\mathbb{E}_{\mathbb{Q}}(.)$. We add the superscript $\mathcal{F}_t$ to indicate the conditional expectations given $\mathcal{F}_t$, i.e. $\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}(.)$ and $\mathbb{E}_{\mathbb{Q}}^{\mathcal{F}_t}(.)$.

The payoff g. In what follows, g denotes the payoff function that defines the option to hedge. The price function of this option is then defined by $u(t,S) := \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}g(S_T)|S_t = S\right]$, for $S = (S^1, ..., S^d) \in \mathbb{R}_+^d$. We assume that $\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0} < \infty$ for some $p_0 > 1$ (for instance, $|g(S)| \leq C(1+|S|^n+|S|^{-n})$ for some $n \in \mathbb{N}$). Without additional regularity assumption on g, it is easy to check that u is a C^∞ function for $t < T$. In particular, for $l, m, n = 1...d$, $\partial_t u$, $\partial_{S^l} u$, $\partial^2_{S^l,S^m} u$ and $\partial^3_{S^l,S^m,S^n} u$ exist and are continuous for $t < T$.

As in Geiss (2002) and Gobet and Makhlouf (2008), the fractional regularity of the payoff function g is measured through

$$V_{t,T}(g) := \mathbb{E}_{\mathbb{P}}\left|g(S_T) - \mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}}(g(S_T))\right|^2,$$

and we consider the following space ($\alpha \in (0,1]$)

$$\mathbf{L}_{2,\alpha} = \{g \text{ s.t. } \mathbb{E}|g(X_T)|^2 + \sup_{0\leq t<T} \frac{V_{t,T}(g)}{(T-t)^\alpha} < +\infty\}.$$

It describes the rate of decreasing of the expected conditional variance of $g(S_T)$ given $\mathcal{F}_t$ as t goes to T. When g belongs to $\mathbf{L}_{2,\alpha}$ and $g(S_T) \in \mathbf{L}_{2p_0}$ (for some $p_0 > 1$), we can define $K^\alpha(g)$ and $K^{\alpha,p_0}(g)$ as

$$\begin{aligned} K^\alpha(g) &:= \mathbb{E}_{\mathbb{P}}|g(S_T)|^2 + \sup_{t\in[0,T)} \frac{V_{t,T}(g)}{(T-t)^\alpha}, \\ K^{\alpha,p_0}(g) &:= \left(\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0}\right)^{\frac{1}{p_0}} + \sup_{t\in[0,T)} \frac{V_{t,T}(g)}{(T-t)^\alpha}. \end{aligned} \tag{1.2}$$

Notice that $\bigcup_{\alpha\in(0,1]} \mathbf{L}_{2,\alpha}$ obviously contains uniformly Hölder continuous functions, but also some non-smooth functions. Some examples are given in the lemma below, whose proof is done in Appendix A.2.

Lemma 1

- Call/Put options. *If* $g(S) = (S^j - K)_+$ *or* $g(S) = (K - S^j)_+$ *(for* $1 \leq j \leq d$*), then*

$g \in \mathbf{L}_{2,\alpha}$ *with* $\alpha = 1$.

- Digital options. *If* $g(S) = \mathbb{1}_{S^j > K}$ *or* $g(S) = \mathbb{1}_{S^j < K}$ *(for* $1 \leq j \leq d$*), then* $g \in \mathbf{L}_{2,\alpha}$ *with* $\alpha = 1/2$.
- Stability by summation and product. *Let* g_1 *and* g_2 *belong resp. to* $\mathbf{L}_{2,\alpha_1}$ *and* $\mathbf{L}_{2,\alpha_2}$. *Then*
  - $g_1 + g_2 \in \mathbf{L}_{2,\alpha}$ *with* $\alpha = \alpha_1 \wedge \alpha_2$;
  - *If in addition* g_1 *and* g_2 *are bounded, then* $g_1 g_2 \in \mathbf{L}_{2,\alpha}$ *with* $\alpha = \alpha_1 \wedge \alpha_2$.

A straightforward application of the above rules shows for instance that Put payoffs with digital triggers (of the form $g(S) = \mathbb{1}_{S^1 < K_1, \cdots, S^{d-1} < K_{d-1}}(K_d - S^d)_+$) belong to $\mathbf{L}_{2,\alpha}$ with $\alpha = 1/2$.

The additional instruments. To perform the DGHS, we need $d(d+1)/2$ extra financial instruments that have non vanishing gammas. To simplify the exposure, we assume that these instruments are options with the same time maturity $T_2 > T$. Their payoffs are given by (for $0 \leq j < k \leq d$)

$$(S^k_{T_2} - K_{j,k} S^j_{T_2})_+ \tag{1.3}$$

($K_{j,k} > 0$). If $j = 0$ and $1 \leq k \leq d$, it corresponds to a Call on the asset S^k with strike $K_{0,k} e^{rT_2}$. If $1 \leq j < k \leq d$, it is an exchange option between S^j and S^k. The price function of each of these instruments writes

$$C^{j,k}(t, S^j, S^k) := \mathbb{E}_{\mathbb{Q}} \left[e^{-r(T_2 - t)} (S^k_{T_2} - K_{j,k} S^j_{T_2})_+ | S^j_t = S^j, S^k_t = S^k \right].$$

Since the model is log-normal with constant correlation, the price functions are explicit and are given by the Black-Scholes and Margrabe formulas (see Appendix A.1).

1.2 Greeks and martingales

In this paragraph, we introduce some specific martingales which are related to the option Greeks. These properties are quite standard, but to our knowledge their importance in the analysis of the tracking error has not been emphasized so far: in particular, leveraging this property, we will deduce that the tracking error is an iterated stochastic integral w.r.t. the $\mathbb{Q}$-Brownian motion W (see Theorem 3).

Notation. For any function $\psi(t, S^1, ..., S^d)$ and a multi-index $(l_1, ..., l_p) \in \{1, ..., d\}^p$, we will denote $\partial^p_{S^{l_1},...,S^{l_p}}\psi(t, S^1_t, ..., S^d_t)$ by $\partial^p_{l_1,...,l_p}\psi(t)$.

For $l, m, n = 1...d$, we define

$$\bar{u}(t) := e^{-rt}u(t); \tag{1.4}$$

$$\bar{u}^{(1)}_l(t) := e^{-rt}\sigma_l S^l_t \partial_l u(t); \tag{1.5}$$

$$\bar{u}^{(2)}_{l,m}(t) := e^{-rt}\sigma_l\sigma_m S^l_t S^m_t \partial^2_{l,m}u(t); \tag{1.6}$$

$$\bar{u}^{(3)}_{l,m,n}(t) := e^{-rt}\sigma_l\sigma_m\sigma_n S^l_t S^m_t S^n_t \partial^3_{l,m,n}u(t). \tag{1.7}$$

In the same way, we define, for $0 \leq j < k \leq d$ and $l, m, n = 1...d$,

$$\begin{aligned}
\bar{C}^{j,k}(t) &:= e^{-rt}C^{j,k}(t);\\
\bar{C}^{j,k,(1)}_l(t) &:= e^{-rt}\sigma_l S^l_t \partial_l C^{j,k}(t);\\
\bar{C}^{j,k,(2)}_{l,m}(t) &:= e^{-rt}\sigma_l\sigma_m S^l_t S^m_t \partial^2_{l,m}C^{j,k}(t);\\
\bar{C}^{j,k,(3)}_{l,m,n}(t) &:= e^{-rt}\sigma_l\sigma_m\sigma_n S^l_t S^m_t S^n_t \partial^3_{l,m,n}C^{j,k}(t).
\end{aligned}$$

The lemma below will be useful.

Lemma 2 *The processes above are $\mathbb{Q}$-martingales on [0,T), and their Itô decompo-*

sitions w.r.t. the $\mathbb{Q}$-Brownian motions W write

$$\mathrm{d}\bar{u}(t) = \sum_{l=1}^{d} \bar{u}_l^{(1)}(t)\mathrm{d}W_t^l; \tag{1.8}$$

$$\mathrm{d}\bar{u}_l^{(1)}(t) = \sum_{m=1}^{d} \left(\bar{u}_{l,m}^{(2)}(t) + \sigma_l \bar{u}_l^{(1)}(t)\mathbb{1}_{m=l}\right)\mathrm{d}W_t^m; \tag{1.9}$$

$$\mathrm{d}\bar{u}_{l,m}^{(2)}(t) = \sum_{n=1}^{d} \left(\bar{u}_{l,m,n}^{(3)}(t) + \sigma_m \bar{u}_{l,m}^{(2)}(t)\mathbb{1}_{n=m} + \sigma_l \bar{u}_{l,m}^{(2)}(t)\mathbb{1}_{n=l}\right)\mathrm{d}W_t^n. \tag{1.10}$$

Similarly, one has

$$\begin{aligned}
\mathrm{d}\bar{C}^{j,k}(t) &= \sum_{l=1}^{d} \bar{C}_l^{j,k,(1)}(t)\mathrm{d}W_t^l; \\
\mathrm{d}\bar{C}_l^{j,k,(1)}(t) &= \sum_{m=1}^{d} \left(\bar{C}_{l,m}^{j,k,(2)}(t) + \sigma_l \bar{C}_l^{j,k,(1)}(t)\mathbb{1}_{m=l}\right)\mathrm{d}W_t^m; \\
\mathrm{d}\bar{C}_{l,m}^{j,k,(2)}(t) &= \sum_{n=1}^{d} \left(\bar{C}_{l,m,n}^{j,k,(3)}(t) + \sigma_m \bar{C}_{l,m}^{j,k,(2)}(t)\mathbb{1}_{n=m} + \sigma_l \bar{C}_{l,m}^{j,k,(2)}(t)\mathbb{1}_{n=l}\right)\mathrm{d}W_t^n.
\end{aligned} \tag{1.11}$$

Proof. We show the result for the martingales involving u. The proof is the same for those involving $C^{j,k}$.

Let us define $\bar{u}(t,S) := e^{-rt}u(t,S)$, for $S = (S^1,...,S^d) \in \mathbb{R}_+^d$. Since $\bar{u}(t,S_t) := \mathbb{E}_{\mathbb{Q}}^{\mathcal{F}_t}\left[e^{-rT}g(S_T)\right]$, clearly $(\bar{u}(t,S_t))_{t\leq T}$ is a $\mathbb{Q}$-martingale for any initial values $(s_0^j)_{1\leq j\leq d}$. Then (see e.g Gobet and Munos (2005)), for a multi-index $l = (l_1,...,l_d)$, the process $\partial^{|l|}_{(s_0^1)^{l_1},...,(s_0^d)^{l_d}}\bar{u}(t,S_t^1,...,S_t^d)$, which is equal to $(\frac{S_t^1}{s_0^1})^{l_1}...(\frac{S_t^d}{s_0^d})^{l_d}\partial^{|l|}_{l_1,\cdots,l_d}\bar{u}(t)$ (since S_t^j is linear w.r.t. s_0^j), is a martingale (the estimates on the regularity for $t<T$ are given later in Lemma 7). Taking $|l|$ equal to 1 (resp. 2, resp. 3), we obtain that the process in (1.5) (resp. (1.6), resp. (1.7)) is a $\mathbb{Q}$-martingale on $[0,T[$.

Concerning the differentials, it is clear that, if $(\psi(t))_t = \left(\psi(t,S_t^1,...,S_t^d)\right)_t$ is a $\mathbb{Q}$-martingale where ψ is a smooth function, then its Itô decomposition contains only terms w.r.t. $\mathrm{d}W$:

$$\mathrm{d}\psi(t) = \sum_{l=1}^{d} \sigma_l S_t^l \partial_l \psi(t)\mathrm{d}W_t^l.$$

By taking $\psi(t) = \bar{u}(t)$ (resp. $\bar{u}_l^{(1)}(t)$, resp. $\bar{u}_{l,m}^{(2)}(t)$), the relation (1.8) (resp. (1.9), resp. (1.10)) follows. □

1.3 Other notations

The time net. In all what follows, $\pi := (t_k)_{k=0\dots N}$ is a deterministic time net, such that $0 = t_0 < t_1 < ... < t_N = T$, and $|\pi| := \sup_{k=0\dots N-1}(t_{k+1} - t_k)$. This defines the N trading dates of the DGHS. We shall use the following net ($\beta \in (0, 1]$)

$$\pi^{(\beta)} := \Big\{ t_k^{(N,\beta)} := T - T\Big(1 - \frac{k}{N}\Big)^{\frac{1}{\beta}}, 0 \le k \le N \Big\}.$$

Note that $\pi^{(1)} = (t_k^{(N,1)})$ coincides with the equidistant net. For $\beta < 1$, the points in $\pi^{(\beta)}$ are more concentrated near T.

The constants. In our estimates, we use numerous constants that will be generically denoted by C or c. We emphasize the fact that, whenever a constant depends on the function g, the dependence will be expressed explicitly, so that all the constants such as C or c do not depend on g, but may depend on $r, (\mu_j)_j, (\sigma_j)_j, (\rho_{j,k})_{j,k}, (K_{j,k})_{j,k}, (s_0^j)_j, \alpha$ and other universal constants. They may also depend on T and on T_2, but remain bounded when $T \to 0$.

$A \le_c B$ means $A \le cB$ with a generic constant c.

2 The Delta-Gamma hedging strategy

2.1 Decomposition of the tracking error as multiple Itô integrals

In order to illustrate the ideas and results of the multidimensional case, which will be studied in paragraph 2.1.2, let us show what happens in the one-dimensional case ($d = 1$), with a single asset S^1.

2.1.1 The one-dimensional case

In that case, the additional hedging instrument is simply a Call on S^1 with strike $K_{0,1}S^0_{T_2}$ at maturity T_2 (see (1.3)). The related price function $C^{0,1}(t, S^0, S^1)$ is the Black-Scholes formula, where the usual variables are t and S^1.

Now let us decompose the tracking error. Taking advantage of the self-financing condition, the discounted terminal value of the hedging portfolio $\bar{V}^N_T = e^{-rT}V^N_T$ is

$$\begin{aligned}\bar{V}^N_T :=& \bar{u}(0, S_0) + \sum_{i=0}^{N-1} \delta^1_{t_i}(\bar{S}^1_{t_{i+1}} - \bar{S}^1_{t_i}) \\ &+ \sum_{i=0}^{N-1} \delta^{0,1}_{t_i}(\bar{C}^{0,1}(t_{i+1}, S^0_{t_{i+1}}, S^1_{t_{i+1}}) - \bar{C}^{0,1}(t_i, S^0_{t_i}, S^1_{t_i})),\end{aligned} \tag{2.1}$$

where

$$\delta^{0,1}_{t_i} = \frac{\bar{u}^{(2)}_{1,1}(t_i)}{\bar{C}^{0,1,(2)}_{1,1}(t_i)}, \tag{2.2}$$

$$\delta^1_{t_i} = \frac{1}{\sigma_1 \bar{S}^1_{t_i}} \bar{u}^{(1)}_1(t_i) - \frac{1}{\sigma_1 \bar{S}^1_{t_i}} \delta^{0,1}_{t_i} \bar{C}^{0,1,(1)}_1(t_i). \tag{2.3}$$

Note that $\delta^{0,1}_{t_i} = \frac{\partial^2_{S^1,S^1} u(t,S^1_t)}{\partial^2_{S^1,S^1} C^{0,1}(t,S^0_t,S^1_t)}$ and $\delta^1_{t_i} = \partial_{S^1} u(t, S^1_t) - \delta^{0,1}_{t_i} \partial_{S^1} C^{0,1}(t, S^0_t, S^1_t)$, which shows that the above representation is equivalent to that mentioned in the introduction (see identities (0.1)). However, representations (2.2) and (2.3) are more convenient for the mathematical analysis below.

The discounted tracking error is defined by

$$\begin{aligned}\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi) &:= \bar{V}_T^N - e^{-rT}g(S_T)\\ &= \bar{V}_T^N - \bar{u}(T,S_T)\\ &= \bar{V}_T^N - (\bar{u}(0,S_0) + \int_0^T \bar{u}_1^{(1)}(t)\mathrm{d}W_t),\end{aligned}$$

where we have used the Itô decomposition (1.8). Combining this with (2.1), one has

$$\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi) = -\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} T^{i,(1)}(t)\mathrm{d}W_t, \tag{2.4}$$

where

$$T^{i,(1)}(t) := \bar{u}_1^{(1)}(t) - \delta_{t_i}^1\sigma_1\bar{S}_t^1 - \delta_{t_i}^{0,1}\bar{C}_1^{0,1,(1)}(t).$$

Then, using (2.3), we observe that

$$T^{i,(1)}(t_i) = \bar{u}_1^{(1)}(t_i) - \left(\frac{\bar{u}_1^{(1)}(t_i)}{\sigma_1\bar{S}_{t_i}^1} - \frac{\delta_{t_i}^{0,1}\bar{C}_1^{0,1,(1)}(t_i)}{\sigma_1\bar{S}_{t_i}^1}\right)\sigma_1\bar{S}_{t_i}^1 - \delta_{t_i}^{0,1}\bar{C}_1^{0,1,(1)}(t_i) = 0.$$

This is the usual Delta-hedging condition. Applying Itô's rule and using Lemma 2, one obtains

$$T^{i,(1)}(t) = \int_{t_i}^t T^{i,(2)}(s)\mathrm{d}W_s \tag{2.5}$$

for $t \in [t_i, t_{i+1})$, with

$$T^{i,(2)}(s) = \bar{u}_{1,1}^{(2)}(s) + \sigma_1\bar{u}_1^{(1)}(s) - \delta_{t_i}^1[\sigma_1]^2\bar{S}_s^1 - \delta_{t_i}^{0,1}\left(\bar{C}_{1,1}^{0,1,(2)}(s) + \sigma_1\bar{C}_1^{0,1,(1)}(s)\right).$$

Then, using (2.3), we obtain

$$\begin{aligned}T^{i,(2)}(t_i) =& \bar{u}_{1,1}^{(2)}(t_i) + \sigma_1\bar{u}_1^{(1)}(t_i) - \left(\frac{\bar{u}_1^{(1)}(t_i)}{\sigma_1\bar{S}_{t_i}^1} - \frac{\delta_{t_i}^{0,1}\bar{C}_1^{0,1,(1)}(t_i)}{\sigma_1\bar{S}_{t_i}^1}\right)[\sigma_1]^2\bar{S}_{t_i}^1\\ &- \delta_{t_i}^{0,1}\left(\bar{C}_{1,1}^{0,1,(2)}(t_i) + \sigma_1\bar{C}_1^{0,1,(1)}(t_i)\right)\\ =& \bar{u}_{1,1}^{(2)}(t_i) - \delta_{t_i}^{0,1}\bar{C}_{1,1}^{0,1,(2)}(t_i).\end{aligned}$$

Using (2.2), it leads to $T^{i,(2)}(t_i) = 0$: this is the Gamma-hedging condition. We apply

once again Itô's rule and Lemma 2 to obtain

$$T^{i,(2)}(s) = \int_{t_i}^{s} T^{i,(3)}(r)\mathrm{d}W_r, \tag{2.6}$$

where $T^{i,(3)}(r)$ can be explicitly written using $\bar{u}^{(3)}_{1,1,1}$, $\bar{u}^{(2)}_{1,1}$, $\bar{u}^{(1)}_{1}$ and the similar Greeks for $C^{0,1}$. For the purpose of the current discussion, the exact espression of $T^{i,(3)}(t)$ is not important (it is given in the general case in Theorem 3). More interesting is to identify the form of the tracking error as a triple iterated stochastic integral w.r.t. W. Indeed, from (2.4), (2.5) and (2.6), one obtains

$$\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi) = -\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \int_{t_i}^{s} T^{i,(3)}(r)\mathrm{d}W_r\mathrm{d}W_s\mathrm{d}W_t. \tag{2.7}$$

The generalization of such a decomposition to the multidimensional asset model will be given in Theorem 3.

Now, let us formally comment the consequences of (2.7). At first sight, if $T^{i,(3)}(r)$ had a $\mathbb{Q}$-$\mathbf{L}_2$ moment uniformly bounded in r, we would deduce that

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi)|^2 &\leq \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \int_{t_i}^{s} \sup_{0\leq r\leq T} |T^{i,(3)}(r)|^2_{\mathbf{L}_2}\mathrm{d}r\mathrm{d}s\mathrm{d}t \\ &\leq \sup_{0\leq i\leq N-1,\ 0\leq r\leq T} |T^{i,(3)}(r)|^2_{\mathbf{L}_2}|\pi|^2\frac{T}{6}, \end{aligned} \tag{2.8}$$

i.e. the tracking error has a $\mathbb{Q}$-$\mathbf{L}_2$ norm of the order of the time step $|\pi|$. In particular, N regular trading dates would lead to an optimal rate of convergence. Actually, this argumentation is not complete because the moments of $T^{i,(3)}(r)$ may explode as r goes to T (because for non smooth payoff, the Greeks may go to $\pm\infty$ as the time to maturity shrinks to 0). It illustrates that a significant piece of work in the Delta-Gamma error analysis lies in the understanding of the behavior of $|T^{i,(3)}(r)|^2_{\mathbf{L}_2}$ as $r \to T$. This feature is tightly estimated through the fractional regularity of g (see Proposition 4 and Corollary 5).

2.1.2 The multidimensional case

Following the previous decomposition for a single asset, we can handle the general case as well. The discounted terminal value of the hedging portfolio is

$$\begin{aligned}\bar{V}_T^N &:= \bar{u}(0,S_0) + \sum_{i=0}^{N-1}\sum_{j=1}^{d}\delta_{t_i}^j(\bar{S}_{t_{i+1}}^j - \bar{S}_{t_i}^j) \\ &+ \sum_{i=0}^{N-1}\sum_{0\le j<k\le d}\delta_{t_i}^{j,k}(\bar{C}^{j,k}(t_{i+1},S_{t_{i+1}}^j,S_{t_{i+1}}^k) - \bar{C}^{j,k}(t_i,S_{t_i}^j,S_{t_i}^k)).\end{aligned} \tag{2.9}$$

Definition 1 (Delta-Gamma hedging strategy) *We define the DGHS by*

$$\delta_{t_i}^{j,k} := \frac{\bar{u}_{j,k}^{(2)}(t_i)}{\bar{C}_{j,k}^{j,k,(2)}(t_i)} \quad (1 \le j < k \le d, \text{ Exchange options}), \tag{2.10}$$

(2.11)

$$\delta_{t_i}^{0,l} := \frac{\bar{u}_{l,l}^{(2)}(t_i)}{\bar{C}_{l,l}^{0,l,(2)}(t_i)} - \frac{1}{\bar{C}_{l,l}^{0,l,(2)}(t_i)}\sum_{1\le j<k\le d}\delta_{t_i}^{j,k}\bar{C}_{l,l}^{j,k,(2)}(t_i) \quad (1 \le l \le d, \text{ Call options}),$$

$$\delta_{t_i}^{l} := \frac{1}{\sigma_l\bar{S}_{t_i}^l}\bar{u}_l^{(1)}(t_i) - \frac{1}{\sigma_l\bar{S}_{t_i}^l}\sum_{0\le j<k\le d}\delta_{t_i}^{j,k}\bar{C}_l^{j,k,(1)}(t_i) \quad (1 \le l \le d, \text{ assets}). \tag{2.12}$$

These are natural extensions of (2.2) and (2.3). Then, similarly, the discounted tracking error is defined by

$$\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi) := \bar{V}_T^N - e^{-rT}g(S_T).$$

Using expression (2.9) and Lemma 2, one has

$$\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi) = -\sum_{i=0}^{N-1}\sum_{l=1}^{d}\int_{t_i}^{t_{i+1}} T_l^{i,(1)}(t)\mathrm{d}W_t^l, \tag{2.13}$$

where

$$T_l^{i,(1)}(t) = \bar{u}_l^{(1)}(t) - \delta_{t_i}^l\sigma_l\bar{S}_t^l - \sum_{0\le j<k\le d}\delta_{t_i}^{j,k}\bar{C}_l^{j,k,(1)}(t) \tag{2.14}$$

for $t \in [t_i, t_{i+1})$.

In the general case, the (discounted) tracking error can be decomposed as a triple stochastic integral w.r.t. W, analogously to (2.7) in dimension 1. In the decomposition below, there are a principal part (related to the third derivative of the option price) and some residual terms.

Theorem 3 *Assume that* $\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0} < \infty$ *for some* $p_0 > 1$. *We have*

$$\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi) = -\sum_{i=0}^{N-1}\sum_{l,m,n=1}^{d}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\int_{t_i}^{s}\left(\bar{u}_{l,m,n}^{(3)}(r) + R_{l,m,n}^{i,(3)}(r)\right)\mathrm{d}W_r^n\mathrm{d}W_s^m\mathrm{d}W_t^l, \tag{2.15}$$

where

$$\begin{aligned}
R_{l,m,n}^{i,(3)}(t) &= \left(\sigma_m\bar{u}_{l,m}^{(2)}(t)\mathbb{1}_{n=m} + \sigma_l\bar{u}_{l,m}^{(2)}(t)\mathbb{1}_{n=l} + \sigma_l\bar{u}_{l,n}^{(2)}(t)\mathbb{1}_{m=l}\right)\\
&\quad + \sigma_l^2\left(\bar{u}_l^{(1)}(t) - \frac{\bar{S}_t^l}{\bar{S}_{t_i}^l}\bar{u}_l^{(1)}(t_i)\right)\mathbb{1}_{n=m=l}\\
&\quad - \sum_{0\le j<k\le d}\delta_{t_i}^{j,k}\bar{C}_{l,m,n}^{j,k,(3)}(t)\\
&\quad - \sum_{0\le j<k\le d}\delta_{t_i}^{j,k}\left(\sigma_m\bar{C}_{l,m}^{j,k,(2)}(t)\mathbb{1}_{n=m} + \sigma_l\bar{C}_{l,m}^{j,k,(2)}(t)\mathbb{1}_{n=l} + \sigma_l\bar{C}_{l,n}^{j,k,(2)}(t)\mathbb{1}_{m=l}\right)\\
&\quad - \sigma_l^2\sum_{0\le j<k\le d}\delta_{t_i}^{j,k}\left(\bar{C}_l^{j,k,(1)}(t) - \frac{\bar{S}_t^l}{\bar{S}_{t_i}^l}\bar{C}_l^{j,k,(1)}(t_i)\right)\mathbb{1}_{n=m=l}.
\end{aligned} \tag{2.16}$$

In addition, the following estimate holds:

$$\mathbb{E}_{\mathbb{P}}\left|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi)\right|^2 \le C\sum_{i=0}^{N-1}\sum_{l,m,n=1}^{d}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\int_{t_i}^{s}\mathbb{E}_{\mathbb{P}}\left|\bar{u}_{l,m,n}^{(3)}(r) + R_{l,m,n}^{i,(3)}(r)\right|^2\mathrm{d}r\mathrm{d}s\mathrm{d}t. \tag{2.17}$$

The proof is postponed to Section 4. We now bring in the next proposition several estimates, as key ingredients of our main results. These estimates are proved in Section 4.

Proposition 4 *Assume that* $\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0} < \infty$ *for some* $p_0 > 1$.

- *For $l, m, n = 1...d$ and $0 \leq t < T$,*

$$\mathbb{E}_{\mathbb{P}} \left| \bar{u}^{(3)}_{l,m,n}(t) \right|^2 \leq C \frac{V_{t,T}(g)}{(T-t)^3}. \tag{2.18}$$

- *There exists a positive $\pi^{\text{threshold}}$ (depending on p_0, T and T_2) such that, if $|\pi| \leq \pi^{\text{threshold}}$, then, for $0 \leq t_i \leq t < t_{i+1} \leq T$,*

$$\mathbb{E}_{\mathbb{P}} \left| R^{i,(3)}_{l,m,n}(t) \right|^2 \leq C \frac{V_{t,T}(g) + \left(\mathbb{E}_{\mathbb{P}} \left| g(S_T) \right|^{2p_0} \right)^{\frac{1}{p_0}}}{(T-t)^2}. \tag{2.19}$$

Thus, an easy and direct consequence of Proposition 4 and the definition (1.2) of $K^{\alpha,p_0}(g)$ is the following corollary.

Corollary 5 *Let $\alpha \in (0,1]$. If $g \in \mathbf{L}_{2,\alpha}$ and $\mathbb{E}_{\mathbb{P}} \left| g(S_T) \right|^{2p_0} < \infty$ for some $p_0 > 1$. Then, when $|\pi| \leq \pi^{\text{threshold}}$, for $0 \leq t < T$,*

$$\mathbb{E}_{\mathbb{P}} \left| \bar{u}^{(3)}_{l,m,n}(t) + R^{i,(3)}_{l,m,n}(t) \right|^2 \leq C \frac{K^{\alpha,p_0}(g)}{(T-t)^{3-\alpha}}. \tag{2.20}$$

Now we are in a position to expose our main results.

Theorem 6 *Let $\alpha \in (0,1]$. Assume $g \in \mathbf{L}_{2,\alpha}$ and $\mathbb{E}_{\mathbb{P}} \left| g(S_T) \right|^{2p_0} < \infty$ for some $p_0 > 1$.*

(1) Regular grid $\pi^{(1)}$. *For N sufficiently large to ensure $|\pi^{(1)}| = \frac{T}{N} \leq \pi^{\text{threshold}}$, one has*

$$\left(\mathbb{E}_{\mathbb{P}} \left| \overline{\mathcal{E}}^{\Delta\Gamma}_N(g, \pi^{(1)}) \right|^2 \right)^{1/2} \leq C \frac{(K^{\alpha,p_0}(g))^{1/2} \; T^{\alpha/2}}{N^{\alpha/2}}. \tag{2.21}$$

(2) Non regular grid $\pi^{(\beta)}$, $\beta \in (0,1)$. *Taking N sufficiently large to ensure $|\pi^{(\beta)}| \leq$*

$\pi^{\text{threshold}}$, *one has*

$$\left(\mathbb{E}_{\mathbb{P}}\left|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi^{(\beta)})\right|^2\right)^{1/2} \leq \begin{cases} C\dfrac{(K^{\alpha,p_0}(g))^{1/2}\ T^{\alpha/2}}{N^{\frac{\alpha}{2\beta}}} \ \textit{if}\ \beta \in (\dfrac{\alpha}{2},1), \\ C\dfrac{(K^{\alpha,p_0}(g))^{1/2}\ T^{\alpha/2}}{N}\sqrt{\log(1+N)} \ \textit{if}\ \beta = \dfrac{\alpha}{2}, \\ C\dfrac{(K^{\alpha,p_0}(g))^{1/2}\ T^{\alpha/2}}{N} \ \textit{if}\ \beta \in (0,\dfrac{\alpha}{2}). \end{cases}$$

Before proving these results, we give few comments.

(1) The use of irregular grids crucially helps to increase the convergence rate of the tracking error, for any fractional regularity $\alpha \in (0,1]$, up to the rate N.

(2) For the regular grid, in general the estimate (2.21) of the rate of convergence $N^{\alpha/2}$ is tight for $\alpha < 1$ (because the estimate (2.20) is tight, see the discussions in Geiss and Hujo (2007) and Gobet and Makhlouf (2008)). In addition, it coincides with the rate of convergence of the Delta hedging strategy: in other words, for non-smooth payoffs, the DGHS used with regular grids does not improve the rate of convergence of the tracking error. This is confirmed by our numerical results in Section 3.

(3) In the case $\alpha = 1$, our estimate (2.21) is not optimal. Indeed, for Call options, from Brodén and Wiktorsson (2008) we expect the rate to be equal to $N^{3/4}$. For smoother payoffs (say C_b^3, leading also to $\alpha = 1$), the first three spatial derivatives of u are bounded and our computations lead to a convergence rate equal to N (see the estimates (2.8)).

Hence, the case $\alpha = 1$ is a singular point, for which the convergence rate may go from $N^{1/2}$ to N. Tight estimates on $(\mathbb{E}_{\mathbb{P}}\left|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi^{(1)})\right|^2)^{1/2}$ can not be given only by assuming $g \in \mathbf{L}_{2,1}$: presumably, an extra condition on ∇g would be necessary. We leave this issue for further investigations.

Proof of Theorem 6. Put $T^{i,(3)}_{l,m,n}(r) = \bar{u}^{(3)}_{l,m,n}(r) + R^{i,(3)}_{l,m,n}(r)$ and define $f : \tau \mapsto \int_{t_i}^{\tau}\int_{t_i}^{t}\int_{t_i}^{s} \mathbb{E}_{\mathbb{P}}\left|T^{i,(3)}_{l,m,n}(r)\right|^2 \mathrm{d}r\mathrm{d}s\mathrm{d}t$. The Taylor expansion of order 2, with integral remainder term, of the function $f(.)$ between $\tau = t_i$ and $\tau = t_{i+1}$ gives

$$\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\int_{t_i}^{s} \mathbb{E}_{\mathbb{P}}\left|T^{i,(3)}_{l,m,n}(r)\right|^2 \mathrm{d}r\mathrm{d}s\mathrm{d}t = \frac{1}{2}\int_{t_i}^{t_{i+1}} (t_{i+1}-t)^2\, \mathbb{E}_{\mathbb{P}}\left|T^{i,(3)}_{l,m,n}(t)\right|^2 \mathrm{d}t.$$

Then, from Theorem 3 and Corollary 5, one obtains

$$\mathbb{E}_{\mathbb{P}}\left|\overline{\mathcal{E}}^{\Delta\Gamma}_N(g,\pi^{(\beta)})\right|^2 \leq C\sum_{i=0}^{N-1}\int_{t_i^{(\beta)}}^{t_{i+1}^{(\beta)}} \left(t_{i+1}^{(\beta)}-t\right)^2 \frac{K^{\alpha,p_0}(g)}{(T-t)^{3-\alpha}}\mathrm{d}t. \tag{2.22}$$

The term $i = N-1$ in the above summation is equal to

$$C\int_{T-t_{N-1}^{(N,\beta)}}^{T} \frac{K^{\alpha,p_0}(g)}{(T-t)^{1-\alpha}}\mathrm{d}t = \frac{CK^{\alpha,p_0}(g)}{\alpha}\frac{T^\alpha}{N^{\alpha/\beta}}, \tag{2.23}$$

where we have used $T - t_{N-1}^{(N,\beta)} = \frac{T}{N^{1/\beta}}$. Furthermore, it is known (see Gobet and Makhlouf (2008)) that

$$\sup_{i=0\ldots N-1}\sup_{t\in[t_i^{(\beta)},t_{i+1}^{(\beta)})}\left(\frac{t_{i+1}^{(\beta)}-t}{(T-t)^{1-\beta}}\right) \leq \frac{T^\beta}{\beta N}.$$

From this, (2.22) and (2.23), it follows that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left|\overline{\mathcal{E}}^{\Delta\Gamma}_N(g,\pi^{(\beta)})\right|^2 &\leq \frac{CK^{\alpha,p_0}(g)}{\alpha}\frac{T^\alpha}{N^{\alpha/\beta}} \\ &\quad + CK^{\alpha,p_0}(g)\Big(\frac{T^\beta}{\beta N}\Big)^2\int_0^{t_{N-1}^{(N,\beta)}}(T-t)^{2-2\beta}\frac{\mathrm{d}t}{(T-t)^{3-\alpha}}.\end{aligned} \tag{2.24}$$

To complete the proof, note that it remains to upper bound the above integral in the three cases $\beta \in (\alpha/2, 1]$, $\beta = \alpha/2$ and $\beta \in (0, \alpha/2)$. Denote by I the second term in the r.h.s. of (2.24).

(1) If $\beta > \alpha/2$, the function $f : t \mapsto \frac{(T-t)^{2-2\beta}}{(T-t)^{3-\alpha}}$ is not integrable at $t = T$ and we easily derive

$$I \leq CK^{\alpha,p_0}(g)\Big(\frac{T^\beta}{\beta N}\Big)^2\frac{T^{\alpha-2\beta}}{(2\beta-\alpha)N^{(\alpha-2\beta)/\beta}} = CK^{\alpha,p_0}(g)\frac{T^\alpha}{\beta^2(2\beta-\alpha)}\frac{1}{N^{\frac{\alpha}{\beta}}}.$$

(2) If $\beta = \alpha/2$, the function f is still not integrable at $t = T$ and we obtain

$$I \leq CK^{\alpha,p_0}(g)\Big(\frac{T^\beta}{\beta N}\Big)^2 \log(N^{1/\beta}).$$

(3) If $\beta < \alpha/2$, the function f is integrable and it gives

$$I \leq CK^{\alpha,p_0}(g)\Big(\frac{T^\beta}{\beta N}\Big)^2 \frac{T^{\alpha-2\beta}}{(\alpha-2\beta)} = CK^{\alpha,p_0}(g)\frac{T^\alpha}{\beta^2(\alpha-2\beta)}\frac{1}{N^2}.$$

□

3 Numerical Results

In this section, we present some experiments of Delta and Delta-Gamma hedging strategies, of both a European Call option ($g(S) = (S-K)_+$) and a European digital Call option ($g(S) = \mathbb{1}_{S\geq K}$) on a single asset ($d = 1$). The two payoff functions belong respectively to $\mathbf{L}_{2,1}$ and $\mathbf{L}_{2,\frac{1}{2}}$. In our experiments, we rely on the known explicit formulas for the prices of these two options. In the following tests, we take $r = 2\%$, $\mu = 1\%$, $\sigma = 25\%$, $s_0^1 = 100$, $T = 1$ year and $K = 100$. The additional hedging instrument is a Call with strike $K_{0,1} = 100$ and maturity $T_2 = 1.25$ year. All the experiments are carried out with 10000 simulations under the historical probability.

Comparison of the order of $\mathbf{L}_2$-convergence. For the figures 1 and 2, the simulations are made with $N = 12, 25, 50, 100, 200, 400, 800$ rebalancing dates, following different β's, i.e. different time nets. Note that, in practice, $N = 12$, $N = 25$ and $N = 50$ correspond approximately to a monthly, fortnightly and weekly rebalancing respectively.

Log-log plots. These figures are the log-log plots of the second moment of the track-

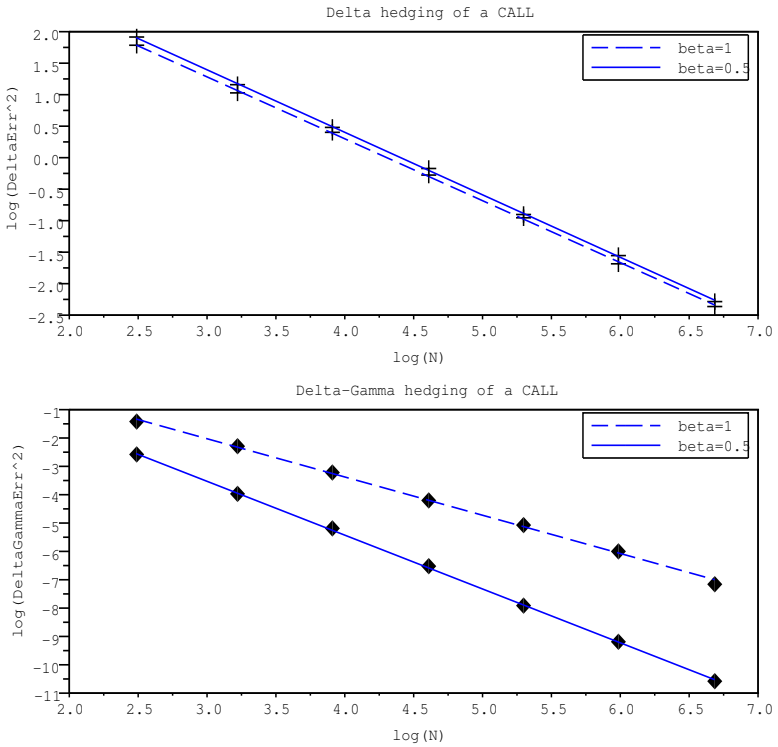

Fig. 1. For a Call: at the top (Delta hedging strategy), $\log(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N^{\Delta}(g,\pi^{(\beta)})|^2)$ vs $\log(N)$. At the bottom (Delta-Gamma hedging strategy), $\log(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi^{(\beta)})|^2)$ vs $\log(N)$.

	Delta hedging		Delta-Gamma hedging	
Beta	1.00	0.50	1.00	0.50
Expected order	0.50 see Zhang (1999)	0.50 see Geiss (2002)	0.75 see Brodén and Wiktorsson (2008)	1.00 (up to a log factor) see our Theorem 6
Empirical order	0.49	0.49	0.68	0.95

Table 1

For a Call: order of convergence of the $\mathbf{L}_2$-norm of tracking errors $(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N(g,\pi^{(\beta)})|^2)^{1/2}$.

ing error versus the number of rebalancing dates (i.e. $\log(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N(g,\pi^{(\beta)})|^2)$ versus $\log(N)$). We observe that they give straight lines, suggesting that not only upper

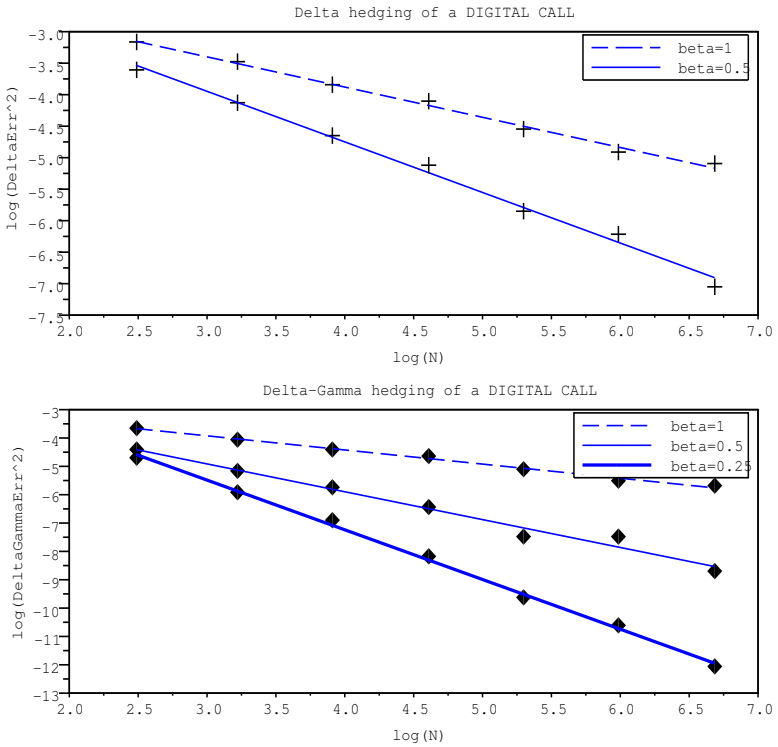

Fig. 2. For a digital Call: at the top (Delta hedging strategy), $\log(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N^{\Delta}(g,\pi^{(\beta)})|^2)$ vs $\log(N)$. At the bottom (Delta-Gamma hedging strategy), $\log(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi^{(\beta)})|^2)$ vs $\log(N)$.

bounds are available but presumably expansion results are also valid in these cases. The resulting slopes of the log-log plots represent twice the convergence order of $(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N(g,\pi^{(\beta)})|^2)^{1/2}$: these empirical convergence orders are reported in Table 1 for the Call and in Table 2 for the digital Call.

Theoretical order of convergence. In these tabulars, we also indicate the convergence order that one can expect from theoretical estimates, together with the reference.

Empirical order of convergence. The high convergence orders are not accurately estimated, possibly because of the relatively large statistical error due to simulations. Generally speaking, we notice that the rate of the Delta-Gamma tracking error is better than that of the Delta tracking error. However, when $\beta = 1$ (equidistant time net), there may be no significative difference between the two rates: consider for in-

	Delta hedging		Delta-Gamma hedging		
Beta	1.00	0.50	1.00	0.50	0.25
Expected order	0.25 see Gobet and Temam (2001)	0.50 (up to a log factor) see Geiss (2002)	0.25 see our Theorem 6	0.50 see our Theorem 6	1.00 (up to a log factor) see our Theorem 6
Empirical order	0.24	0.40	0.25	0.49	0.88

Table 2

For a Digital Call: order of convergence of the $\mathbf{L}_2$-norm of tracking errors $(\mathbb{E}_{\mathbb{P}}|\overline{\mathcal{E}}_N(g,\pi^{(\beta)})|^2)^{1/2}$.

stance the digital Call (see Table 2). This shows the advantage of hedging at non equidistant rebalancing dates when the fractional regularity index α of the payoff function is smaller than 1.

Furthermore, the smaller β, the better the rate of convergence of the tracking error, either for the hedging of the Call or for that of the digital Call. However, one should not take β too small since the time net points become too close to each other near the maturity, which might cause numerical instabilities in the simulations or might lead to unrealistic trading dates. From Tables 1 and 2, observe that one obtains approximately the right convergence order for the $\mathbf{L}_2$-norm of tracking errors (order 1/2 with Delta hedging and order 1 with Delta-Gamma hedging) by taking the critical threshold for β, i.e. α for Delta hedging and $\frac{\alpha}{2}$ for Delta-Gamma hedging.

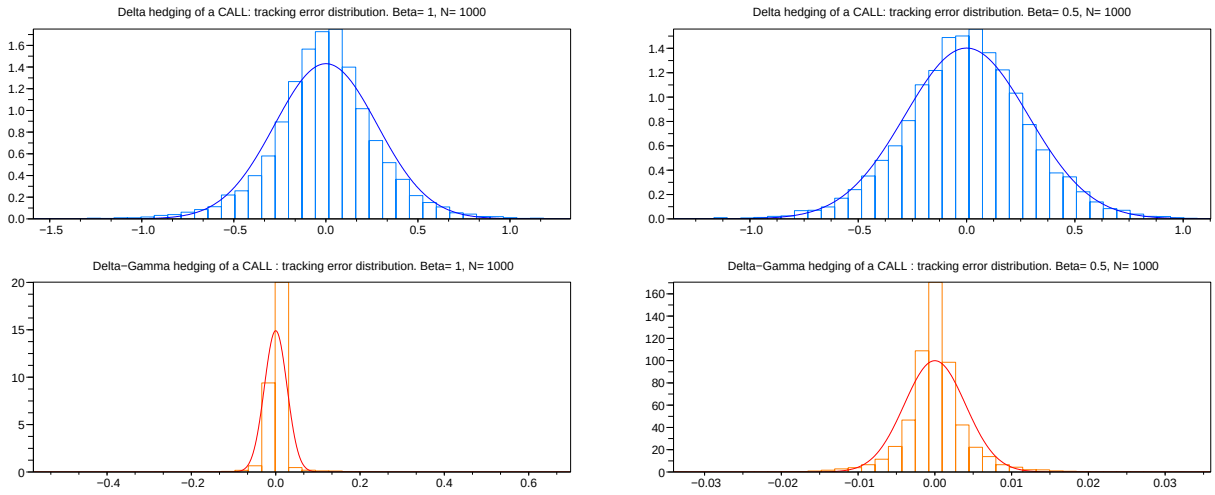

Fig. 3. Distributions of the tracking errors for a Call

Comparison of convergence in distribution. Figures 3 and 4 show, for each option and for different β's, the histogram of the tracking error, superposed to the Gaussian density with mean zero and variance equal to the observed empirical variance of the error. It is obtained for $N = 1000$ trading dates. This comparison is aimed at checking two features:

(1) is the asymptotic distribution Gaussian? the answer is generally no, as it can be easily observed in the figures. Additionally, it is known (cf Gobet and Temam (2001), Hayashi and Mykland (2005)) that the limiting distribution is not Gaussian but mixed Gaussian. Hence, we should consider the Gaussian distribution with the right variance as a benchmark and not as the true distribution: this representation helps to answer the next item.

(2) do the convergence rates in $\mathbf{L}_2$ and in distribution differ? In introduction, we mention that it can happen.

When $\beta > \frac{\alpha}{2}$ with Delta-Gamma hedging and $\beta > \alpha$ with Delta hedging, one notices from Figures 3 and 4 that the empirical distribution of the tracking error and the related Gaussian distribution seem to be not in the same scale, which corroborates the fact that the convergence in $\mathbf{L}_2$ and in distribution hold at different rates.

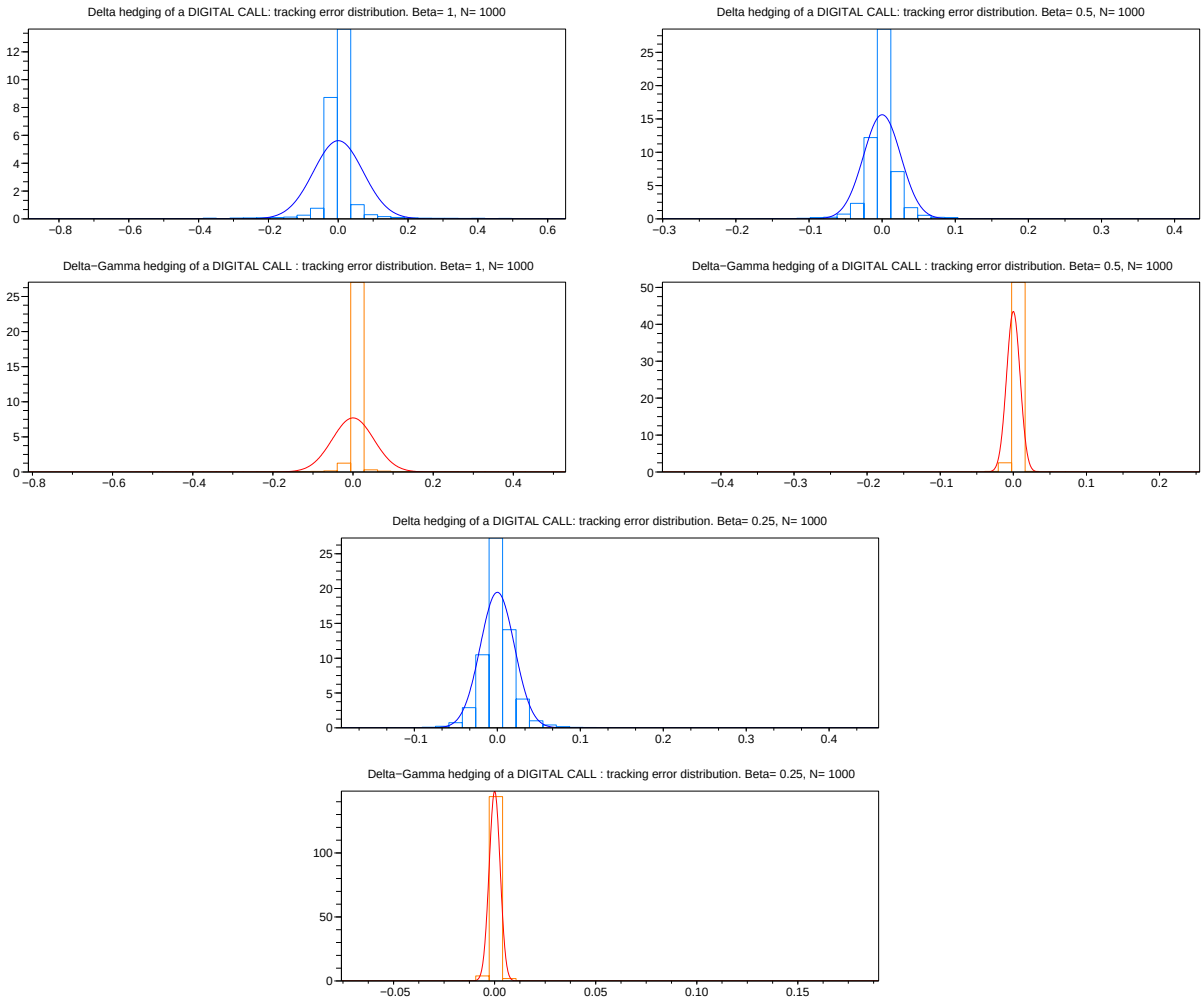

Fig. 4. Distributions of the tracking errors for a Digital Call

4 Proofs

4.1 Proof of Theorem 3.

Decomposition (2.15). It is clear that the definition of $\delta_{t_i}^l$ given by (2.12) ensures that $T_l^{i,(1)}(t_i) = 0$ for any $l = 1...d$, where $T_l^{i,(1)}(t)$ is defined by (2.14). Thus $T_l^{i,(1)}(t) = \int_{t_i}^t \mathrm{d}T_l^{i,(1)}(s)$, $t \in [t_i, t_{i+1})$. Applying Itô's rule and using Lemma 2, one obtains

$$\mathrm{d}T_l^{i,(1)}(t) = \sum_{m=1}^{d} T_{l,m}^{i,(2)}(t)\mathrm{d}W_t^m,$$

with

$$T^{i,(2)}_{l,m}(t) = \bar{u}^{(2)}_{l,m}(t) + \sigma_l \bar{u}^{(1)}_l(t)\mathbb{1}_{m=l} - \delta^l_{t_i}\sigma_l^2 \bar{S}^l_t \mathbb{1}_{m=l} \\ - \sum_{0\le j<k\le d} \delta^{j,k}_{t_i}\left(\bar{C}^{j,k,(2)}_{l,m}(t) + \sigma_l \bar{C}^{j,k,(1)}_l(t)\mathbb{1}_{m=l}\right).$$

Now, as in the one-dimensional case (see paragraph 2.1.1), we can check that the definitions of $\delta^{j,k}_{t_i}$ given by (2.10) (resp. (2.11)) ensure that $T^{i,(2)}_{l,m}(t_i) = 0$ for $1 \le l \ne m \le d$ (resp. $1 \le l = m \le d$) (in the previous cancellations, we strongly use that the price function $C^{j,k}$ depends on S only through the variables S^j and S^k). We apply once again Itô's rule and Lemma 2 to obtain

$$\mathrm{d}T^{i,(2)}_{l,m}(t) = \sum_{n=1}^{d} T^{i,(3)}_{l,m,n}(t)\mathrm{d}W^n_t$$

for $t \in [t_i, t_{i+1})$, with

$$\begin{aligned} T^{i,(3)}_{l,m,n}(t) &= \bar{u}^{(3)}_{l,m,n}(t) + \sigma_m \bar{u}^{(2)}_{l,m}(t)\mathbb{1}_{n=m} + \sigma_l \bar{u}^{(2)}_{l,m}(t)\mathbb{1}_{n=l} \\ &\quad + \left(\sigma_l \bar{u}^{(2)}_{l,n}(t) + \sigma_l^2 \bar{u}^{(1)}_l(t)\mathbb{1}_{n=l}\right)\mathbb{1}_{m=l} \\ &\quad - \delta^l_{t_i}\sigma_l^3 \bar{S}^l_t \mathbb{1}_{n=m=l} \\ &\quad - \sum_{0\le j<k\le d} \delta^{j,k}_{t_i}\left(\bar{C}^{j,k,(3)}_{l,m,n}(t) + \sigma_m \bar{C}^{j,k,(2)}_{l,m}(t)\mathbb{1}_{n=m} + \sigma_l \bar{C}^{j,k,(2)}_{l,m}(t)\mathbb{1}_{n=l}\right) \\ &\quad - \sum_{0\le j<k\le d} \delta^{j,k}_{t_i}\left(\sigma_l \bar{C}^{j,k,(2)}_{l,n}(t) + \sigma_l^2 \bar{C}^{j,k,(1)}_l(t)\mathbb{1}_{n=l}\right)\mathbb{1}_{m=l}. \end{aligned} \tag{4.1}$$

Substituting the expression (2.12) for $\delta^l_{t_i}$ in line (4.1) above, and rearranging the different terms, one readily obtains (2.16).

Estimate (2.17). The required estimate is computed under the historical probability $\mathbb{P}$. Note that an estimation of the $\mathbf{L}_2$-norm under the risk-neutral measure $\mathbb{Q}$ is straightforward using the Itô isometry.

Writing $\mathrm{d}W_t^l = \mathrm{d}\hat{W}_t^l + \lambda_l \mathrm{d}t$ in (2.13) yields (using the inequality $(a+b)^2 \leq 2(a^2+b^2)$)

$$\mathbb{E}_{\mathbb{P}}\left|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi)\right|^2 \leq 2\mathbb{E}_{\mathbb{P}}\left|\sum_{i=0}^{N-1}\sum_{l=1}^{d}\int_{t_i}^{t_{i+1}} T_l^{i,(1)}(t)\mathrm{d}\hat{W}_t^l\right|^2 + 2(\sup_{l=1\ldots d}\lambda_l^2)\mathbb{E}_{\mathbb{P}}\left|\sum_{i=0}^{N-1}\sum_{l=1}^{d}\int_{t_i}^{t_{i+1}}\left|T_l^{i,(1)}(t)\right|\mathrm{d}t\right|^2.$$

One can apply Itô's isometry to the first term ($(\hat{W}_t)_t$ being a $\mathbb{P}$-Brownian motion) and Cauchy-Schwartz inequality to the second to obtain

$$\mathbb{E}_{\mathbb{P}}\left|\overline{\mathcal{E}}_N^{\Delta\Gamma}(g,\pi)\right|^2 \leq \left(2 + 2Td(\sup_{l=1\ldots d}\lambda_l^2)\right)\sum_{i=0}^{N-1}\sum_{l=1}^{d}\int_{t_i}^{t_{i+1}}\mathbb{E}_{\mathbb{P}}\left|T_l^{i,(1)}(t)\right|^2\mathrm{d}t. \tag{4.2}$$

Now, since $T_l^{i,(1)}(t) = \sum_{m=1}^d \int_{t_i}^t T_{l,m}^{i,(2)}(s)\mathrm{d}W_s^m$ and $T_{l,m}^{i,(2)}(s) = \sum_{n=1}^d \int_{t_i}^s T_{l,m,n}^{i,(3)}(r)\mathrm{d}W_r^n$ with $T_{l,m,n}^{i,(3)}(r) = \bar{u}_{l,m,n}^{(3)}(r) + R_{l,m,n}^{i,(3)}(r)$, using the same arguments as above, we show that

$$\mathbb{E}_{\mathbb{P}}\left|T_l^{i,(1)}(t)\right|^2 \leq_c \sum_{m=1}^{d}\int_{t_i}^{t}\mathbb{E}_{\mathbb{P}}\left|T_{l,m}^{i,(2)}(s)\right|^2\mathrm{d}s,$$
$$\mathbb{E}_{\mathbb{P}}\left|T_{l,m}^{i,(2)}(s)\right|^2 \leq_c \sum_{n=1}^{d}\int_{t_i}^{s}\mathbb{E}_{\mathbb{P}}\left|T_{l,m,n}^{i,(3)}(r)\right|^2\mathrm{d}r.$$

Plugging these inequalities into (4.2) leads to (2.17). □

4.2 Proof of Proposition 4.

Estimate (2.18). It is part of the statement of the following lemma.

Lemma 7 *For $1 \leq l,m,n \leq d$ and $0 \leq t < T$,*

$$\mathbb{E}_{\mathbb{P}}\left|\bar{u}_l^{(1)}(t)\right|^2 \leq C\frac{V_{t,T}(g)}{(T-t)}, \tag{4.3}$$

$$\mathbb{E}_{\mathbb{P}}\left|\bar{u}_{l,m}^{(2)}(t)\right|^2 \leq C\frac{V_{t,T}(g)}{(T-t)^2}, \tag{4.4}$$

$$\mathbb{E}_{\mathbb{P}}\left|\bar{u}_{l,m,n}^{(3)}(t)\right|^2 \leq C\frac{V_{t,T}(g)}{(T-t)^3}. \tag{4.5}$$

Proof of Lemma 7. For $1 \leq l \leq d$, set $X_t^l := \ln S_t^l$ and $X_t := (X_t^1, ..., X_t^d)$: this defines the log-price process. We define the function $v(t, X)$ for $t \geq 0$ and $X = (X^1, ..., X^d) \in \mathbb{R}^d$ by $v(t, X) = u(t, e^{X^1}, ..., e^{X^d})$. Then $v(t, X_t) = u(t, S_t)$. For notation simplicity, $(e^{X^1}, ..., e^{X^d})$ will be denoted by e^X. For $1 \leq l, m, n \leq d$,

$$\begin{aligned}
\partial_l^1 v(t,X) &= e^{X^l}\partial_l^1 u(t,e^X),\\
\partial_{lm}^2 v(t,X) &= e^{X^l}e^{X^m}\partial_{lm}^2 u(t,e^X) + e^{X^l}\partial_l^1 u(t,e^X)\mathbb{1}_{m=l},\\
\partial_{lmn}^3 v(t,X) &= e^{X^l}e^{X^m}e^{X^n}\partial_{lmn}^3 u(t,e^X) + e^{X^l}e^{X^m}\partial_{lm}^2 u(t,e^X)(\mathbb{1}_{n=l}+\mathbb{1}_{n=m})\\
&\quad + e^{X^l}e^{X^n}\partial_{ln}^2 u(t,e^X)\mathbb{1}_{m=l} + e^{X^l}\partial_l^1 u(t,e^X)\mathbb{1}_{n=m=l}.
\end{aligned}$$

Previous relations yield

$$\begin{aligned}
e^{X^l}\partial_l^1 u(t,e^X) &= \partial_l^1 v(t,X),\\
e^{X^l}e^{X^m}\partial_{lm}^2 u(t,e^X) &= \partial_{lm}^2 v(t,X) - \partial_l^1 v(t,X)\mathbb{1}_{m=l},\\
e^{X^l}e^{X^m}e^{X^n}\partial_{lmn}^3 u(t,e^X) &= \partial_{lmn}^3 v(t,X) - \partial_{lm}^2 v(t,X)(\mathbb{1}_{n=l}+\mathbb{1}_{n=m})\\
&\quad - \partial_{ln}^2 v(t,X)\mathbb{1}_{m=l} + 2\partial_l^1 v(t,X)\mathbb{1}_{n=m=l}.
\end{aligned} \tag{4.6}$$

We only prove (4.5), the proof is exactly the same for (4.3) and (4.4).

From (4.6), we readily get

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left|\bar{u}_{l,m,n}^{(3)}(t)\right|^2 \leq 16 \sup_{1\leq j\leq d}|\sigma_j|^6 e^{-2rt}\Big(&\mathbb{E}_{\mathbb{P}}\left|\partial_{lmn}^3 v(t,X_t)\right|^2 + \mathbb{E}_{\mathbb{P}}\left|\partial_{lm}^2 v(t,X_t)\right|^2\\
&+ \mathbb{E}_{\mathbb{P}}\left|\partial_{ln}^2 v(t,X_t)\right|^2 + \mathbb{E}_{\mathbb{P}}\left|\partial_l^1 v(t,X_t)\right|^2\Big).
\end{aligned} \tag{4.7}$$

Now, write $v(t,x) = e^{-r(T-t)}\int_{\mathbb{R}^d} g\Big(\cdots, e^{x^l+(r-\frac{1}{2}\sigma_l^2)(T-t)+\sigma_l y^l}, \cdots\Big)p(t,x;T,y)\mathrm{d}y$, where $p(t,x;T,y)$ denotes the transition density function of the d-dimensional $\mathbb{Q}$-Brownian motion W. Thus we can explicitly differentiate this Gaussian density with respect to

any component of x, and show that

$$\begin{aligned}
\partial^1_l v(t,X_t) &= e^{-r(T-t)}\mathbb{E}^{\mathcal{F}_t}_{\mathbb{Q}}\left(g(S_T)H^{(1)}_{t,T}\right),\\
\partial^2_{lm} v(t,X_t) &= e^{-r(T-t)}\mathbb{E}^{\mathcal{F}_t}_{\mathbb{Q}}\left(g(S_T)H^{(2)}_{t,T}\right),\\
\partial^3_{lmn} v(t,X_t) &= e^{-r(T-t)}\mathbb{E}^{\mathcal{F}_t}_{\mathbb{Q}}\left(g(S_T)H^{(3)}_{t,T}\right)
\end{aligned}$$

(this is the usual representation of Greeks using Malliavin calculus weights, see Fournié et al.(1999)). In the one dimensional case ($d=1$), one has:

$$\begin{aligned}
H^{(1)}_{t,T} &= \frac{W_T - W_t}{T-t},\\
H^{(2)}_{t,T} &= \frac{(W_T - W_t)^2}{(T-t)^2} - \frac{1}{T-t},\\
H^{(3)}_{t,T} &= \frac{(W_T - W_t)^3}{(T-t)^3} - 3\frac{W_T - W_t}{T-t}.
\end{aligned}$$

In our more general setting, the random variables $H^{(1)}_{t,T}$, $H^{(2)}_{t,T}$ and $H^{(3)}_{t,T}$ are independent of $\mathcal{F}_t$, have zero mean and satisfy to the following estimates (for $q>1$):

$$\mathbb{E}^{\mathcal{F}_t}_{\mathbb{Q}}\left|H^{(i)}_{t,T}\right|^q \leq \frac{C_q}{(T-t)^{i\frac{q}{2}}}, \quad 1\leq i\leq 3. \tag{4.8}$$

To pass to $\mathbb{P}$-conditional expectations, we introduce $Z_t = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}|_{\mathcal{F}_t}$ the Radon-Nikodym density of $\mathbb{Q}$ w.r.t. $\mathbb{P}$ on $\mathcal{F}_t$. Therefore, using the zero-mean property of the weights $H^{(i)}_{t,T}$, one can write

$$\begin{aligned}
\partial^1_l v(t,X_t) &= e^{-r(T-t)}\mathbb{E}^{\mathcal{F}_t}_{\mathbb{Q}}\left[\left(g(S_T) - \mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}} g(S_T)\right)H^{(1)}_{t,T}\right]\\
&= e^{-r(T-t)}\mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}}\left[\left(g(S_T) - \mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}} g(S_T)\right)H^{(1)}_{t,T}\frac{Z_T}{Z_t}\right],\\
\partial^2_{lm} v(t,X_t) &= e^{-r(T-t)}\mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}}\left[\left(g(S_T) - \mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}} g(S_T)\right)H^{(2)}_{t,T}\frac{Z_T}{Z_t}\right],\\
\partial^3_{lmn} v(t,X_t) &= e^{-r(T-t)}\mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}}\left[\left(g(S_T) - \mathbb{E}^{\mathcal{F}_t}_{\mathbb{P}} g(S_T)\right)H^{(3)}_{t,T}\frac{Z_T}{Z_t}\right].
\end{aligned}$$

Thus, using (4.8) (together with the standard inequality $\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}\left|\frac{Z_T}{Z_t}\right|^q \leq C_q$, for $q > 1$) and applying Cauchy-Schwartz inequality, one readily obtains

$$\begin{aligned}\left|\partial_l^1 v(t,X_t)\right|^2 &\leq C\frac{\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}\left|g(S_T)-\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}g(S_T)\right|^2}{T-t},\\ \mathbb{E}_{\mathbb{P}}\left|\partial_l^1 v(t,X_t)\right|^2 &\leq C\frac{V_{t,T}(g)}{(T-t)}.\end{aligned} \tag{4.9}$$

Similarly,

$$\mathbb{E}_{\mathbb{P}}\left|\partial_{lm}^2 v(t,X_t)\right|^2 \leq C\frac{V_{t,T}(g)}{(T-t)^2}, \quad \mathbb{E}_{\mathbb{P}}\left|\partial_{lmn}^3 v(t,X_t)\right|^2 \leq C\frac{V_{t,T}(g)}{(T-t)^3}. \tag{4.10}$$

Then, (4.9), (4.10), together with (4.7), yield (4.5). □

Estimate (2.19). According to (2.16), one can write

$$R_{l,m,n}^{i,(3)}(t) = \bar{U}^{(2)}(t) + \bar{U}^{(1)}(t) - \sum_{0\leq j<k\leq d}\Big(\bar{U}^{C,j,k,(3)}(t) + \bar{U}^{C,j,k,(2)}(t) + \bar{U}^{C,j,k,(1)}(t)\Big),$$

with

$$\begin{aligned}\bar{U}^{(2)}(t) &:= \Big(\sigma_m \bar{u}_{l,m}^{(2)}(t)\mathbb{1}_{n=m} + \sigma_l \bar{u}_{l,m}^{(2)}(t)\mathbb{1}_{n=l} + \sigma_l \bar{u}_{l,n}^{(2)}(t)\mathbb{1}_{m=l}\Big),\\ \bar{U}^{(1)}(t) &:= \sigma_l^2\Big(\bar{u}_l^{(1)}(t) - \frac{\bar{S}_t^l}{\bar{S}_{t_i}^l}\bar{u}_l^{(1)}(t_i)\Big)\mathbb{1}_{n=m=l},\\ \bar{U}^{C,j,k,(3)}(t) &:= \delta_{t_i}^{j,k}\bar{C}_{l,m,n}^{j,k,(3)}(t),\\ \bar{U}^{C,j,k,(2)}(t) &:= \delta_{t_i}^{j,k}\Big(\sigma_m \bar{C}_{l,m}^{j,k,(2)}(t)\mathbb{1}_{n=m} + \sigma_l \bar{C}_{l,m}^{j,k,(2)}(t)\mathbb{1}_{n=l} + \sigma_l \bar{C}_{l,n}^{j,k,(2)}(t)\mathbb{1}_{m=l}\Big),\\ \bar{U}^{C,j,k,(1)}(t) &:= \sigma_l^2\delta_{t_i}^{j,k}\Big(\bar{C}_l^{j,k,(1)}(t) - \frac{\bar{S}_t^l}{\bar{S}_{t_i}^l}\bar{C}_l^{j,k,(1)}(t_i)\Big)\mathbb{1}_{n=m=l}.\end{aligned}$$

Then,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left|R_{l,m,n}^{i,(3)}(t)\right|^2 \leq_c &\left(\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{(2)}(t)\right|^2 + \mathbb{E}_{\mathbb{P}}\left|\bar{U}^{(1)}(t)\right|^2\right)\\ &+ \sup_{0\leq j<k\leq d}\left(\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(3)}(t)\right|^2 + \mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(2)}(t)\right|^2 + \mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(1)}(t)\right|^2\right).\end{aligned} \tag{4.11}$$

Let us upper bound each term appearing in (4.11). From Lemma 7, one has

$$\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{(2)}(t)\right|^2 \leq 3C(\sigma_l^2+\sigma_m^2)\frac{V_{t,T}(g)}{(T-t)^2}, \tag{4.12}$$

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{(1)}(t)\right|^2 &\leq 2\sigma_l^4\mathbb{E}_{\mathbb{P}}\left|\bar{u}_l^{(1)}(t)\right|^2 + 2\sigma_l^4\mathbb{E}_{\mathbb{P}}\Big(\left|\bar{u}_l^{(1)}(t_i)\right|^2\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_{t_i}}\left|\frac{\bar{S}_t^l}{\bar{S}_{t_i}^l}\right|^2\Big)\\ &\leq_c \mathbb{E}_{\mathbb{P}}\left|\bar{u}_n^{(1)}(t)\right|^2 + \mathbb{E}_{\mathbb{P}}\left|\bar{u}_n^{(1)}(t_i)\right|^2\\ &\leq_c \frac{V_{t,T}(g)}{(T-t)}, \qquad (4.13)\end{aligned}$$

where we have used the fact that $\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_{t_i}}\left|\frac{\bar{S}_t^n}{\bar{S}_{t_i}^n}\right|^2 = \mathbb{E}_{\mathbb{P}}\left|\frac{\bar{S}_t^n}{\bar{S}_{t_i}^n}\right|^2 \leq C$.

To handle the terms $\bar{U}^{C,j,k,(1)}(t)$, $\bar{U}^{C,j,k,(2)}(t)$, $\bar{U}^{C,j,k,(3)}(t)$ in (4.11), we need extra intermediate results, that we present as lemmas.

Lemma 8 *For $1 \leq l,m,n \leq d$ and $0 \leq t < T$, one has*

$$\mathbb{E}_{\mathbb{P}}\left|\bar{u}_l^{(1)}(t)\right|^{2p_0} \leq C\frac{\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0}}{(T-t)^{p_0}}, \qquad \mathbb{E}_{\mathbb{P}}\left|\bar{u}_{l,m}^{(2)}(t)\right|^{2p_0} \leq C\frac{\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0}}{(T-t)^{2p_0}}.$$

The proof of this lemma is very similar to that of Lemma 7. One has just to substitute the $\mathbf{L}_{2p_0}$-norm for the $\mathbf{L}_2$-norm and $\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0}$ for $V_{t,T}(g)$. We skip details.

Lemma 9 *For $0 \leq j < k \leq d$, $(l,m,n) \in \{(j,j,j),(j,j,k),(j,k,k),(k,k,k)\}$, $p > 1$ and $0 \leq t_i \leq t < t_{i+1} \leq T$, one has*

$$\mathbb{E}_{\mathbb{P}}\left|\frac{\bar{C}_{l,m}^{j,k,(2)}(t)}{\bar{C}_{l,n}^{j,k,(2)}(t)}\right|^p \leq C. \tag{4.14}$$

Proof of Lemma 9. Estimate (4.14) is a direct consequence from $(A.3)$, $(A.4)$, $(A.5)$, together with $\frac{1}{\sqrt{T_2-t}} \leq \frac{1}{\sqrt{T_2-T}}$ and the fact that

$$\mathbb{E}_{\mathbb{P}}\left|S_t^j\right|^q + \mathbb{E}_{\mathbb{P}}\left|S_t^j\right|^{-q} \leq C_q \tag{4.15}$$

for $0 \leq j \leq d$, $t \in [0,T]$ and $q > 0$. □

Lemma 10 *For $0 \leq j < k \leq d$, $(l,m,n) \in \{(j,j,j),(j,j,k),(j,k,k),(k,k,k)\}$ and $p > 1$, there exists a positive constant $\pi^{\text{threshold}}$ (which depends on p, T and T_2) such that, for any time net whose size $|\pi| \leq \pi^{\text{threshold}}$ and for any $0 \leq t_i \leq t < t_{i+1} \leq T$,*

$$\mathbb{E}_{\mathbb{P}} \left| \frac{\bar{C}^{j,k,(2)}_{l,m}(t)}{\bar{C}^{j,k,(2)}_{l,m}(t_i)} \right|^p \leq C, \tag{4.16}$$

$$\mathbb{E}_{\mathbb{P}} \left| \frac{\bar{C}^{j,k,(3)}_{l,m,n}(t)}{\bar{C}^{j,k,(2)}_{l,m}(t_i)} \right|^p \leq C. \tag{4.17}$$

Proof of Lemma 10. We only detail the proof of (4.16), the proof of (4.17) is similar. The proof is divided into two steps. We first establish intermediate controls, then we complete the proof.

Step 1. In the following, q is a real number greater than p. One has

$$\begin{aligned}
d_1^{j,k}(t,S_t) &= \frac{1}{\sigma_{j,k}\sqrt{T_2-t}} \ln\left(\frac{S_t^k}{K_{j,k}S_t^j}\right) + \frac{1}{2}\sigma_{j,k}\sqrt{T_2-t} \\
&= \frac{1}{\sigma_{j,k}\sqrt{T_2-t}} \left(\ln s_0^k + (\mu_k - \frac{1}{2}\sigma_k^2)t + \sigma_k \hat{W}_t^k\right) \\
&- \frac{1}{\sigma_{j,k}\sqrt{T_2-t}} \left(\ln K_{j,k} + \ln s_0^j + (\mu_j - \frac{1}{2}\sigma_j^2)t + \sigma_j \hat{W}_t^j\right) + \frac{1}{2}\sigma_{j,k}\sqrt{T_2-t} \\
&= \frac{1}{\sigma_{j,k}\sqrt{T_2-t}} \left(\sigma_{j,k}\hat{W}_t^{j,k} - m^{j,k}(t)\right),
\end{aligned}$$

where $\sigma_{j,k}$ is defined by (A.2), and

$$\begin{aligned}
\sigma_{j,k}\hat{W}_t^{j,k} &:= \sigma_k \hat{W}_t^k - \sigma_j \hat{W}_t^j, \\
m^{j,k}(t) &:= \frac{1}{\sigma_{j,k}\sqrt{T_2-t}} \left(\ln \frac{s_0^k}{s_0^j K_{j,k}} + \left(\mu_k - \mu_j + \frac{1}{2}\sigma_j^2 - \frac{1}{2}\sigma_k^2\right)t\right) + \frac{1}{2}\sigma_{j,k}\sqrt{T_2-t}.
\end{aligned}$$

Note that $(\hat{W}_t^{j,k})_t$ is a $\mathbb{P}$-Brownian motion. We will denote $e^{-\frac{q}{2}\left|d_1^{j,k}(t,S_t)\right|^2}$ by $\phi^{j,k,q}(t)$. Then,

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_{t_i}}\left(\phi^{j,k,q}(t)\right) &= \frac{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2-t}}{\sqrt{q}} \mathbb{E}_{\mathbb{P}}^{\mathcal{F}_{t_i}} \left(\frac{\sqrt{q}}{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2-t}} \exp - \frac{\left(\sigma_{j,k}\hat{W}_t^{j,k} - m^{j,k}(t)\right)^2}{2\left(\sigma_{j,k}\sqrt{T_2-t}\right)^2/q}\right) \\
&= \frac{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2-t}}{\sqrt{q}} (f_X * f_Y)(m^{j,k}(t) - \sigma_{j,k}\hat{W}_{t_i}^{j,k}),
\end{aligned}$$

where f_X and f_Y are the respective probability density functions of two independent random variables X and Y having normal distributions $\mathcal{N}\Big(0; \frac{\left(\sigma_{j,k}\sqrt{T_2-t}\right)^2}{q}\Big)$ and $\mathcal{N}\Big(0; \sigma_{j,k}^2(t-t_i)\Big)$. Therefore,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_{t_i}}\left(\phi^{j,k,q}(t)\right) &= \frac{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2-t}}{\sqrt{q}} f_{X+Y}(m^{j,k}(t) - \sigma_{j,k}\hat{W}_{t_i}^{j,k}) \\ &= c(t_i,t)\exp\frac{-q\left(\sigma_{j,k}\hat{W}_{t_i}^{j,k} - m^{j,k}(t)\right)^2}{2\left(\sigma_{j,k}^2(T_2-t) + q\sigma_{j,k}^2(t-t_i)\right)}\end{aligned}$$

with

$$c(t_i,t) := \frac{\sigma_{j,k}\sqrt{T_2-t}}{\sqrt{\sigma_{j,k}^2(T_2-t) + q\sigma_{j,k}^2(t-t_i)}},$$

which is uniformly bounded by 1. Then

$$\frac{\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_{t_i}}\left(\phi^{j,k,q}(t)\right)}{\phi^{j,k,q}(t_i)} = c(t_i,t)e^{\varphi(t_i,t,\hat{W}_{t_i}^{j,k})},$$

where, for $0 \le t_i \le t \le t_{i+1} \le T$ and $y \in \mathbb{R}$,

$$\begin{aligned}\varphi(t_i,t,y) &:= \frac{-q\left(\sigma_{j,k}y - m^{j,k}(t)\right)^2}{2\left(\sigma_{j,k}^2(T_2-t) + q\sigma_{j,k}^2(t-t_i)\right)} + \frac{q\left(\sigma_{j,k}y - m^{j,k}(t_i)\right)^2}{2\sigma_{j,k}^2(T_2-t)} \\ &= c_2(t_i,t)y^2 + c_1(t_i,t)y + c_0(t_i,t),\end{aligned}$$

and $c_2(t_i,t)$, $c_1(t_i,t)$, $c_0(t_i,t)$ are uniformly bounded and deterministic, with

$$c_2(t_i,t) := \frac{q}{2(T_2-t)} - \frac{q}{2\left(T_2 - t + q(t-t_i)\right)}.$$

Then, the quantity $\mathbb{E}_{\mathbb{P}}\left[\frac{\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_{t_i}}\left(\phi^{j,k,q}(t)\right)}{\phi^{j,k,q}(t_i)}\right]$ is finite (and uniformly bounded also) for all $t_i \le t \le t_{i+1} \le T$ if and only if

$$c_2(t_i,t) - \frac{1}{2t_i} < 0, \forall t_i \le t \le t_{i+1} \le T. \tag{4.18}$$

From the expression of $c_2(t_i,t)$ taken at $t = t_{i+1}$, we see that, with a mesh size $|\pi| \ge (t_{i+1} - t_i)$ large enough, the condition (4.18) may fail. A sufficient condition on

$|\pi|$ is $|\pi| \leq \frac{(T_2-T)^2}{Tq^2} := \pi^{\text{threshold}}$. Indeed, under this condition, we have

$$c_2(t_i,t) = \frac{2q^2(t-t_i)}{4(T_2-t)\left(T_2-t+q(t-t_i)\right)} < \frac{2q^2|\pi|}{4(T_2-T)^2} \leq \frac{1}{2T} \leq \frac{1}{2t_i}.$$

Step 2. Now, it is clear from $(A.3)$, $(A.4)$ and $(A.5)$, together with (4.15) that, for $(l,m) \in \{(j,j),(j,k),(k,k)\}$, one can apply Hölder's inequality to obtain

$$\mathbb{E}_{\mathbb{P}}\left|\frac{\bar{C}^{j,k,(2)}_{l,m}(t)}{\bar{C}^{j,k,(2)}_{l,m}(t_i)}\right|^p \leq C\left(\mathbb{E}_{\mathbb{P}}\left(\frac{\phi^{j,k,q}(t)}{\phi^{j,k,q}(t_i)}\right)\right)^{\frac{p}{q}} = C\left(\mathbb{E}_{\mathbb{P}}\left[\frac{\mathbb{E}^{\mathcal{F}_{t_i}}_{\mathbb{P}}\left(\phi^{j,k,q}(t)\right)}{\phi^{j,k,q}(t_i)}\right]\right)^{\frac{p}{q}}$$

(where $1<p<q$), which has just been shown to be finite under the condition above on the time net size. The lemma is proved. □

Completion of the upper bounds of (4.11). Let $0 \leq j < k \leq d$.

If $j \neq 0$, use (2.10) to obtain

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(3)}(t)\right|^2 &= \mathbb{E}_{\mathbb{P}}\left|\delta^{j,k}_{t_i}\bar{C}^{j,k,(3)}_{l,m,n}(t)\right|^2 = \mathbb{E}_{\mathbb{P}}\left|\bar{u}^{(2)}_{j,k}(t_i)\frac{\bar{C}^{j,k,(3)}_{l,m,n}(t)}{\bar{C}^{j,k,(2)}_{j,k}(t_i)}\right|^2 \\ &\leq \left(\mathbb{E}_{\mathbb{P}}\left|\bar{u}^{(2)}_{j,k}(t_i)\right|^{2p_0}\right)^{\frac{1}{p_0}}\left(\mathbb{E}_{\mathbb{P}}\left|\frac{\bar{C}^{j,k,(3)}_{l,m,n}(t)}{\bar{C}^{j,k,(2)}_{j,k}(t_i)}\right|^{2q}\right)^{\frac{1}{q}},\end{aligned}$$

where p_0 and q are conjugate real numbers. Thus, from Lemma 8 and Lemma 10, and for a time net π whose size is sufficiently small, we readily obtain

$$\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(3)}(t)\right|^2 \leq C\frac{\left(\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0}\right)^{\frac{1}{p_0}}}{(T-t)^2}.$$

If $j = 0$, by using (2.11) we derive

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,0,k,(3)}(t)\right|^2 = \mathbb{E}_{\mathbb{P}}\left|\delta_{t_i}^{0,k}\bar{C}_{l,m,n}^{0,k,(3)}(t)\right|^2 \\
&= \mathbb{E}_{\mathbb{P}}\left|\frac{\bar{C}_{l,m,n}^{0,k,(3)}(t)}{\bar{C}_{k,k}^{0,k,(2)}(t_i)}\bar{u}_{k,k}^{(2)}(t_i) - \frac{\bar{C}_{l,m,n}^{0,k,(3)}(t)}{\bar{C}_{k,k}^{0,k,(2)}(t_i)}\sum_{1\le j_1<j_2\le d}\frac{\bar{C}_{k,k}^{j_1,j_2,(2)}(t_i)}{\bar{C}_{j_1,j_2}^{j_1,j_2,(2)}(t_i)}\bar{u}_{j_1,j_2}^{(2)}(t_i)\right|^2 \\
&\le 2\left(\mathbb{E}_{\mathbb{P}}\left|\frac{\bar{C}_{l,m,n}^{0,k,(3)}(t)}{\bar{C}_{k,k}^{0,k,(2)}(t_i)}\right|^{2q}\right)^{\frac{1}{q}}\left(\mathbb{E}_{\mathbb{P}}\left|\bar{u}_{k,k}^{(2)}(t_i)\right|^{2p_0}\right)^{\frac{1}{p_0}} \qquad (4.19) \\
&+ 2d^2\left(\mathbb{E}_{\mathbb{P}}\left|\frac{\bar{C}_{l,m,n}^{0,k,(3)}(t)}{\bar{C}_{k,k}^{0,k,(2)}(t_i)}\right|^{2q_1}\right)^{\frac{1}{q_1}} \qquad (4.20) \\
&\times \sup_{1\le j_1<j_2\le d}\left\{\left(\mathbb{E}_{\mathbb{P}}\left|\frac{\bar{C}_{k,k}^{j_1,j_2,(2)}(t_i)}{\bar{C}_{j_1,j_2}^{j_1,j_2,(2)}(t_i)}\right|^{2q_2}\right)^{\frac{1}{q_2}}\left(\mathbb{E}_{\mathbb{P}}\left|\bar{u}_{j_1,j_2}^{(2)}(t_i)\right|^{2p_0}\right)^{\frac{1}{p_0}}\right\}, \qquad (4.21)
\end{aligned}$$

where p_0, q, q_1 and q_2 are positive real numbers such that $\frac{1}{p_0}+\frac{1}{q} = \frac{1}{p_0}+\frac{1}{q_1}+\frac{1}{q_2} = 1$. Now, we apply Lemma 10 and Lemma 8 for term (4.19), Lemma 10 for term (4.20) and Lemmas 9 and 8 for term (4.21) to conclude that, for $|\pi|$ small enough, $\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,0,k,(3)}(t)\right|^2$ has the same bound as $\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(3)}(t)\right|^2$ with $j \neq 0$. Then, for every $0 \le j < k \le d$,

$$\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(3)}(t)\right|^2 \le C\frac{\left(\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0}\right)^{\frac{1}{p_0}}}{(T-t)^2}. \qquad (4.22)$$

The same arguments yield, for every $0 \le j < k \le d$,

$$\mathbb{E}_{\mathbb{P}}\left|\bar{U}^{C,j,k,(2)}(t)\right|^2 \le C\frac{\left(\mathbb{E}_{\mathbb{P}}\left|g(S_T)\right|^{2p_0}\right)^{\frac{1}{p_0}}}{(T-t)^2}. \qquad (4.23)$$

To handle $\bar{U}^{C,j,k,(1)}(t)$, we need to proceed a bit differently. We take advantage of the

linear SDE satisfied by $\bar{U}^{C,j,k,(1)}$. Indeed, using (1.11), one has

$$\begin{aligned}
\mathrm{d}\bar{U}^{C,j,k,(1)}(t) &= \sigma_l^2 \delta_{t_i}^{j,k} \mathbb{1}_{n=m=l} \left(\mathrm{d}\bar{C}_l^{j,k,(1)}(t) - \frac{\bar{C}_l^{j,k,(1)}(t_i)}{\bar{S}_{t_i}^l} \mathrm{d}\bar{S}_t^l \right) \\
&= \sigma_l^2 \delta_{t_i}^{j,k} \mathbb{1}_{n=m=l} \sum_{m\prime=1}^{d} \left(\bar{C}_{l,m\prime}^{j,k,(2)}(t) + \sigma_l \bar{C}_l^{j,k,(1)}(t) \mathbb{1}_{m\prime=l} \right) \mathrm{d}W_t^{m\prime} \\
&- \sigma_l^2 \delta_{t_i}^{j,k} \mathbb{1}_{n=m=l} \frac{\bar{C}_l^{j,k,(1)}(t_i)}{\bar{S}_{t_i}^l} \sigma_l \bar{S}_t^l \mathrm{d}W_t^l \\
&= \mathrm{d}\bar{V}^{C,j,k,(2)}(t) + \sigma_l \bar{U}^{C,j,k,(1)}(t) \mathrm{d}W_t^l,
\end{aligned}$$

where

$$\bar{V}^{C,j,k,(2)}(t) := \sigma_l^2 \delta_{t_i}^{j,k} \mathbb{1}_{n=m=l} \sum_{m\prime=1}^{d} \int_{t_i}^{t} \bar{C}_{l,m\prime}^{j,k,(2)}(s) \mathrm{d}W_s^{m\prime}.$$

Since $\bar{U}^{C,j,k,(1)}(t_i) = 0$, we obtain

$$\mathbb{E}_{\mathbb{P}} \left| \bar{U}^{C,j,k,(1)}(t) \right|^2 \leq 2\mathbb{E}_{\mathbb{P}} \left| \bar{V}^{C,j,k,(2)}(t) \right|^2 + 2\sigma_l^2 \int_{t_i}^{t} \mathbb{E}_{\mathbb{P}} \left| \bar{U}^{C,j,k,(1)}(s) \right|^2 \mathrm{d}s. \tag{4.24}$$

Following the ideas used to establish (2.17), we prove that

$$\mathbb{E}_{\mathbb{P}} \left| \bar{V}^{C,j,k,(2)}(t) \right|^2 \leq_c \sum_{m\prime=1}^{d} \int_{t_i}^{t} \mathbb{E}_{\mathbb{P}} \left| \delta_{t_i}^{j,k} \bar{C}_{l,m\prime}^{j,k,(2)}(s) \right|^2 \mathrm{d}s.$$

Then, the same arguments used for the proof of (4.22) and (4.23) yield

$$\mathbb{E}_{\mathbb{P}} \left| \bar{V}^{C,j,k,(2)}(t) \right|^2 \leq_c \int_{t_i}^{t} \frac{\left(\mathbb{E}_{\mathbb{P}} \left| g(S_T) \right|^{2p_0} \right)^{\frac{1}{p_0}}}{(T-s)^2} \mathrm{d}s \leq_c \frac{\left(\mathbb{E}_{\mathbb{P}} \left| g(S_T) \right|^{2p_0} \right)^{\frac{1}{p_0}}}{(T-t)^2} |\pi|.$$

Therefore, by invoking Gronwall's lemma in (4.24), we get

$$\mathbb{E}_{\mathbb{P}} \left| \bar{U}^{C,j,k,(1)}(t) \right|^2 \leq_c \frac{\left(\mathbb{E}_{\mathbb{P}} \left| g(S_T) \right|^{2p_0} \right)^{\frac{1}{p_0}}}{(T-t)^2} |\pi|.$$

Plugging this estimate, the inequalities (4.12), (4.13), (4.22) and (4.23) into (4.11), we complete the proof of (2.19) and that of Proposition 4. □

A Appendix

A.1 Margrabe formula

The following result is the so-called Margrabe formula (cf Margrabe (1978)): it is a generalization of the Black-Scholes formula, and gives an explicit expression for the Call and Exchange options.

Proposition 11 (The Margrabe Formula) *For $t \in [0, T_2)$ and positive initial spot values S^j, S^k, one has*

$$C^{j,k}(t, S^j, S^k) = S^k \mathcal{N}(d_1^{j,k}(t, S)) - K_{j,k} S^j \mathcal{N}(d_2^{j,k}(t, S)), \tag{A.1}$$

where $\mathcal{N}(.)$ is the cumulative distribution function of the standard Gaussian distribution, and

$$\begin{aligned} d_1^{j,k}(t, S) &:= \frac{1}{\sigma_{j,k}\sqrt{T_2 - t}} \ln\left(\frac{S^k}{K_{j,k} S^j}\right) + \frac{1}{2}\sigma_{j,k}\sqrt{T_2 - t}, \\ d_2^{j,k}(t, S) &:= \frac{1}{\sigma_{j,k}\sqrt{T_2 - t}} \ln\left(\frac{S^k}{K_{j,k} S^j}\right) - \frac{1}{2}\sigma_{j,k}\sqrt{T_2 - t}, \\ \sigma_{j,k} &:= \sqrt{\sigma_j^2 + \sigma_k^2 - 2\rho_{j,k}\sigma_j\sigma_k}. \end{aligned} \tag{A.2}$$

We note that $|d_2^{j,k}(t, S)|^2 = |d_1^{j,k}(t, S)|^2 - 2\ln\left(\frac{S^k}{K_{j,k}S^j}\right)$, or equivalently, $e^{-\frac{\left|d_2^{j,k}(t,S)\right|^2}{2}} = \frac{S^k}{K_{j,k}S^j} e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}$. Then, straightforward computations give the following derivatives expressions for $C^{j,k}$.

a) For $0 \leq j < k \leq d$,

$$
\begin{aligned}
\partial^1_k C^{j,k}(t, S^j, S^k) &= \mathcal{N}(d_1^{j,k}(t,S)); \\
\partial^2_{kk} C^{j,k}(t, S^j, S^k) &= \frac{1}{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2 - t}} \frac{1}{S^k} e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}; \qquad \text{(A.3)} \\
\partial^3_{kkk} C^{j,k}(t, S^j, S^k) &= \frac{1}{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2 - t}} \frac{1}{|S^k|^2} \left(-1 - \frac{d_1^{j,k}(t,S)}{\sigma_{j,k}\sqrt{T_2 - t}} \right) e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}.
\end{aligned}
$$

b) For $1 \leq j < k \leq d$,

$$
\begin{aligned}
\partial^1_j C^{j,k}(t, S^j, S^k) &= -K_{j,k}\mathcal{N}(d_2^{j,k}(t,S)); \\
\partial^2_{jj} C^{j,k}(t, S^j, S^k) &= \frac{1}{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2 - t}} \frac{S^k}{|S^j|^2} e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}; \qquad \text{(A.4)} \\
\partial^2_{jk} C^{j,k}(t, S^j, S^k) &= -\frac{1}{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2 - t}} \frac{1}{S^j} e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}; \qquad \text{(A.5)} \\
\partial^3_{jjk} C^{j,k}(t, S^j, S^k) &= -\frac{1}{\sqrt{2\pi}\sigma^2_{j,k}(T_2 - t)} \frac{d_2^{j,k}(t,S)}{|S^j|^2} e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}; \\
\partial^3_{jkk} C^{j,k}(t, S^j, S^k) &= \frac{1}{\sqrt{2\pi}\sigma^2_{j,k}(T_2 - t)} \frac{d_1^{j,k}(t,S)}{S^j S^k} e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}; \\
\partial^3_{jjj} C^{j,k}(t, S^j, S^k) &= \frac{1}{\sqrt{2\pi}\sigma_{j,k}\sqrt{T_2 - t}} \frac{S^k}{|S^j|^3} \left(-1 + \frac{d_2^{j,k}(t,S)}{\sigma_{j,k}\sqrt{T_2 - t}} \right) e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}.
\end{aligned}
$$

The interest for writing the above derivatives according only to $e^{-\frac{\left|d_1^{j,k}(t,S)\right|^2}{2}}$ (and not $e^{-\frac{\left|d_2^{j,k}(t,S)\right|^2}{2}}$) is to clearly see that in the ratio of two derivatives of order 2 or 3, the exponential terms cancel. For instance, $\frac{\partial^3_{jjk} C^{j,k}(t,S^j,S^k)}{\partial^2_{jk} C^{j,k}(t,S^j,S^k)} = \frac{1}{\sigma_{j,k}\sqrt{T_2 - t}} \frac{d_2^{j,k}(t,S)}{S^j}$. These features are crucial in the convergence analysis (see the proof of Proposition 4).

A.2 Proof of Lemma 1

• Suppose that g is Lipschitz (including the case of Call/Put payoffs). Then observe that

$$V_{t,T}(g) = \mathbb{E}_{\mathbb{P}}|g(S_T) - \mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}(g(S_T))|^2 \leq \mathbb{E}_{\mathbb{P}}|g(S_T) - g(S_t)|^2$$
$$\leq C\mathbb{E}_{\mathbb{P}}|S_T - S_t|^2 \leq C(T-t)$$

proving that $g \in \mathbf{L}_{2,\alpha}$ with $\alpha = 1$. The first item of Lemma 1 is proved.

• Now, consider $g(S) = \mathbb{1}_{S \geq K}$ and a single log-normal asset S. Then, passing to the log variables, we set

$$v(t,x) = \mathbb{P}(\log(S_T) \geq \log(K) | \log(S_t) = x)$$
$$= \mathcal{N}(\frac{1}{\sigma\sqrt{T-t}} \log(\frac{e^x}{Ke^{-\mu(T-t)}}) - \frac{1}{2}\sigma\sqrt{T-t}).$$

An application of Itô formula to the $\mathbb{P}$-martingale $[v(s, \log(S_s)) = \mathbb{P}^{\mathcal{F}_s}(S_T \geq K)]_{0 \leq s \leq T}$ leads to

$$V_{t,T}(g) = \int_t^T \mathbb{E}_{\mathbb{P}}\Big(\frac{1}{2\pi(T-s)} e^{-(\frac{1}{\sigma\sqrt{T-s}} \log(\frac{S_s}{Ke^{-\mu(T-s)}}) - \frac{1}{2}\sigma\sqrt{T-s})^2}\Big) \mathrm{d}s.$$

Then, standard computations give $V_{t,T}(g) \leq \int_t^T \frac{C}{\sqrt{T-s}} \mathrm{d}s \leq C\sqrt{T-t}$, which shows that $g \in \mathbf{L}_{2,\alpha}$ with $\alpha = 1/2$.

• The stability by summation is obvious. Regarding the stability by product, write

$$V_{t,T}(g) = \mathbb{E}_{\mathbb{P}}|g_1(S_T)g_2(S_T) - \mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}(g_1(S_T)g_2(S_T))|^2$$
$$\leq \mathbb{E}_{\mathbb{P}}|g_1(S_T)g_2(S_T) \pm g_1(S_T)\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}(g_2(S_T)) - \mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}(g_1(S_T))\mathbb{E}_{\mathbb{P}}^{\mathcal{F}_t}(g_2(S_T))|^2$$
$$\leq 2|g_1|_\infty^2 V_{t,T}(g_2) + 2|g_2|_\infty^2 V_{t,T}(g_1).$$

Our statement readily follows from this. □.

References

[BW08] M. Brodén and M. Wiktorsson. On the convergence of higher order hedging schemes. *Preprint*, 2008.

[FLL+99] E. Fournié, J.M. Lasry, J. Lebuchoux, P.L. Lions, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3(4):391–412, 1999.

[Gei02] S. Geiss. Quantitative approximation of certain stochastic integrals. *Stoch. Stoch. Rep.*, 73(3-4):241–270, 2002.

[GH07] S. Geiss and M. Hujo. Interpolation and approximation in $L_2(\gamma)$. *Journal of Approximation Theory*, 144:213–232, 2007.

[GM05] E. Gobet and R. Munos. Sensitivity analysis using Itô-Malliavin calculus and martingales, and application to stochastic optimal control. *SIAM Journal on control and optimization*, 43(5):1676–1713, 2005.

[GM08] E. Gobet and A. Makhlouf. L2-time regularity of BSDEs with irregular terminal functions. *In revision for Stochastic Processes and their Applications, available at http://hal.archives-ouvertes.fr/hal-00291768/fr/*, 2008.

[GT01] E. Gobet and E. Temam. Discrete time hedging errors for options with irregular pay-offs. *Finance and Stochastics*, 5(3):357–367, 2001.

[GT08] S. Geiss and A. Toivola. Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces. *Preprint, arXiv:0711.1439v2*, 2008.

[HM05] T. Hayashi and P. Mykland. Evaluating hedging errors: an asymptotic approach. *Mathematical Finance*, 15(2):309–343, 2005.

[Hul09] J.C. Hull. *Options, futures, and other derivatives. 7th ed.* Prentice-Hall International Editions. Upper Saddle River, NJ: Prentice Hall., 2009.

[KJPS98] N. El Karoui, M. Jeanblanc-Picqué, and S.E. Shreve. Robustness of the Black and Scholes formula. *Math. Finance*, 8(2):93–126, 1998.

[KS97] Y.M. Kabanov and M. Safarian. On Leland's strategy of option pricing with transaction costs. *Finance and Stochastics*, 1(3):239–250, 1997.

[KS98] I. Karatzas and S.E. Shreve. *Methods of mathematical finance*. Springer-Verlag, New York, 1998.

[Mar78] W. Margrabe. The value of an option to exchange one asset for another. *Journal of Finance*, 33(1):177–186, 1978.

[Mar99] C. Martini. Propagation of convexity by Markovian and martingalian semigroups. *Potential Anal.*, 10(2):133–175, 1999.

[Per03] S. Pergamenshchikov. Limit theorem for Leland's strategy. *The Annals of Applied Probability*, 13:1099–1118, 2003.

[Zha99] R. Zhang. *Couverture approchée des options Européennes*. PhD thesis, Ecole Nationale des Ponts et Chaussées, http://cermics.enpc.fr/theses/99/zhang-ruotao.ps.gz, 1999.