PARTIAL HEDGING IN ROUGH VOLATILITY MODELS

Edouard Motte, Donatien Hainaut

LIDAM Discussion Paper ISBA
2023 / 26

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

Partial hedging in rough volatility models

Motte Edouard[*], Donatien Hainaut[†]

June 29, 2023

Abstract

This paper studies the problem of partial hedging within the framework of rough volatility models in an incomplete market setting. We employ a stochastic control problem formulation to minimize the discrepancy between a stochastic target and the terminal value of a hedging portfolio. As rough volatility models are neither Markovian nor semi-martingale, stochastic control problems associated to rough models are quite complex to solve. Therefore, we propose a multifactor approximation of the rough volatility model and introduce the associated Markov stochastic control problem. We establish the convergence of the optimal solution for the Markov partial hedging problem to the optimal solution of the original problem as the number of factors tends to infinity. Furthermore, the optimal solution of the Markov problem can be derived solving a Hamilton-Jacobi-Bellman (HJB) equation and more precisely a nonlinear partial differential equation (PDE). Due to the inherent complexity of this nonlinear PDE, an explicit formula for the optimal solution is generally unattainable. By introducing the dual solution of the Markov problem and expressing the primal solution as a function of the dual solution, we derive approximate solutions to the Markov problem using a dual control method. This method enables for sub-optimal choices of dual control to deduce lower and upper bounds on the optimal solution as well as sub-optimal hedging ratios. In particular, explicit formulas for partial hedging strategies in rough Heston model are derived.

KEYWORDS: *Partial hedging, rough volatility, rough Heston, stochastic control, Hamilton-Jacobi-Bellman, Markov approximation, dual control method.*

1 Introduction

Rough volatility models have gained significant popularity in quantitative finance since the pioneering work of Gatheral et al. [22]. These models incorporate long-range dependence, capturing important empirical stylized facts such as volatility clustering and roughness, which are often neglected in classical volatility models. In option pricing, rough volatility models generate implied volatility surfaces that are consistent with observed volatility surfaces, as shown in subsequent papers [7, 14, 20, 22, 28]. With only few parameters, they can effectively capture implied volatility smile as well as at-the-money skew overcoming stochastic volatility models that produce inconsistent volatility surfaces. In addition, we observe that the interest in rough processes extends to other domains such as insurance, whether in terms of their impact on pricing and insurance portfolios [11, 12] or on claims modelling [26].

In this paper, we investigate the problem of hedging in rough volatility models. While previous research [13, 21] have explored this matter in the context of complete market, where the volatility risk can be hedged either by trading forward variance curve or variance swap, our paper takes a different approach. We relax the complete market assumption and focus on an incomplete market, considering only underlying assets as hedging instruments. Since the market is incomplete, a perfect hedging strategy cannot be implemented, this is why we are interested in partial hedging strategies. Partial hedging strategies introduced by Föllmer and Leukert [15, 16], are powerful techniques for minimizing hedging losses at a fixed cost lower than the super-replication price. Their results have next been applied to various markets and various risk processes, we can mention among others [9, 10, 23, 29, 30, 33]. Notably, [29] extend the theory of partial hedging to stochastic volatility environment by formulating the problem as a stochastic control problem. However, the

[*] *LIDAM-ISBA, Université Catholique de Louvain, Voie du Roman Pays 20, Louvain-La-Neuve, 1348, Belgium (Email: edouard.motte@uclouvain.be)*

[†] *LIDAM-ISBA, Université Catholique de Louvain, Voie du Roman Pays 20, Louvain-La-Neuve, 1348, Belgium (Email: donatien.hainaut@uclouvain.be)*

problem of partial hedging in rough volatility models has not yet been investigated, so this article aims to fill this gap.

To this end, we introduce a stochastic control problem under rough volatility models. While the literature has studied stochastic control problems in rough volatility models primarily focusing on portfolio optimization, see [3, 4, 6, 18, 19, 24, 32], problems involving hedging or stochastic targets have received less attention. The non-Markovian nature of rough volatility processes poses significant challenges in solving these control problems. In portfolio optimization problem, [4, 18, 19] propose first order approximate solutions by relying on martingale distortion transformation of the value function while [6, 32] employ a Markov approximation of the rough volatility models. It is the latter technique that is developed in this paper to solve the control problem. Relying on several papers [1, 2, 5, 8, 26], we introduce a Markov multifactor approximation of rough volatility models based on the representation of the kernel function in terms of a Laplace transform. Then, we consider the Markov control problem associated with the approximate volatility model and show, with the help of convergence results stated in [2], that instead of solving the initial non-Markovian problem, we can solve the Markovian problem with negligible error.

The introduced Markov stochastic control problem is similar to a stochastic control problem associated to a partial hedging problem in multivariate stochastic volatility. Previous studies [17, 29, 31] have shown that the optimal value function for such problems satisfies a nonlinear partial differential equation that cannot be completely linearized, even by switching to the dual formulation of the problem. There are mainly two techniques developed in the literature to overcome this nonlinearity issue. [17, 29] consider fast-mean reverting volatility models to propose asymptotic solutions while [31] consider a dual control method to provide approximate solutions of the optimal solution. In this paper, we adopt a similar approach to the dual control method introduced by [31] to propose approximate solutions of the Markov problem for sub-optimal choices of dual control. Our approach has several advantages: it works with general classes of volatility models, gives lower and upper bounds to the optimal solution and allows to deduce convergence results toward the optimal solution.

The paper is outlined as follows. First, in Section 2, the mathematical framework is presented. We introduce the class of rough volatility models studied and we formulate the partial hedging problem. Next, in Sections 3 and 4, we discuss the multifactor approximation of the rough volatility model and introduce the associated Markov stochastic control problem. Moreover, we demonstrate the convergence of the optimal solution of the Markov problem to the optimal solution of the original problem. Then, in Section 5, we solve the Markov problem by introducing the Hamilton-Jacobi-Bellman (HJB) equation and deduce that the optimal solution satisfies a nonlinear PDE. Consequently, by expressing the primal solution in terms of the dual solution, we derive approximate solutions using a dual control method. Notably, we provide explicit formulas for sub-optimal partial hedging strategies in the rough Heston model. Finally, in Section 6, we conclude the paper by presenting a numerical application that focuses on the partial hedging of linear and vanilla options within the rough Heston model.

2 Statement of the problem

Consider a finite horizon $T > 0$ and a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ where $\mathbb{P}$ stands for the real measure and the filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ denotes all information known over time. Assume an arbitrage-free financial market in which we have a cash-account and a risky asset denoted respectively by $(S_t^0)_{0\leq t\leq T}$ and $(S_t)_{0\leq t\leq T}$. We suppose that those processes have the following dynamics

$$dS_t^0 = r\, S_t^0 dt$$

and

$$dS_t = \mu_t S_t dt + \sqrt{\nu_t} S_t dW_S(t)$$

where r is the risk-free interest rate, W_S is a standard brownian motion, $\mu_t := r + A\nu_t$, $A \in \mathbb{R}$ and $(\nu_t)_{0\leq t\leq T}$ a rough volatility process. The rough volatility satisfies a stochastic Volterra equation of the form

$$\nu_t = \nu_0 + \int_0^t G(t-s)b(\nu_s)ds + \int_0^t G(t-s)\sigma(\nu_s)dW_v(s) \tag{1}$$

where $\nu_0 \in \mathbb{R}_+$, $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ are Lipschitz continuous, W_v is a standard brownian motion such that $d < W_v(t), W_s(t) >= \rho\, dt$ and G a kernel assumed to be completely monotone. The rough volatility model (1) is a general model that incorporates well known rough models such for example the rough Heston model introduced in [13] with volatility process given by

$$\nu_t = \nu_0 + \int_0^t G(t-s)\,\kappa(\theta - \nu_s)ds + \int_0^t G(t-s)\,\zeta\sqrt{\nu_s}dW_v(s).$$

In the following, we consider the fractional kernel defined by

$$G(t) := \frac{t^{H-1/2}}{\Gamma(H+1/2)} \tag{2}$$

where H is the Hurst coeficient such that $H \in (0, 1/2)$ in order to consider rough volatility models. Note that using Corollary B.2. in [2], we can prove the existence of an unconstrained weak solution of the stochastic differention equation (1) when $\sigma(.)$ is continuous with linear growth and the fractional kernel satisfies (2). Moreover, the existence of a non-negative solution can be obtained by relying on Theorem B.4 in [2] if $\sigma(0) = 0$.

In this financial market with rough volatility, we are concerned with the hedging of contingent claims of maturity $T > 0$ of the form

$$H_T = h(S_T),$$

with $h(.)$ a continous function. El Euch and Rosenbaum [13] already tackle the question of hedging in rough volatility environment. They prove that perfect hedging is possible in rough Heston model provided that the forward variance curve can be taken as hedging instrument. However, this assumption is quite strong, which is why we are interested in the question of hedging in a financial market with only underlying assets as hedging instruments. As the market is incomplete, we already know that perfect hedging is not more possible, but we can still stay on the safe size by super-hedging the contingent claims. However, super-hedging strategies lead generally to super-hedging prices that are too high to be considered in practice. In the case where the initial capital available is smaller than the super-hedging prices, we know that we are not hedged in 100% of the cases. We can nevertheless define hedging strategies that aim at minimizing a loss arising from the hedging operation. This type of hedging strategy is called partial hedging strategy and was introduced in [15, 16]. It is this kind of strategy that we consider in the following. In this perspective, we consider a self-financing hedging portfolio denoted by $(V_t)_{0\le t\le T}$ and defined by investment in the assets available in the market (cash-account and the risky asset). The amount invested at time $t \in [0, T]$ in the risky asset is denoted by ξ_t and the portfolio evolves according to the following SDE

$$\begin{aligned} dV_t &= r\,(V_t - \xi_t S_t)dt + \xi_t\, dS_t, \\ V_0 &= v. \end{aligned}$$

The hedging ratio process $(\xi_t)_{0\le t\le T}$ is admissible if $(\xi_t)_{0\le t\le T}$ is a progressively measurable process in regards to $\mathcal{F}_t$ such that $E(\int_0^T \xi_t^2 S_t^2 \nu_t dt) < +\infty$. Similarly, we can also define the Profit and Loss (P&L) at maturity T denoted by π_T and defined by

$$\pi_T := V_T - H_T.$$

Let $\mathcal{R}$ be the set of all progressively measurable processes $(\xi_t)_{0\le t\le T}$ valued in $\mathbb{R}$ such that $E(\int_0^T \xi_t^2 S_t^2 \nu_t dt) < +\infty$, since we already mentionned that perfect hedging is not possible in our market, we have that

$$\nexists (\xi_t)_{0\le t\le T} \in \mathcal{R}\ s.t.\ \pi_T = 0\ a.s.$$

We thus consider the partial hedging problem and define in this sense an optimal hedging strategy satisfying the following optimization problem

$$l(t, s, \nu, v) := \inf_{\xi_t \in \mathcal{R}} E\Bigg(L(h(S_T), V_T)|S_t = s, \nu_t = \nu, V_t = v\Bigg), \tag{3}$$

where $L(.)$ is a continuous proper convex loss function satisfying quadratic growth condition. As example of loss function, we have the power loss function of the form

$$L_{power}(x, y) := \frac{1}{p}(x - y)^p,\ p = 2n,\ n \in \mathbb{N}_0, \tag{4}$$

but other types of loss functions can be considered such as exponential or shortfall losses. At this stage, the main problem in solving the introduced stochastic control problem is that the rough volatility model (1) is neither Markovian nor a semi-martingale. Thus, the principle of dynamic programming cannot be applied to solve the stochastic control problem. To overcome this problem, and following the idea of [6], we will consider a Markovian approximation of our initial problem and then solve the Markovian problem using the principle of dynamic programming.

3 Markov approximation

The non-Markovian structure of the rough volatility prevents from directly solving the partial hedging problem using classical stochastic control techniques. However, as shown in [6] in the case of portfolio optimization, the problem can be solved with a small error by considering a Markov approximation of the volatility process. As shown in several papers [2, 5, 8, 26, 27], the starting point of the Markovian approximation is the representation of the kernel $G(t)$ in terms of a Laplace transform such that

$$G(t) = \int_0^{+\infty} e^{-tx}\lambda(dx),$$

where λ is a measure on $\mathbb{R}_+$. In the case of fractional kernel, we have that

$$\begin{aligned} G(t) =& \frac{t^{H-1/2}}{\Gamma(H+1/2)} \\ =& \underbrace{\frac{1}{\Gamma(H+1/2)\Gamma(1/2-H)}}_{:=C_H} \int_0^{+\infty} e^{-tx}x^{-H-1/2}dx \\ =& C_H \int_0^{+\infty} e^{-tx}x^{-H-1/2}dx, \end{aligned}$$

where for $x > 0$,

$$\mu(x) = C_H x^{-H-1/2}dx.$$

Then, we approximate the integral by a finite sum and consider the approximate kernel $\hat{G}$ defined by

$$\hat{G}(t) := \sum_{i=1}^{n} w_i e^{-tx_i} \approx G(t), \tag{5}$$

where $(w_i)_{i=1,\dots,n}$ are the weights and $(x_i)_{i=1,\dots,n}$ the mean reversion terms that should be appropriately defined, we discuss later on the choice of these parameters. In this way, we can approximate the rough volatility process by defining a new stochastic process denoted by $(\hat{\nu}_t)_{0\le t\le T}$ satisfying for $t \in [0,T]$,

$$\hat{\nu}_t = \nu_0 + \int_0^t \hat{G}(t-s)b(\hat{\nu}_s)ds + \int_0^t \hat{G}(t-s)\sigma(\hat{\nu}_s)dW_v(s). \tag{6}$$

The following proposition states that the stochastic Volterra equation (6) can reduced to a $n-$dimensional stochastic differential equation.

Proposition 1. *The solution of (6) is given by* $\hat{\nu}_t = \nu_0 + \sum_{i=1}^n w_i \nu_t^i$ *where* $(\boldsymbol{\nu_t})_{0\le t\le T} := \Big((\nu_t^1, \nu_t^2, ..., \nu_t^n)\Big)_{0\le t\le T}$ *is solution of the n-dimensional SDE defined by*

$$\begin{aligned} \nu_t^i &= -\int_0^t x_i\nu_s^i ds + \int_0^t b(\hat{\nu}_s)ds + \int_0^t \sigma(\hat{\nu}_s)dW_v(s),\ i = 1, ..., n, \\ \nu_0^i &= 0. \end{aligned} \tag{7}$$

Proof. We refer the reader to the proof of Proposition 2.1. in [5]. □

Based on the approximate volatility process (6) and its Markov representation induced by SDEs (7), we can define the Markovian approximation of the stochastic control problem introduced in (3). First, we consider the approximate processes $(S_t^n)_{0\leq t\leq T}$ for which its SDE can either be written in terms of the approximate volatility process $(\hat{\nu}_t)_{0\leq t\leq T}$ or in terms of its Markov representation. Thus, for $t\in[0,T]$, the dynamic of $(S_t^n)_{0\leq t\leq T}$ is given by

$$dS_t^n = \hat{\mu}_t S_t^n dt + \sqrt{\hat{\nu}_t} S_t^n dW_S(t), \tag{8}$$

but also by

$$dS_t^n = \hat{\mu}_t S_t^n dt + \sqrt{\nu_0 + \sum_{i=1}^n w_i \nu_t^i} S_t^n dW_S(t), \tag{9}$$

with $\hat{\mu}_t = r + A\hat{\nu}_t = r + A\left(\nu_0 + \sum_{i=1}^n w_i\nu_t^i\right)$. In the same way, denoting the approximate hedging process by $(\xi_t^n)_{0\leq t\leq T}$, we define the associated hedging portfolio $(V_t^n)_{0\leq t\leq T}$ satisfying the following SDE

$$\begin{aligned} dV_t^n &= r(V_t^n - \xi_t^n S_t^n)dt + \xi_t^n dS_t^n, \\ V_0^n &= v, \end{aligned} \tag{10}$$

where $(\xi_t^n)_{0\leq t\leq T}$ is admissible if $(\xi_t^n)_{0\leq t\leq T}$ is a progressively measurable process in regards to $\mathcal{F}_t$ such that $E(\int_0^T (\xi_t^n)^2 (S_t^n)^2 \hat{\nu}_t dt) < +\infty$. The approximate P&L is defined by

$$\pi_T^n := V_T^n - H_T^n, \tag{11}$$

and as the market is incomplete,

$$\nexists(\xi_t^n)_{0\leq t\leq T} \in \mathcal{R}_n \ s.t.\ \pi_T^n = 0 \ a.s.\ ,$$

where $\mathcal{R}_n$ the set of all progressively measurable processes $(\xi_t^n)_{0\leq t\leq T}$ with regards to $\mathcal{F}_t$ valued in $\mathbb{R}$ such that $E(\int_0^T (\xi_t^n)^2 (S_t^n)^2 \hat{\nu}_t dt) < +\infty$. Thus, we introduce the approximate partial hedging problem that can be written either by considering the dependence on the approximate volatility process $(\hat{\nu}_t)_{0\leq t\leq T}$ or the dependence on its Markovian representation. In fact, by considering the dynamic (8) of $(S_t^n)_{0\leq t\leq T}$ written in terms of $(\hat{\nu}_t)_{0\leq t\leq T}$, we define the stochastic control problem

$$l_n(t,s,\nu,v) := \inf_{\xi_t^n\in\mathcal{R}_n} E\bigg(L(h(S_T^n), V_T^n) | S_t^n = s, \hat{\nu}_t = \nu, V_t^n = v\bigg). \tag{12}$$

Considering now the dynamic (9) of $(S_t^n)_{0\leq t\leq T}$ written in terms of $(\boldsymbol{\nu_t})_{0\leq t\leq T}$, we have the Markovian stochastic control problem

$$l_n(t,s,\boldsymbol{\nu},v) := \inf_{\xi_t^n\in\mathcal{R}_n} E\bigg(L(h(S_T^n), V_T^n) | S_t^n = s, \boldsymbol{\nu_t} = \boldsymbol{\nu}, V_t^n = v\bigg) \tag{13}$$

such that

$$l_n(t,s,\nu,v) = l_n(t,s,\boldsymbol{\nu},v).$$

Thanks to the approximation of the volatility process (6) and its Markov representation, we obtain a Markovian framework in which we solve the stochastic control problem (13) using the principle of dynamic programming. Nevertheless, before solving the control problem, it is interesting to consider the question of convergence of the approximate solution $l_n(.)$ toward $l(.)$. Indeed, without proof of convergence, solving the approximate problem would be pointless, this is why we dedicate a section to this issue.

4 Convergence results

In this section, we prove a convergence between the approximate solution $l_n(.)$ and $l(.)$. The first step is to prove the almost sure convergence of the approximate volatility process $(\hat{\nu}_t)_{0\leq t\leq T}$ to $(\nu_t)_{0\leq t\leq T}$, to prove it we rely on [2]. First, under specific assumptions on the weights $(w_i)_{i=1,\ldots,n}$ and mean reversion terms $(x_i)_{i=1,\ldots,n}$, we can prove that $\hat{G}$ converges to G in L^2.

Assumption 2. Suppose that the weights $(w_i)_{i=1,...,n}$ and mean reversion terms $(x_i)_{i=1,...,n}$ are given by

$$w_i = \int_{\eta_{i-1}}^{\eta_i} \lambda(dx),\ x_i = \frac{1}{w_i}\int_{\eta_{i-1}}^{\eta_i} x\,\lambda(dx),\ i = 1, ..., n,$$

with $(\eta_i)_{i=1,...,n}$ auxiliary mean reversion terms such that

$$0 = \eta_0 \leq x_1 \leq \eta_1 \leq x_2 \leq \eta_2 < ... < x_n \leq \eta_n,$$

and as n goes to infinity

$$\eta_n \to +\infty,\ \sum_{i=1}^{n}\int_{\eta_{i-1}}^{\eta_i} (x_i - x)^2\,\lambda(dx) \to 0.$$

Remark. As shown in [2], Assumption 2 is satisfied if we consider auxiliary mean reversion terms $(\eta_i)_{i=1,...,n}$ of the form

$$\eta_i = i \times \frac{n^{-\frac{1}{5}}}{T}\left(\frac{\sqrt{10}(1-2H)}{5-2H}\right),\ i = 1, ..., n.$$

Therefore, without loss of generality, we consider this form of auxiliary mean reversion terms in the numerical results.

Proposition 3. *(Proposition 3.3 in [2]) Suppose that for all $n \geq 1$, $(w_i)_{i=1,...,n}$ and $(x_i)_{i=1,...,n}$ satisfy Assumption 2 and that $\hat{G}$ is defined by (5). Then $\hat{G}$ converges in $L^2[0,T]$ to G when n goes to infinity i.e.*

$$||\hat{G} - G||_{L^2} \to 0.$$

Proof. We refer to [2] for the proof. □

Based on this proposition, we now establish a convergence of the approximate volatility process $\hat{\nu}_t$ toward ν_t.

Theorem 4. *Assume that $n \geq 1$, $(w_i)_{i=1,...,n}$ and $(x_i)_{i=1,...,n}$ satisfy Assumption 2, $\hat{G}$ is defined by (5) and that there exists positive constant δ and C such that*

$$\sup_{n\geq 1}\left(\int_0^h |\hat{G}(s)|^2 ds + \int_0^{T-h} |\hat{G}(h+s) - \hat{G}(s)|^2 ds\right) \leq Ch^{2\delta},$$

for any $t, h \geq 0$ with $t + h \leq T$, then

$$\hat{\nu}_t \xrightarrow{a.s.} \nu_t$$

as n goes to infinity.

Proof. The proof is an immediate consequence of Theorem 3.6. in [2] for one dimension. We just need to check that

$$\int_0^T |G(s) - \hat{G}(s)|^2 ds \to 0,\ \hat{\nu}_0 \to \nu_0$$

as n goes to infinity. The first convergence is obtained by Proposition 3 since assumptions of this proposition are fullfilled, we know that $\hat{G}$ converges in $L^2[0,T]$ to G, therefore, we have that

$$\int_0^T |G(s) - \hat{G}(s)|^2 ds \to 0.$$

Similarly, the second convergence is direct since, by definition, we consider for all $n \geq 1$,

$$\hat{\nu}_0 = \nu_0.$$

As the assumption of Theorem 3.6 in [2] are valid for one dimension, we can conclude that $\hat{\nu}$ is tight for the uniform topology and any point ν is solution of the Voltera equation (1) and thus

$$\hat{\nu}_t \xrightarrow{\mathcal{L}} \nu_t,$$

as n goes to infinity. Moreover, by the Skorokhod representation theorem, as convergence in law implies almost sure convergence on a suitable probability space, we finally obtain that

$$\hat{\nu}_t \xrightarrow{a.s.} \nu_t,$$

as n goes to infinity. □

Now that we establish a convergence for the volatility process, we can go a step further and show that the approximate processes $(S_t^n)_{0\leq t\leq T}$ and $(V_t^n)_{0\leq t\leq T}$ converge almost surely respectively toward $(S_t)_{0\leq t\leq T}$ and $(V_t)_{0\leq t\leq T}$.

Proposition 5. *Assume that assumptions of Therorem 4 are fulfilled and consider the approximate processes $(S_t^n)_{0\leq t\leq T}$ and $(V_t^n)_{0\leq t\leq T}$ satisfying SDEs (8) and (10) such that $\forall t\in[0,T]$, $\xi_t^n=\xi_t$ a.s., then we have the following convergence results*

$$S_t^n \xrightarrow{a.s.} S_t,$$
$$V_t^n \xrightarrow{a.s.} V_t,$$

as n goes to infinity.

Proof. The proof is provided in Appendix. □

With Proposition 5, we have shown that approximate processes converge to the rough volatility dependent processes. Using these results, we are now able to show the convergence of the approximate solution to the solution of the control problem under rough volatility. To do so, inspired by [6], we define $l^{\xi}(t,s,\nu,v)$ and $l_n^{\xi^n}(t,s,\nu,v)$ by

$$\begin{aligned} l^{\xi}(t,s,\nu,v) &:= E_{t,s,\nu,v}\bigg(L(h(S_T),V_T(\xi))\bigg)\\ &:= E\bigg(L(h(S_T),V_T(\xi))|S_t=s,\nu_t=\nu,V_t=v\bigg),\\ l_n^{\xi^n}(t,s,\nu,v) &:= E_{t,s,\nu,v}\bigg(L(h(S_T^n),V_T^n(\xi^n))\bigg)\\ &:= E\bigg(L(h(S_T^n),V_T^n(\xi^n))|S_t^n=s,\hat{\nu}_t=\nu,V_t^n=v\bigg), \end{aligned}$$

such that

$$l(t,s,\nu,v)=\inf_{\xi_t\in\mathcal{R}} l^{\xi}(t,s,\nu,v),$$
$$l_n(t,s,\nu,v)=\inf_{\xi_t^n\in\mathcal{R}_n} l_n^{\xi^n}(t,s,\nu,v).$$

We first consider a lemma before before stating the convergence result we wish to achieve.

Lemma 6. *Assume that assumptions of Therorem 4 are fulfilled and fix admissible hedging strategies $(\xi_t)_{0\leq t\leq T}$, $(\xi_t^n)_{0\leq t\leq T}$ such that $\forall t\in[0,T]$, $\xi_t^n=\xi_t$ a.s., if the sequence $\bigg(L(h(S_T^n),V_T^n)\bigg)_{n\geq1}$ is uniformly integrable, then*

$$l_n^{\xi}(t,s,\nu,v)\to l^{\xi}(t,s,\nu,v),$$

as n goes to infinity.

Proof. Using Proposition 5, since we assume that $\forall t\in[0,T]$, $\xi_t^n=\xi_t$ *a.s.*, we have that

$$S_t^n \xrightarrow{a.s.} S_t,$$
$$V_t^n \xrightarrow{a.s.} V_t,$$

as the loss function $L(.)$ and $h(.)$ are continous, we deduce that

$$L(h(S_T^n),V_T^n) \xrightarrow{a.s.} L(h(S_T),V_T),$$

and the uniform integrability of $\bigg(L(h(S_T^n),V_T^n)\bigg)_{n\geq1}$ implies that

$$E_{t,s,\hat{\nu},v}\bigg(L(h(S_T^n),V_T^n)\bigg)\to E_{t,s,\nu,v}\bigg(L(h(S_T),V_T)\bigg).$$

Therefore, we have proved the statement

$$l_n^{\xi}(t,s,\nu,v)\to l^{\xi}(t,s,\nu,v).$$

□

We are now able to consider the statement of desired convergence result.

Theorem 7. *Let $(\xi_t^{n\,*})_{0\le t\le T}$ be the optimal hedging ratio associated to the $n-$approximate stochastic control problem (13). $\forall t\in[0,T]$, for every $\varepsilon>0$, $\exists N\in\mathbb{N}$, such that $\forall n\ge N$,*

$$|l(t,s,\nu,v)-l^{\xi^{n\,*}}(t,s,\nu,v)|<\varepsilon,$$

i.e.

$$\lim_{n\to+\infty} l^{\xi^{n\,*}}(t,s,\nu,v)=l(t,s,\nu,v).$$

Moreover

$$\lim_{n\to+\infty} l_n(t,s,\nu,v)=\lim_{n\to+\infty} l_n(t,s,\boldsymbol{\nu},v)=l(t,s,\nu,v).$$

Proof. Let fix $\varepsilon>0$ and $t\in[0,T]$. Suppose that $(\xi_t^{n\,*})_{0\le t\le T}$ is the optimal hedging ratio associated to the $n-$approximate stochastic control problem (13), as $\forall n\in\mathbb{N}$,

$$l_n(t,s,\nu,v)=l_n(t,s,\boldsymbol{\nu},v),$$

then we have that

$$l_n(t,s,\nu,v)=l_n^{\xi^{n\,*}}(t,s,\nu,v).$$

Using Lemma 6, we have that

$$\begin{aligned}\lim_{n\to\infty} l_n(t,s,\nu,v)&=\lim_{n,m\to\infty} l_m^{\xi^{n\,*}}(t,s,\nu,v)\\&=\lim_{n\to\infty} l^{\xi^{n\,*}}(t,s,\nu,v)\end{aligned}$$

or equivalently

$$\lim_{n\to\infty}\Big(l_n(t,s,\nu,v)-l^{\xi^{n\,*}}(t,s,\nu,v)\Big)=0. \tag{14}$$

Therefore by definition of the limit, $\exists N_1\in\mathbb{N}$, such that $\forall n\ge N_1$,

$$|l_n(t,s,\nu,v)-l^{\xi^{n\,*}}(t,s,\nu,v)|<\frac{\varepsilon}{2}. \tag{15}$$

Moreover, considering $\underline{l}(t,s,\nu,v)$ defined by

$$\underline{l}(t,s,\nu,v):=\lim_{n\to\infty} l_n(t,s,\nu,v)=\lim_{n\to\infty}\inf_{\xi_t^n\in\mathcal{R}_n} l_n^{\xi^n}(t,s,\nu,v),$$

thus $\exists N_2\in\mathbb{N}$, such that $\forall n\ge N_2$,

$$|\underline{l}(t,s,\nu,v)-l_n(t,s,\nu,v)|\le\frac{\varepsilon}{2} \tag{16}$$

Note that since

$$l(t,s,\nu,v)=\inf_{\xi_t\in\mathcal{R}}\lim_{n\to\infty} E\Big(L(h(S_T^n),V_T^n)|S_t^n=s,\hat{\nu}_t=\nu,V_t^n=v\Big),$$

we have that

$$l(t,s,\nu,v)\ge\underline{l}(t,s,\nu,v).$$

By choosing $N:=\max(N_1,N_2)$, we have that for $n\ge N$ inequalities (15) and (16) are satisfied. In this case, we have that

$$\begin{aligned}|l(t,s,\nu,v)-l^{\xi^{n\,*}}(t,s,\nu,v)|&\le|\underline{l}(t,s,\nu,v)-l^{\xi^{n\,*}}(t,s,\nu,v)|\\&\le\underbrace{|\underline{l}(t,s,\nu,v)-l_n(t,s,\nu,v)|}_{<\frac{\varepsilon}{2}}+\underbrace{|l_n(t,s,\nu,v)-l^{\xi^{n\,*}}(t,s,\nu,v)|}_{<\frac{\varepsilon}{2}}\\&<\varepsilon.\end{aligned}$$

Therefore, $\exists N\in\mathbb{N}$ such that $\forall n\ge N$,

$$|l(t,s,\nu,v)-l^{\xi^{n\,*}}(t,s,\nu,v)|<\varepsilon$$

or equivalently

$$\lim_{n\to+\infty} l^{\xi^{n\,*}}(t,s,\nu,v) = l(t,s,\nu,v).$$

Moreover, using (14), we deduce that $\forall t\in[0,T]$,

$$\begin{aligned}\lim_{n\to+\infty} l_n(t,s,\boldsymbol{\nu},v) = \lim_{n\to+\infty} l_n(t,s,\nu,v) &= \lim_{n\to+\infty} l^{\xi^{n\,*}}(t,s,\nu,v)\\ &= l(t,s,\nu,v).\end{aligned}$$

This completes the proof as we have proved the two stated convergence results. □

These convergence results are crucial for the following. On the one hand, it means that the optimal hedging ratio $(\xi_t^{n\,*})_{0\le t\le T}$ associated to the $n-$approximate stochastic control problem (13) is $\varepsilon-$optimal for the original problem. One the other hand, we know that the solution of the approximate control problem $l_n(.)$ converges toward the solution of the initial control problem $l(.)$. Therefore, thanks to these results, we know that instead of solving the original non-Markovian problem, we can solve the approximate Markovian problem with an error that can be relatively small if n is large enough.

5 Solution of the approximate Markovian problem

We have just shown that we can solve the optimal problem with a small error by solving the Markovian problem. In this section, we thus solve this problem using classical dynamic programming techniques and more precisely the Hamilton-Jacobi-Bellman (HJB) equation. The Markovian problem is equivalent to solving a partial hedging problem in a multidimensional stochastic volatility environment. The partial hedging problem under stochastic (one dimensional) volatility model has already been investigated in the literature by [29]. It follows that the problem requires solving a nonlinear partial differential equation and therefore the solution cannot be reduced to an expectation by the Feynman-Kac theorem. In our multidimensional volatility case, we will also observe that the control problem involves solving a nonlinear PDE making it quite complex and not allowing to deduce an explicit form of the optimal solution. Inspired by [31], we propose a dual control method to obtain approximate solutions of the problem.

Recall that we define the approximate partial hedging problem by a Markovian stochastic control problem of the form:

$$l_n(t,s,\boldsymbol{\nu},v) = \inf_{\xi_t^n\in\mathcal{R}_n} E\bigg(L(h(S_T^n),V_T^n)|S_t^n=s,\boldsymbol{\nu_t}=\boldsymbol{\nu},V_t^n=v\bigg), \tag{17}$$

where $\mathcal{R}_n$ the set of all progressively measurable processes $(\xi_t^n)_{0\le t\le T}$ with regards to $\mathcal{F}_t$ valued in $\mathbb{R}$ such that $E(\int_0^T(\xi_t^n)^2(S_t^n)^2\hat{\nu}_t dt) < +\infty$. As the stochastic control problem (17) is Markovian, we can solve it using the HJB equation. Assuming that $l_n(t,s,\boldsymbol{\nu},v)$ is locally bounded on $[0,T)\times\mathbb{R}^{2+n}$ and the hamiltonian associated to the problem (17) is finite and upper semicontinous on $[0,T)\times\mathbb{R}^{2+n}\times\mathbb{R}^{2+n}\times\mathcal{S}_{2+n}$, classic results from dynamic programming (see Theorem 7.4. in [35]) imply that $l_n(t,s,\boldsymbol{\nu},v)$ is the viscosity solution of the following HJB[1]:

$$\begin{aligned}
-\partial_t l_n - \inf_{\xi^n\in\mathbb{R}}\bigg\{&\partial_s l_n\,\hat{\mu}s + \sum_{i=1}^n \partial_{\nu_i} l_n\,\bigg(-x_i\nu_i + b(\hat{\nu})\bigg) + \partial_v l_n\,\bigg(rv + (\hat{\mu}-r)\,\xi^n s\bigg)\\
&+\frac{1}{2}\partial_{ss} l_n\,\hat{\nu}s^2 + \frac{1}{2}\partial_{vv} l_n\,(\xi^n)^2\,\hat{\nu}s^2 + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \partial_{\nu_i\nu_j} l_n\,\sigma^2(\hat{\nu}) + \partial_{sv} l_n\,\xi^n\,\hat{\nu}s^2\\
&+\rho\sum_{i=1}^n \partial_{\nu_i s} l_n\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu}) + \rho\sum_{i=1}^n \partial_{\nu_i v} l_n\,\xi^n\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})\bigg\} = 0,\\
l_n(T,s,\nu,v) = L\bigg(h(s),v\bigg).&
\end{aligned} \tag{18}$$

Proposition 8. *The primal optimal control $(\xi_t^{n\,*})_{0\le t\le T}$ is given by*

$$\xi_t^{n\,*} = -\frac{\partial_v l_n\,(\hat{\mu}_t - r)S_t^n + \partial_{sv} l_n\,\hat{\nu}_t(S_t^n)^2 + \rho\sum_{i=1}^n \partial_{\nu_i v} l_n\,\sqrt{\hat{\nu}_t}S_t^n\,\sigma(\hat{\nu}_t)}{\partial_{vv} l_n\,\hat{\nu}_t(S_t^n)^2}$$

[1]For the sake of clarity, we write l_n instead of $l_n(t,s,\boldsymbol{\nu},v)$ and $\hat{\nu}$ instead of $\nu_0 + \sum_{i=1}^n w_i\nu_i$

with $\hat{\nu}_t = \nu_0 + \sum_{i=1}^n w_i \nu_t^i$ and the associated solution $l_n(.)$ solves a nonlinear PDE of the form

$$\partial_t l_n + \mathcal{L}_{s,\boldsymbol{\nu}} l_n - \frac{\Big(\partial_v l_n\,(\hat{\mu}-r)s + \partial_{sv} l_n\,\hat{\nu}s^2 + \rho\sum_{i=1}^n \partial_{\nu_i v} l_n\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})\Big)^2}{2\partial_{vv} l_n\,\hat{\nu}s^2} = 0, \tag{19}$$

with $\mathcal{L}_{s,\boldsymbol{\nu}}$ the generator associated to S^n and $\boldsymbol{\nu}$.

Proof. To deduce the optimal control associated to the Markovian control problem, we solve the HJB equation (18). The HJB equation has a solution if the infimum is different from $-\infty$, it is the case if $\partial_{vv} l_n \geq 0$. In this case, assuming that $\partial_{vv} l_n \geq 0$, the infimum is obtained by the first order condition i.e. by cancelling the derivative of the function with respect to ξ^n. Therefore, the optimal ξ^n denoted by $\xi^{n\,*}$ is such that

$$\partial_{vv} l_n\,\xi^n \hat{\nu}s^2 + \partial_v l_n\,(\hat{\mu}-r)s + \partial_{sv} l_n\,\hat{\nu}s^2 + \rho\sum_{i=1}^n \partial_{\nu_i v} l_n\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu}) = 0,$$

we deduce that

$$\xi^{n\,*} = -\frac{\partial_v l_n\,(\hat{\mu}-r)s + \partial_{sv} l_n\,\hat{\nu}s^2 + \rho\sum_{i=1}^n \partial_{\nu_i v} l_n\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})}{\partial_{vv} l_n\,\hat{\nu}s^2}.$$

Plugging the optimal control into the HJB equation and consider the generator $\mathcal{L}_{s,\boldsymbol{\nu}}$ defined by

$$\mathcal{L}_{s,\boldsymbol{\nu}} := \partial_s\,\hat{\mu}s + \sum_{i=1}^n \partial_{\nu_i}\Big(-x_i\nu_i + b(\hat{\nu})\Big) + \frac{1}{2}\partial_{ss}\,\hat{\nu}s^2 + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \partial_{\nu_i\nu_j}\,\sigma^2(\hat{\nu}) + \rho\sum_{i=1}^n \partial_{\nu_i s}\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu}),$$

the optimal solution satisfies the nonlinear PDE given by

$$\partial_t l_n + \mathcal{L}_{s,\boldsymbol{\nu}} l_n - \frac{\Big(\partial_v l_n\,(\hat{\mu}-r)s + \partial_{sv} l_n\,\hat{\nu}s^2 + \rho\sum_{i=1}^n \partial_{\nu_i v} l_n\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})\Big)^2}{2\partial_{vv} l_n\,\hat{\nu}s^2} = 0.$$

□

The PDE satisfied by the optimal solution is nonlinear, therefore we cannot reduce l_n as an expectation using Feynman-Kac theorem. The dual problem is a way to overcome this nonlinearity problem as it usually allows to transform a nonlinear PDE into a linear one. In our problem, the dual transformation does not allow to obtain a linear PDE. Nevertheless, we still consider the dual approach as it will allow to deduce approximate solutions to our problem by applying a dual control method. To this end, we apply the Legendre-Fenchel transform to the problem (17) and consider the concave dual $\hat{l}_n(.)$ of $l_n(.)$ with respect to the variable v as the opposite of the Legendre-Fenchel transform, such that

$$\begin{aligned}\hat{l}_n(t,s,\boldsymbol{\nu},z) &:= -\sup_v\{zv - l_n(t,s,\boldsymbol{\nu},v)\},\\ &= \inf_v\{l_n(t,s,\boldsymbol{\nu},v) - zv\}.\end{aligned}$$

We observe that as $l_n(t,s,\boldsymbol{\nu},v)$ is convex in v then $\hat{l}_n(t,s,\boldsymbol{\nu},z)$ is concave in z. We also associate the terminal value to the dual solution $\hat{l}_n$ given by

$$\hat{l}_n(T,s,\boldsymbol{\nu},z) = \hat{L}(h(s),z).$$

Based on the PDE satisfied by the primal solution $l_n(t,s,\boldsymbol{\nu},v)$, we deduce the PDE satisfied by $\hat{l}_n(t,s,\boldsymbol{\nu},z)$.

Proposition 9. *The dual solution $\hat{l}_n(t,s,\boldsymbol{\nu},z)$ satisfies the nonlinear PDE*

$$\begin{aligned}0 =& \partial_t\hat{l}_n + \mathcal{L}_{s,\boldsymbol{\nu}}\hat{l}_n - zr\partial_z\hat{l}_n + \frac{1}{2\hat{\nu}s^2}z^2\,(\hat{\mu}-r)^2 s^2\partial_{zz}\hat{l}_n - z\,(\hat{\mu}-r)s\partial_{sz}\hat{l}_n,\\ &-\frac{1}{\sqrt{\hat{\nu}}}\rho\sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n\,z\,(\hat{\mu}-r)\sigma(\hat{\nu}) - \frac{1}{2\partial_{zz}\hat{l}_n}\,\sigma(\hat{\nu})^2(1-\rho^2)\,\sum_{i=1}^n\sum_{j=1}^n \partial_{\nu_i z}\hat{l}_n\partial_{\nu_j z}\hat{l}_n,\end{aligned} \tag{20}$$

with the associated terminal value $\hat{l}_n(T,s,\boldsymbol{\nu},z) = \hat{L}(h(s),z)$. Moreover, the optimal primal control $(\xi_t^{n\,*})_{0\leq t\leq T}$ can be expressed in term of dual solution as, for $t\in[0,T]$,

$$\xi_t^{n\,*} = \frac{Z_t\,\partial_{zz}\hat{l}_n\,(\hat{\mu}_t - r)S_t^n - \partial_{sz}\hat{l}_n\,\hat{\nu}_t(S_t^n)^2 - \rho\sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n\,\sqrt{\hat{\nu}_t}S_t^n\,\sigma(\hat{\nu}_t)}{\hat{\nu}_t(S_t^n)^2}. \tag{21}$$

Proof. The proof is provided in Appendix. □

Actually, the dual solution is the solution of a new stochastic control problem. To prove it, we introduce the dual process $(Z_t)_{0\leq t\leq T}$ controlled by the dual control process $(\gamma_t)_{0\leq t\leq T}$ and defined, for $t\in[0,T]$, by the following SDE

$$dZ_t = -rZ_t dt - Z_t \frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}} dW_S(t) + \gamma_t dB_v(t),$$

where B_v is a standard brownian motion, independent from W_s defined such that for $t\in[0,T]$,

$$W_v(t) = \rho W_s(t) + \sqrt{1-\rho^2} B_v(t).$$

The dual control process $(\gamma_t)_{0\leq t\leq T}$ is admissible if $(\gamma_t)_{0\leq t\leq T}$ is a progressively measurable and square integrable process in regards to $\mathcal{F}_t$. We now define the dual stochastic control problem. In addition, we show that there is no duality gap as the primal solution can be written in terms of the dual solution.

Proposition 10. *The dual solution $\hat{l}_n(t,s,\boldsymbol{\nu},z)$ is solution of a stochastic control problem such that*

$$\hat{l}_n(t,s,\boldsymbol{\nu},z) = \sup_{\gamma_t\in\mathcal{D}} E_{t,s,\boldsymbol{\nu},z}\Big(\hat{L}(h(S_T^n), Z_T)\Big),$$

with $\mathcal{D}$ be the set of all progressively measurable and square integrable processes in regards to $\mathcal{F}_t$ valued in $\mathbb{R}$. Moreover the optimal dual control $(\gamma_t^*)_{0\leq t\leq T}$ is given by

$$\gamma_t^* = -\sigma(\hat{\nu}_t)\sqrt{1-\rho^2}\sum_{i=1}^n \frac{\partial_{\nu_i z}\hat{l}_n}{\partial_{zz}\hat{l}_n}. \tag{22}$$

Proof. Assume that $\hat{l}_n^{bis}(t,s,\nu,z)$ is defined by

$$\hat{l}_n^{bis}(t,s,\nu,z) := \sup_{\gamma_t\in\mathcal{D}} E_{t,s,\boldsymbol{\nu},z}\Big(\hat{L}(h(S_T^n), Z_T)\Big). \tag{23}$$

We just have to prove that the HJB equation associated to $\hat{l}_n^{bis}$ matches the PDE (20). The HJB equation associated to the control problem (23) is given by

$$\begin{aligned}
0 = \partial_t\hat{l}_n^{bis} &+ \mathcal{L}_{s,\hat{\nu}}\hat{l}_n^{bis} - z^\gamma r\partial_z\hat{l}_n^{bis} + \sup_{\gamma\in\mathbb{R}}\Bigg\{\frac{1}{2}\partial_{zz}\hat{l}_n^{bis}\Big(z^2\frac{(\hat{\mu}-r)^2}{\hat{\nu}^2}+\gamma^2\Big) \\
&-\partial_{zs}\hat{l}_n^{bis}\, z\frac{(\hat{\mu}-r)}{\sqrt{\hat{\nu}_t}}\sqrt{\hat{v}}s + \sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n^{bis}\Big(\rho z\frac{(\hat{\mu}-r)}{\sqrt{\hat{\nu}_t}}\sqrt{\hat{v}}s + \sqrt{1-\rho^2}\gamma\sigma(\hat{\nu})\Big)\Bigg\}, \\
\hat{l}_n^{bis}(T,s,\nu,z) &= \hat{L}(h(s),z).
\end{aligned}$$

The supremum is different of $+\infty$ if $\partial_{zz}\hat{l}_n^{bis}\leq 0$. In this case, using the first order condition, the optimal dual control $(\gamma_t^*)_{0\leq t\leq T}$ is given by

$$\gamma^* = -\frac{\sqrt{1-\rho^2}\sigma(\hat{\nu})\sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n^{bis}}{\partial_{zz}\hat{l}_n^{bis}}.$$

Thus, the PDE satisfied by $\hat{l}_n^{bis}$ becomes

$$\begin{aligned}
0 =& \partial_t\hat{l}_n^{bis} + \mathcal{L}_{s,\boldsymbol{\nu}}\hat{l}_n^{bis} - z\, r\partial_z\hat{l}_n^{bis} + \frac{1}{2}\partial_{zz}\hat{l}_n^{bis} z^2\frac{(\hat{\mu}-r)^2}{\hat{\nu}^2} - \partial_{zs}\hat{l}_n^{bis}\, z(\hat{\mu}-r)s \\
&+\sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n^{bis}\,\rho\, z\,\frac{(\hat{\mu}-r)}{\sqrt{\hat{\nu}}}\,\sigma(\hat{\nu})\, s - \frac{1}{2\partial_{zz}\hat{l}_n^{bis}}\sigma^2(\hat{\nu})(1-\rho^2)\sum_{i=1}^n\sum_{j=1}^n \partial_{\nu_i z}\hat{l}_n^{bis}\partial_{\nu_j z}\hat{l}_n^{bis},
\end{aligned} \tag{24}$$

We can observe that (24) is exactly the same PDE as (20) and as the two PDE's have the same terminal value, we can conclude that

$$\hat{l}_n(t,s,\boldsymbol{\nu},z) = \hat{l}_n^{bis}(t,s,\boldsymbol{\nu},z) = \sup_{\gamma_t\in\mathcal{D}} E_{t,s,\boldsymbol{\nu},z}\Big(\hat{L}(h(S_T^n), Z_T)\Big).$$

In this case, the optimal dual control is given by

$$\gamma_t^* = -\frac{\sqrt{1-\rho^2}\sigma(\hat{\nu}_t)\sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n}{\partial_{zz}\hat{l}_n}.$$

□

Proposition 11. *By choosing $z(t,s,v)$ solution of*

$$\partial_z \hat{l}_n(t,s,\boldsymbol{\nu},z) + v = 0, \tag{25}$$

then

$$l_n(t,s,\boldsymbol{\nu},v) = \hat{l}_n(t,s,\boldsymbol{\nu},z) + zv, \tag{26}$$

with $z = z(t,s,\boldsymbol{\nu},v)$ the value at time t of the dual process $(Z_t)_{0\le t\le T}$.

Proof. Consider $l_n^{bis}(t,s,\nu,v)$ to be the dual of the dual solution and defined by

$$l_n^{bis}(t,s,\boldsymbol{\nu},v) := \sup_z \left\{\hat{l}_n(t,s,\boldsymbol{\nu},z) + zv\right\}.$$

Our goal is to prove that the dual of the dual is the primal. First, using the first order condition since $\hat{l}_n(t,s,\boldsymbol{\nu},z)$ is concave in z, we have that $z(t,s,\boldsymbol{\nu},v)$ is solution of

$$\partial_z \hat{l}_n(t,s,\boldsymbol{\nu},z) + v = 0, \tag{27}$$

in this case l_n^{bis} reduces to

$$l_n^{bis}(t,s,\boldsymbol{\nu},v) = \hat{l}_n(t,s,\boldsymbol{\nu},z) + zv$$

with z satisfying (27). Now, we just have to prove that $l_n^{bis}(t,s,\boldsymbol{\nu},v) = l_n(t,s,\boldsymbol{\nu},v)$. The proof of this equality is similar to the proof of Proposition 9 this is why we have decided not to go into detail but it is easy to show that $l_n^{bis}(t,s,\boldsymbol{\nu},v)$ satisfies the same PDE then the primal solution $l_n(t,s,\boldsymbol{\nu},v)$ given by (19). Moreover, the two PDEs have the same terminal value. In fact, we can rewrite the terminal value of l_n^{bis} as

$$\begin{aligned} L^{bis}(h(s),v) &= \sup_z \left\{ \inf_v \left(L(h(s),v) - zv \right) + zv \right\} \\ &= \sup_z \left\{ zv - \sup_v \left(zv - L(h(s),v) \right) \right\} \end{aligned}$$

As $\sup_v \left(zv - L(h(s),v) \right)$ is the Legendre transform of $L(h(s),v)$, $L^{bis}(h(s),v)$ is the Legendre transform of the Legendre transform of $L(h(s),v)$. By the Theorem of Fenchel-Moreau, as $L(.)$ is a proper continous convex function, we obtain that

$$L^{bis}(h(s),v) = L(h(s),v).$$

Therefore, as $l_n(t,s,\boldsymbol{\nu},v)$ and $l_n^{bis}(t,s,\boldsymbol{\nu},v)$ satisfy the same PDE with the same terminal value, we conclude that

$$l_n^{bis}(t,s,\boldsymbol{\nu},v) = l_n(t,s,\boldsymbol{\nu},v).$$

Thus we obtain that given $z(t,s,\boldsymbol{\nu},v)$ is solution of

$$\partial_z \hat{l}_n(t,s,\boldsymbol{\nu},z) + v = 0,$$

then

$$l_n(t,s,\boldsymbol{\nu},v) = \hat{l}_n(t,s,\boldsymbol{\nu},z) + zv.$$

□

As the primal solution is a function of the dual solution, we can derive the primal solution in the case where the dual solution admits a closed formula. However, in our case, we observe that by switching to the dual problem, although the nonlinear term of the PDE is less important, the PDE (20) satisfied by the dual solution $\hat{l}_n$ remains nonlinear. Thus, as the nonlinearity problem persists, we are not able, in general, to express the dual solution as an expectation. The partial hedging problem is still complicated to solve. There is nevertheless a specific case for which a closed formula of the dual solution can be obtained. Indeed, if we consider a linear payoff defined as

$$H_T^{linear} = \alpha + \beta S_T, \tag{28}$$

a power loss function and a rough Heston model then the solution of the dual problem is obtained by closed formula. This particular case is a toy case since most of payoffs are generally not linear functions of the underlyings. However, it will allow to quantify the errors made when considering sub-optimal hedging strategies deduced by the dual control method and therefore to benchmark this dual control approach.

Lemma 12. *Suppose a power loss of the form $L(h,v) = \frac{1}{p}(h-v)^p$ with $p = 2n$, $n \in \mathbb{N}_0$, then*

$$\hat{L}(h(s), z) = -\frac{1}{q} z^q - h(s)\, z,$$

with $q = \frac{p}{p-1}$.

Proof. The proof is provided in Appendix. □

Proposition 13. *Consider a linear payoff H_T^{linear} given by (28), a power loss of the form $L(h(s), v) = \frac{1}{p}(h(s)-v)^p$, suppose that the volatility is modeled by a rough Heston model such that $b(x) = \kappa(\theta - x)$ and $\sigma(x) = \zeta\sqrt{x}$. Therefore,*

$$\hat{l}_n(t, s, \boldsymbol{\nu}, z) = -\frac{z^q}{q} \exp\left(C_t + \sum_{i=1}^n D_t^i \nu_t^i\right) - \left(e^{-r(T-t)}\alpha + \beta S_t^n\right) z,$$

where C_t and $(D_t)_{i=1,..,n}$ are time-dependent functions solution of Riccati ODEs given respectively by

$$\partial_t C_t = rq - \frac{1}{2}\, q(q-1)\, A^2\, \nu_0 - \sum_{i=1}^n D_t^i \Big(\kappa(\theta - \nu_0) - q\rho A \zeta\, \nu_0\Big) - \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n D_t^i D_t^j \zeta^2\, \nu_0 \Big(1 - (1-\rho^2)\frac{q}{q-1}\Big),$$
$$C_T = 0,$$

and for $i = 1, .., n$,

$$\begin{aligned}
\partial_t D_t^i =& x_i D_t^i + w_i \sum_{j=1}^n D_t^j (\kappa + q\rho A\zeta) - \frac{1}{2}\zeta^2 w_i \sum_{j=1}^n\sum_{k=1}^n D_t^j D_t^k \Big(1 - (1-\rho^2)\frac{q}{q-1}\Big) \\
& - \frac{1}{2} w_i\, q(q-1)\, A^2, \\
D_T^i =& 0.
\end{aligned}$$

Moreover, the primal solution is given by

$$l_n(t, s, \boldsymbol{\nu}, v) = -\frac{Z_t^q}{q} \exp\left(C_t + \sum_{i=1}^n D_t^i \nu_t^i\right) - \left(e^{-r(T-t)}\alpha + \beta S_t^n\right) Z_t + Z_t\, V_t^n,$$

with

$$Z_t = \left(V_t^n - \left(e^{-r(T-t)}\alpha + \beta S_t^n\right)\right)^{\frac{1}{q-1}} \exp\left(-\frac{1}{q-1}(C_t + \sum_{i=1}^n D_t^i \nu_t^i)\right)$$

and the associated optimal hedging ratio is such that

$$\xi_t^{n\,*} = \frac{1}{\hat{\nu}_t S_t^n}\left(e^{-r(T-t)}\alpha + \beta S_t^n - V_t^n\right)\left(\frac{1}{p-1}(\hat{\mu}_t - r) - \rho\zeta\hat{\nu}_t \sum_{i=1}^n D_t^i\right) + \beta.$$

Proof. The full proof is in Appendix. Here is a summary of the steps in the proof. Firstly, to obtain the form of $\hat{l}_n(.)$, we just need to consider the following ansatz

$$\hat{l}_n(t,s,\boldsymbol{\nu},z) = -\frac{z^q}{q}\exp\left(C_t + \sum_{i=1}^{n} D_t^i \nu_t^i\right) + \left(e^{-r(T-t)}\alpha + \beta s\right) z,$$

and plug it into PDE (20). Then using the form of $\hat{l}_n(.)$ and the relation (26) between the primal and the dual, we deduce the form of $l_n(.)$. Finally, the form of the optimal heding ratio is deduced using the form (21) of $\xi_t^{n\,*}$ and the closed form of the dual solution. $\square$

Remark. Using the same assumptions as Proposition 13 but instead of the rough Heston model, consider the classical Heston model, i.e. $n = 1$, then C_t and D_t admit closed formulas.

To approximate the solution of our problem for general payoffs, we rely on the expression of the primal as a function of the dual. For that, define a set $U \subseteq \mathbb{R}$ such that U is a convex compact subset of $\mathbb{R}$ with non-empty interior and denote $\mathcal{U}$ a set of progressively measurable and square integrable processes valued in U such that $\mathcal{U} \subseteq \mathcal{D}$. Using the relation between the primal and the dual solution, we deduce that

$$\begin{aligned} l_n(t,s,\boldsymbol{\nu},v) &= \sup_z \left\{ \sup_{\gamma_t \in \mathcal{D}} E_{t,s,\boldsymbol{\nu},z}\left(\hat{L}(h(S_T^n), Z_T)\right) + zv \right\} \\ &\geq \sup_z \left\{ \sup_{\gamma_t \in \mathcal{U}} E_{t,s,\boldsymbol{\nu},z}\left(\hat{L}(h(S_T^n), Z_T)\right) + zv \right\}. \end{aligned} \tag{29}$$

Based on inequality (29) and inspired by the dual control method stated in [31], we will define lower and upper bounds for the primal solution of the Markov partial hedging problem. For that, for every fixed admissible dual control $(\gamma_t)_{0\leq t\leq T}$, we define

$$Y(t,s,\boldsymbol{\nu},z;\gamma) := E_{t,s,\boldsymbol{\nu},z}\left(\hat{L}(\,h(S_T^n), Z_T(\gamma)\,)\right), \tag{30}$$

We derive now a theorem that states how to deduce upper and lower bounds for the primal solution.

Theorem 14. *Let $\mathcal{U} \subseteq \mathcal{D}$ be a set of admisible dual controls and define $\underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z)$ by*

$$\underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) := \sup_{\gamma_t \in \mathcal{U}} Y(t,s,\boldsymbol{\nu},z;\gamma) \leq \hat{l}_n(t,s,\boldsymbol{\nu},z).$$

Therefore, defining $\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v)$ by

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) := \sup_z \left\{ \underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) + zv \right\},$$

we have that

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) \leq l_n(t,s,\boldsymbol{\nu},v).$$

Moreover, suppose that $\underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v)$ is twice continously differentiable, stricly concave and $z(t,s,\boldsymbol{\nu},v)$ is the solution of

$$\partial_z \underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) + v = 0. \tag{31}$$

We define the primal control by $\bar{\xi}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z)$ such that

$$\bar{\xi}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) = \frac{z\,\partial_{zz}\underline{\hat{l}}_n^{\mathcal{U}}\,(\hat{\mu}-r)s - \partial_{sz}\underline{\hat{l}}_n^{\mathcal{U}}\,\hat{\nu}s^2 - \rho\sum_{i=1}^{n}\partial_{\nu_i z}\underline{\hat{l}}_n^{\mathcal{U}}\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})}{\hat{\nu}s^2}, \tag{32}$$

with $z = z(t,s,\boldsymbol{\nu},v)$ solution of (31) and we consider the associated self-financing portfolio denoted by $(\bar{V}_t^{\mathcal{U}})_{0\leq t\leq T}$. If we define

$$\bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) := E_{t,s,\hat{\nu},v}\left(L(h(S_T^n), \bar{V}_T^{\mathcal{U}})\right),$$

then the primal solution satisfies

$$l_n(t,s,\boldsymbol{\nu},v) \leq \bar{l}_n(t,s,\boldsymbol{\nu},v).$$

Therefore, we obtain lower and upper bounds for the primal solution such that

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) \leq l_n(t,s,\boldsymbol{\nu},v) \leq \bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v).$$

Proof. The proof is almost direct. Fix a set of admissible dual controls $\mathcal{U} \subseteq \mathcal{D}$, as

$$\begin{aligned}\hat{l}_n(t,s,\boldsymbol{\nu},z) &= \sup_{\gamma_t \in \mathcal{D}} E_{t,s,\boldsymbol{\nu},z}\left(\hat{L}(\,h(S_T^n), Z_T(\gamma)\,)\right),\\ &= \sup_{\gamma_t \in \mathcal{D}} Y(t,s,\boldsymbol{\nu},z;\gamma),\end{aligned}$$

we immediately obtain that

$$\underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) \leq \hat{l}_n(t,s,\boldsymbol{\nu},z).$$

Moreover, we observe that

$$\begin{aligned}l_n(t,s,\boldsymbol{\nu},v) &= \sup_z \left\{\hat{l}_n(t,s,\boldsymbol{\nu},v) + zv\right\}\\ &\geq \sup_z \left\{\underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) + zv\right\}\\ &:= \underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v).\end{aligned}$$

The inequality of the upper bound is obvious since using the definition of the optimal solution,

$$\begin{aligned}l_n(t,s,\boldsymbol{\nu},v) &= \inf_{\xi_t^n \in \mathcal{R}_n} E_{t,s,\boldsymbol{\nu},v}\left(L(h(S_T^n), V_T^n)\right)\\ &\leq E_{t,s,\boldsymbol{\nu},v}\left(L(h(S_T^n), \bar{V}_T^{\mathcal{U}})\right)\\ &= \bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v).\end{aligned}$$

Therefore, we prove

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) \leq l_n(t,s,\boldsymbol{\nu},v) \leq \bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v).$$

□

Remark. The way we define the upper bound is different than in [31], we have made this choice in order to prove a convergence result of the bounds toward the primal solution when considering large dual control subsets $\mathcal{U}$. Note also that the approximate hedging ratio $\bar{\xi}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z)$ defined by (32) has the same form as the optimal hedging ratio $\xi_t^{n\,*}$ defined in (21) with the difference that we consider the subset of admissible dual control $\mathcal{U}$ instead of $\mathcal{D}$.

Theorem 14 is important since it allows to approximate the primal by different sub-optimal choices of dual controls and thus, enables to easily deduce sub-optimal hedging strategies that can be computed relying on Monte Carlo simulations. Note that there is a wide range of possible subset choices, but depending on the choice of the subset $\mathcal{U}$, the computation time of the bounds can be quite substantial. We refer to [31] for the algorithm allowing to compute the bounds via a Monte Carlo approach. In this paper, we decide to only focus on a particular dual control subset that allows to express explicit formulas for the lower bound and the approximate hedging ratio. We discuss later the choice of the dual control subset.

Theorem 14 introduces a sub-optimal hedging strategy associated with the approximate hedging ratio $\bar{\xi}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z)$ and the upper bound $\bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v)$ that can be implemented in practice and for which we can obtain a bound on the error made by considering this strategy instead of the optimal since:

$$\begin{aligned}|\bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) - l_n(t,s,\boldsymbol{\nu},v)| &\leq |\bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) - \underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v)|\\ &:= C_{UL}^{\mathcal{U}}.\end{aligned}$$

Note that $C_{UL}^{\mathcal{U}}$ only depends on the set $\mathcal{U}$ and not necessary on n. In practice, by computing upper and lower bounds, we can deduce an upper bound on the error between the dual control approximate solution and the optimal solution of the Markov problem. Therefore, it enables to verify that the error is acceptable and that the proposed dual control method is relevant. Moreover, we show that if we consider a set sequence of admissible dual control $(\mathcal{U}_i)_{i\in\mathbb{N}}$ such that $\mathcal{U}_i \subseteq \mathcal{U}_{i+1}$ and $\lim_{i\to+\infty} \mathcal{U}_i = \mathcal{D}$ then the approximate solution of the Markov problem also converges to the primal solution.

Proposition 15. *Consider a compact set sequence of admissible dual controls $(\mathcal{U}_i)_{i\in\mathbb{N}}$ such that $\mathcal{U}_i \subseteq \mathcal{U}_{i+1}$ with $\lim_{i\to+\infty}\mathcal{U}_i = \mathcal{D}$ and the associated sequence of functions $\left(\underline{\tilde{l}}_n^{\mathcal{U}_i}(.)\right)_{i\in\mathbb{N}}$ twice continously differentiable with second derivatives that converge uniformly in $\mathbb{R}$, then $\forall t\in[0,T]$, $\forall n\in\mathbb{N}$,*

$$\lim_{i\to+\infty} \underbrace{|\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v) - \underline{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)|}_{=C_{UL}^{\mathcal{U}_i}} = 0,$$

i.e.

$$\lim_{i\to+\infty} \bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v) = l_n(t,s,\boldsymbol{\nu},v).$$

Proof. Consider a compact set sequence of admissible dual controls $(\mathcal{U}_i)_{i\in\mathbb{N}}$ such that $\mathcal{U}_i \subsetneq \mathcal{U}_{i+1}$ and $\lim_{i\to+\infty}\mathcal{U}_i = \mathcal{D}$ and fix $n\in\mathbb{N}$. First, we can show the convergence of $\underline{\tilde{l}}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)$ toward $\hat{l}_n(t,s,\boldsymbol{\nu},v)$. In fact, as for $i\in\mathbb{N}$, $\mathcal{U}_i\subseteq\mathcal{U}_{i+1}$, we have that $\forall t\in[0,T]$,

$$\underline{\tilde{l}}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},z) \le \underline{\tilde{l}}_n^{\mathcal{U}_{i+1}}(t,s,\boldsymbol{\nu},z) \le \hat{l}_n^{\mathcal{D}}(t,s,\boldsymbol{\nu},z),$$

and as $(\mathcal{U}_i)_{i=1,\dots,n}$ is a sequence of compact set, the infimum fonction over $\mathcal{U}_i$ is continous for $i\in\mathbb{N}$. Thus taking the limit of $i\to+\infty$, we have that $\forall t\in[0,T]$,

$$\lim_{i\to+\infty}\underline{\tilde{l}}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},z) = \hat{l}_n(t,s,\boldsymbol{\nu},z).$$

In this case, we deduce the convergence of the lower bound of the primal solution toward the primal solution since $\forall t\in[0,T]$,

$$\begin{aligned}
\lim_{i\to+\infty}\underline{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v) &= \lim_{i\to+\infty}\left(\sup_z\left\{\underline{\tilde{l}}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},z)+zv\right\}\right)\\
&= \sup_z\left\{\lim_{i\to+\infty}\underline{\tilde{l}}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},z)+zv\right\}\\
&= \sup_z\left\{\hat{l}_n(t,s,\boldsymbol{\nu},z)+zv\right\}\\
&= l_n(t,s,\boldsymbol{\nu},v).
\end{aligned}$$

It remains to show the convergence of the upper bound to the primal solution. For this purpose, we need to show that the approximate hedge ratio converges to the optimal hedge ratio. As the sequence of functions $\left(\underline{\tilde{l}}_n^{\mathcal{U}_i}(.)\right)_{i\in\mathbb{N}}$ is twice continously differentiable with second derivatives that converge uniformly in $\mathbb{R}$, standard result in Analysis states that $\forall t\in[0,T]$,

$$\begin{aligned}
\lim_{i\to+\infty}\bar{\xi}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},z) &= \lim_{i\to+\infty}\frac{z\,\partial_{zz}\underline{\tilde{l}}_n^{\mathcal{U}_i}\,(\hat{\mu}-r)s - \partial_{sz}\underline{\tilde{l}}_n^{\mathcal{U}_i}\,\hat{\nu}s^2 - \rho\sum_{i=1}^n\partial_{\nu_i z}\underline{\tilde{l}}_n^{\mathcal{U}_i}\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})}{\hat{\nu}s^2}\\
&= \frac{z\,\partial_{zz}\hat{l}_n\,(\hat{\mu}-r)s - \partial_{sz}\hat{l}_n\,\hat{\nu}s^2 - \rho\sum_{i=1}^n\partial_{\nu_i z}\hat{l}_n\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})}{\hat{\nu}s^2}\\
&= \xi_n^*(t,s,\boldsymbol{\nu},z),
\end{aligned}$$

with $z=z(t,s,\boldsymbol{\nu},v)$ solution of

$$\underbrace{\lim_{i\to\infty}\partial_z\underline{\tilde{l}}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},z)}_{=\partial_z\hat{l}_n(t,s,\boldsymbol{\nu},z)} + v = 0.$$

Using similar arguments as for the proof of Proposition 5 and Lemma 6, we deduce that almost surely

$$\bar{V}_T^{\mathcal{U}_i} \to V_T^{n\,*},$$

as i goes to infinity and then $\forall t\in[0,T]$,

$$\lim_{i\to+\infty}\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v) = l_n(t,s,\boldsymbol{\nu},v).$$

We therefore conclude that $\forall t \in [0,T]$,

$$\begin{aligned}\lim_{n\to+\infty}\left(\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)-\underline{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v))\right) &= l_n(t,s,\boldsymbol{\nu},v)-l_n(t,s,\boldsymbol{\nu},v)\\ &=0.\end{aligned}$$

That concludes the proof since we prove the stated proposition. □

The previous proposition shows that if we consider a large enough set of admissible dual controls, the approximate solution converges to the primal solution of the Markov problem. In practice, we observe that even if the set of admissible dual controls $\mathcal{U}$ is small, the error is small, which seems to show that the choice of the dual control does not significantly impact the value of the primal solution.

Let's go back to the original hedging problem under rough volatility, the initial control problem posed was

$$l(t,s,\nu,v)=\inf_{\xi_t\in\mathcal{R}} E\bigg(L(h(S_T),V_T)|S_t=s,\nu_t=\nu,V_t=v\bigg).$$

The proposed approximate solution $\bar{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v)$ is a two-fold approximate solution, on the one hand by the Markov discretization of the volatility process and on the other hand by the sub-optimal choice of the dual control. However, we can show that the error with respect to the optimal solution of the initial problem can be small if n and $\mathcal{U}$ are large enough, this is the purpose of the following proposition.

Proposition 16. *Consider a compact set sequence of admissible dual controls $(\mathcal{U}_i)_{i\in\mathbb{N}}$ and the associated sequence of functions $\left(\underline{l}_n^{\mathcal{U}_i}(.)\right)_{i\in\mathbb{N}}$ satisfying assumptions of Proposition 15. $\forall t\in[0,T]$, $\forall\varepsilon>0$, $\exists N\in\mathbb{N}$, such that $\forall\, n\geq N$, $\exists M\in\mathbb{N}$ such that $\forall i\geq M$,*

$$|l(t,s,\nu,v)-\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)|<\varepsilon$$

and

$$|l(t,s,\nu,v)-l^{\bar{\xi}_n^{\mathcal{U}_i}}(t,s,\boldsymbol{\nu},v)|<\varepsilon.$$

It means that the approximate hedging ratio $\bar{\xi}_n^{\mathcal{U}_i}$ associated to $\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)$ is $\varepsilon-$optimal for the original problem.

Proof. The proof is almost direct. Fix $\varepsilon>0$ and $t\in[0,T]$, from Theorem 7, we know that $\exists N\in\mathbb{N}$ such that $\forall n\geq N$,

$$|l(t,s,\nu,v)-l_n(t,s,\boldsymbol{\nu},v)|<\frac{\varepsilon}{2},$$

and

$$|l(t,s,\nu,v)-l^{\xi_n}(t,s,\boldsymbol{\nu},v)|<\frac{\varepsilon}{2}.$$

Moreover, from Proposition 15, we have that $\forall n\in\mathbb{N}$, $\exists M_1\in\mathbb{N}$ such that $\forall i\geq M_1$,

$$|l_n(t,s,\boldsymbol{\nu},v)-\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)|<\frac{\varepsilon}{2},$$

and since almost surely $\bar{\xi}_n^{\mathcal{U}_i}(t)\to\xi_n(t)$ as $i\to+\infty$, we deduce using a similar argument as in Lemma 6 that $\forall t\in[0,T]$, $\forall n\in\mathbb{N}$, $\exists M_2\in\mathbb{N}$ such that $\forall i\geq M_2$,

$$|l^{\xi_n}(t,s,\boldsymbol{\nu},v)-l^{\bar{\xi}_n^{\mathcal{U}_i}}(t,s,\boldsymbol{\nu},v)|<\frac{\varepsilon}{2}.$$

Therefore choosing $M:=\max(M_1,M_2)$, we know that $\forall n\geq N$, $\exists M\in\mathbb{N}$ such that $\forall i\geq M$,

$$\begin{aligned}|l(t,s,\nu,v)-\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)| =&|l(t,s,\nu,v)-l_n(t,s,\boldsymbol{\nu},v)+l_n(t,s,\boldsymbol{\nu},v)-\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)|\\ \leq&|l(t,s,\nu,v)-l_n(t,s,\boldsymbol{\nu},v)|+|l_n(t,s,\boldsymbol{\nu},v)-\bar{l}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)|\\ <&\varepsilon,\end{aligned}$$

and

$$
\begin{aligned}
|l(t,s,\nu,v) - l^{\bar{\xi}_n^{\mathcal{U}_i}}(t,s,\boldsymbol{\nu},v)| &= |l(t,s,\nu,v) - l^{\xi_n}(t,s,\boldsymbol{\nu},v) + l^{\xi_n}(t,s,\boldsymbol{\nu},v) - l^{\bar{\xi}_n^{\mathcal{U}_i}}(t,s,\boldsymbol{\nu},v)| \\
&\leq |l(t,s,\nu,v) - l^{\xi_n}(t,s,\boldsymbol{\nu},v)| + |l^{\xi_n}(t,s,\boldsymbol{\nu},v) - l^{\bar{\xi}_n^{\mathcal{U}_i}}(t,s,\boldsymbol{\nu},v)| \\
&< \varepsilon.
\end{aligned}
$$

□

The result of Proposition 16 is of course a theoretical result. It is not necessarily satisfied if, for example, we only consider a single set of dual control $\mathcal{U}$ and not a sequence of dual control set. Nevertheless, if n and the dual control set $\mathcal{U}$ are large enough then the error with respect to the original problem should be quite small. In practice, for a fixed set of dual control $\mathcal{U}$, the error is controlled by the number of factors n and the gap $C_{UL}^{\mathcal{U}}$ between lower and upper bounds.

Appropriate choice of the dual control subset $\mathcal{U} \subseteq \mathcal{D}$

Now, we consider a particular subset of dual controls for which explicit formulas can be obtained. Thus for the following, the dual control subset considered is defined as

$$
\mathcal{U} = \left\{ (\gamma_t)_{0 \leq t \leq T} = \Big(c \times Z_t \times \sigma(\hat{\nu}_t) \Big)_{0 \leq t \leq T}, \ c \in U \subseteq \mathbb{R} \right\} \subseteq \mathcal{D}. \tag{33}
$$

We notice that the chosen form of the dual controls belonging to $\mathcal{U}$ is similar to the form of the optimal dual control (22). This particular subset (33) enables to interpret the sub-optimal hedging strategy as well as obtain closed forms for the lower bound and the approximate hedging ratio. First, assuming this subset of admissible dual control, we observe that

$$
\begin{aligned}
\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) &= \sup_z \left\{ \sup_{\gamma_t \in \mathcal{U}} Y(t,s,\boldsymbol{\nu},z;\gamma) + zv \right\} \\
&= \sup_z \left\{ \max_{c \in U} Y(t,s,\boldsymbol{\nu},z;c) + zv \right\} \qquad (34) \\
&= \max_{c \in U} \sup_z \Big\{ Y(t,s,\boldsymbol{\nu},z;c) + zv \Big\}. \qquad (35)
\end{aligned}
$$

Defining, for $c \in U$, $\underline{l}_n(t,s,\boldsymbol{\nu},z;c)$ by

$$
\underline{l}_n(t,s,\boldsymbol{\nu},v;c) := \sup_z \Big\{ Y(t,s,\boldsymbol{\nu},z;c) + zv \Big\},
$$

we have that

$$
\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) = \underline{l}_n(t,s,\boldsymbol{\nu},v;c^*)
$$

with

$$
c^* := \arg\max_{c \in U} \ \underline{l}_n(t,s,\boldsymbol{\nu},v;c).
$$

In this case, the sub-optimal hedging strategy $\bar{V}_t^{\mathcal{U}}$ is easily interpreted as a perfect hedging strategy of a modified payoff. This is similar to the idea of [15, 16] who present partial hedging strategies as perfect hedging of knock-out options.

Proposition 17. *If we consider a subset of admissible dual control $\mathcal{U}$ of the form (33) then*

$$
\bar{V}_t^{\mathcal{U}} = E_{t,s,\boldsymbol{\nu},z}^{\mathbb{Q}(c^*)} \bigg(e^{-r(T-t)} \Big(- \partial_{Z_T} \hat{L}(h(S_T^n), Z_T(c^*)) \Big) \bigg),
$$

where $(\mathbb{Q}(c))_{c \in U}$ called "risk-neutral" measures are $\mathbb{P}-$equivalent measures such that the processes have the following dynamics

$$
dS_t^n = rS_t^n + \sqrt{\hat{\nu}_t} S_t^n dW_s^{\mathbb{Q}(c)}(t),
$$

$$
d\nu_t^i = \bigg(- x_i \nu_t^i + b(\hat{\nu}_t) + \sigma(\hat{\nu}_t) \Big(- \rho \frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}} + \sqrt{1-\rho^2} c\, \sigma(\hat{\nu}_t) \Big) \bigg) dt + \sigma(\hat{\nu}_t) dW_v^{\mathbb{Q}(c)}(t), \ i = 1, ..., n,
$$

$$dZ_t = Z_t\Bigg(\bigg(-r + \frac{(\hat{\mu}_t - r)^2}{\hat{\nu}_t} + c^2\sigma^2(\hat{\nu}_t)\bigg)dt - \frac{(\mu - r)}{\sqrt{\hat{\nu}_t}}dW_S^{\mathbb{Q}(c)}(t) + c\,\sigma(\hat{\nu}_t)dB_v^{\mathbb{Q}(c)}(t)\Bigg)$$

where $W_s^{\mathbb{Q}(c)}$, $W_v^{\mathbb{Q}(c)}$ are standard brownian motions under $\mathbb{Q}(c)-$measure with $d < W_s^{\mathbb{Q}(c)}, W_v^{\mathbb{Q}(c)} >_t = \rho\, dt$ and $B_v^{\mathbb{Q}(c)}$ is a standard brownian motion, independent from $W_s^{\mathbb{Q}(c)}$ defined such that for $t \in [0,T]$,

$$W_v^{\mathbb{Q}(c)}(t) = \rho W_s^{\mathbb{Q}(c)}(t) + \sqrt{1-\rho^2}B_v^{\mathbb{Q}(c)}(t).$$

In particular, for a power loss function of the form $L(h(s), v) = \frac{1}{p}(h(s) - v)^p$ with $p = 2n$, $n \in \mathbb{N}_0$, the value sub-optimal hedging portfolio is given by

$$\bar{V}_t^{\mathcal{U}} = E_{t,s,\boldsymbol{\nu},z}^{\mathbb{Q}(c^*)}\bigg(e^{-r(T-t)}\Big(H_T^n + Z_T^{q-1}\Big)\bigg),$$

with $q = \frac{p}{p-1}$.

Proof. From the Theorem 14, we know, using (31), that, at time t, $z = z^*$ with z^* solution of

$$v = -\partial_z \underline{l}_n^{\mathcal{U}}(t, s, \boldsymbol{\nu}, z),$$

Thus, by (34), we have that

$$\underline{l}_n^{\mathcal{U}}(t, s, \boldsymbol{\nu}, v) = Y(t, s, \boldsymbol{\nu}, z^*; c^*(z^*)) + z^* v,$$

with

$$c^*(z) := \arg\max_{c \in U} Y(t, s, \boldsymbol{\nu}, z; c).$$

But using (35), we also have that

$$\underline{l}_n^{\mathcal{U}}(t, s, \boldsymbol{\nu}, v) = Y(t, s, \boldsymbol{\nu}, z^*(c^*); c^*) + z^*(c^*)v$$

with $z^*(c)$ solution of

$$v = -\partial_z Y(t, s, \boldsymbol{\nu}, z; c),$$

we conclude by unicity that, at time t, $z = z^* = z^*(c^*)$ and $c^*(z^*) = c^*$. Therefore, we obtain, at time t, the following relation

$$\bar{V}_t^{\mathcal{U}} = -\partial_z E_{t,s,\boldsymbol{\nu},z}\Big(\hat{L}(h(S_T^n), Z_T(c^*)\Big).$$

Using the theorem of exchanging expectation and derivative, we have that

$$\begin{aligned}
\bar{V}_t^{\mathcal{U}} &= E_{t,s,\boldsymbol{\nu},z}\Big(-\partial_{Z_t}\hat{L}(h(S_T^n), Z_T(c^*))\Big) \\
&= E_{t,s,\boldsymbol{\nu},z}\Big(-\partial_{Z_T}\hat{L}(h(S_T^n), Z_T(c^*)) \times \partial_{Z_t} Z_T(c^*)\Big).
\end{aligned}$$

But as, for $t \in [0,T]$, the dual control is given by $\gamma_t = c \times Z_t \times \sigma(\hat{\nu}_t), c \in U$, we have that

$$dZ_t(c) = Z_t(c)\bigg(-r dt - \frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}}dW_S(t) + c\,\sigma(\hat{\nu}_t)dB_v(t)\bigg)$$

and therefore, for $0 \leq t \leq T$,

$$\frac{Z_T(c)}{Z_t(c)} = e^{-r(T-t)} \times \underbrace{\exp\bigg(-\frac{1}{2}\int_t^T \Big(\frac{(\hat{\mu}_s - r)^2}{\hat{\nu}_s} + c^2\sigma^2(\hat{\nu}_s)\Big)ds - \int_t^T \frac{(\hat{\mu}_s - r)}{\sqrt{\hat{\nu}_s}}dW_S(s) + \int_t^T c\,\sigma(\hat{\nu}_s)dB_v(s)\bigg)}_{:= \frac{d\mathbb{Q}(c)}{d\mathbb{P}}|_{\mathcal{F}_t}}.$$

Since $\bigg(\frac{(\mu_t - r)}{\sqrt{\hat{\nu}_t}},\ c\,\sigma(\hat{\nu}_t)\bigg)_{0 \leq t \leq T}$ is a 2 dimensional vector of adapted and square integrable processes, using Girsanov's Theorem, we can define $\mathbb{P}-$equivalent probability measures $(\mathbb{Q}(c))_{c \in U}$ with change of measure defined by $\frac{d\mathbb{Q}(c)}{d\mathbb{P}}|_{\mathcal{F}_t}$ such that

$$\begin{aligned}
dW_s^{\mathbb{Q}(c)}(t) =& dW_s(t) + \frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}}dt, \\
dB_v^{\mathbb{Q}(c)}(t) =& dB_v(t) - c\,\sigma(\hat{\nu}_t)dt.
\end{aligned}$$

Therefore, dynamics of processes under the $\mathbb{Q}(c)-$measure are given by

$$dS_t^n = rS_t^n + \sqrt{\hat{\nu}_t}S_t^n dW_s^{\mathbb{Q}(c)}(t),$$

$$d\nu_t^i = \Big(- x_i\nu_t^i + b(\hat{\nu}_t) + \sigma(\hat{\nu}_t)\Big(- \rho\frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}} + \sqrt{1-\rho^2}c\,\sigma(\hat{\nu}_t)\Big)\Big)dt + \sigma(\hat{\nu}_t)dW_v^{\mathbb{Q}(c)}(t),\ i = 1, ..., n,$$

$$dZ_t = Z_t\bigg(\Big(- r + \frac{(\hat{\mu}_t - r)^2}{\hat{\nu}_t} + c^2\sigma^2(\hat{\nu}_t)\Big)dt - \frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}}dW_S^{\mathbb{Q}}(t) + c\,\sigma(\hat{\nu}_t)dB_v^{\mathbb{Q}(c)}(t)\bigg)$$

where $W_s^{\mathbb{Q}(c)}$, $W_v^{\mathbb{Q}(c)}$ are standard brownian motions under $\mathbb{Q}(c) - measure$ with $d < W_s^{\mathbb{Q}(c)}, W_v^{\mathbb{Q}(c)} >_t = \rho\, dt$ and $B_v^{\mathbb{Q}(c)}$ is a standard brownian motion, independent from $W_s^{\mathbb{Q}(c)}$ defined such that for $t \in [0, T]$,

$$W_v^{\mathbb{Q}(c)}(t) = \rho W_s^{\mathbb{Q}(c)}(t) + \sqrt{1-\rho^2}B_v^{\mathbb{Q}(c)}(t).$$

In this case we have that

$$\begin{aligned}
\bar{V}_t^{\mathcal{U}} &= E_{t,s,\boldsymbol{\nu},z}\Big(- \partial_{Z_T}\hat{L}(h(S_T^n), Z_T(c^*)) \times \partial_{Z_t}Z_T(c^*)\Big) \\
&= E_{t,s,\boldsymbol{\nu},z}\Big(- \partial_{Z_T}\hat{L}(h(S_T^n), Z_T(c^*)) \times \frac{Z_T(c^*)}{Z_t(c^*)}\Big) \\
&= E_{t,s,\boldsymbol{\nu},z}\Big(e^{-r(T-t)} - \partial_{Z_T}\hat{L}(h(S_T^n), Z_T(c^*)) \times \frac{d\mathbb{Q}(c^*)}{d\mathbb{P}}|_{\mathcal{F}_t}\Big) \\
&= E_{t,s,\boldsymbol{\nu},z}^{\mathbb{Q}(c^*)}(e^{-r(T-t)}\Big(- \partial_{Z_T}\hat{L}(h(S_T^n), Z_T(c^*))\Big)\Big).
\end{aligned}$$

Moreover, if we consider a power loss, we know that

$$\hat{L}(h(S_T^n), Z_T(c^*)) = -\frac{Z_T^q(c^*)}{q} - H_T^n Z_T(c^*),$$

we deduce that in this case,

$$\partial_{Z_T}\hat{L}(h(S_T^n), Z_T(c^*)) = -Z_T^{q-1}(c^*) - H_T^n,$$

Therefore, we obtain that

$$\bar{V}_t^{\mathcal{U}} = E_{t,s,\boldsymbol{\nu},z}^{\mathbb{Q}(c^*)}\Big(e^{-r(T-t)}\Big(H_T^n + Z_T^{q-1}(c^*)\Big)\Big).$$

□

Still assuming that the subset of admissible dual controls $\mathcal{U}$ has the form (33), we next show that, for the rough Heston model, the lower bond as well as the approximate hedging ratio associated to a power loss function have explicit forms.

Proposition 18. *Consider a power loss of the form $L(h(s), v) = \frac{1}{p}(h(s) - v)^p$, suppose that the volatility is modeled by a rough Heston model such that $b(x) = \kappa(\theta - x)$ and $\sigma(x) = \zeta\sqrt{x}$. Moreover, assume that the subset of admissible dual control is given by (33). Therefore $Y(t, s, \boldsymbol{\nu}, z; \gamma)$ defined by (30) is such that*

$$Y(t, s, \boldsymbol{\nu}, z; c) = -\frac{1}{q}z^q \exp(C_t(c) + \sum_{i=1}^n D_t^i(c)\nu_t^i) - E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c)}(e^{-r(T-t)}H_T^n)\, z,$$

where $C_t(c)$ and $(D_t(c))_{i=1,..,n}$ are time-dependent functions, solutions of Riccati ODEs given respectively by

$$\begin{aligned}
\partial_t C_t(c) &= r\,q - \frac{1}{2}\,q(q-1)\,(A^2 + c^2)\nu_0 - \sum_{i=1}^n D_t^i\Big(\kappa(\theta - \nu_0) + q\zeta\,\nu_0\Big(- \rho A + \sqrt{1-\rho^2}c\Big)\Big) \\
&\quad - \frac{1}{2}\nu_0\zeta^2\sum_{i=1}^n\sum_{j=1}^n D_t^i D_t^j, \\
C_T(c) &= 0,
\end{aligned}$$

and for $i = 1, ..., n$,

$$\begin{aligned}\partial_t D^i_t(c) &= x_i D^i_t + w_i \sum_{j=1}^n D^j_t \Big(\kappa - q\Big(-\rho A\zeta + \sqrt{1-\rho^2}\zeta c\Big)\Big) - \frac{1}{2} w_i \zeta^2 \sum_{j=1}^n \sum_{k=1}^n D^j_t D^k_t \\ &\quad - \frac{1}{2} w_i\, q(q-1)\,(A^2 + c^2), \\ D^i_T(c) &= 0.\end{aligned}$$

In this case, the lower bound satisfies

$$\underline{l}^{\mathcal{U}}_n(t,s,\boldsymbol{\nu},v) = -\frac{1}{q} Z^q_t(c^*) \exp(C_t(c^*) + \sum_{i=1}^n D^i_t(c^*)\nu^i_t) - E^{\mathbb{Q}(c^*)}_{t,s,\boldsymbol{\nu}}(e^{-r(T-t)} H^n_T) Z_t(c^*) + Z_t(c^*) \bar{V}^{\mathcal{U}}_t, \tag{36}$$

with

$$Z_t(c^*) = = \exp\Big(-\frac{1}{q-1}(C_t(c^*) + \sum_{i=1}^n D^i_t(c^*)\nu^i_t)\Big) \times \Big(\bar{V}^{\mathcal{U}}_t - E^{\mathbb{Q}(c^*)}_{t,s,\boldsymbol{\nu}}(e^{-r(T-t)} H^n_T))\Big)^{\frac{1}{q-1}}.$$

Moreover, if $z = Z_t(c^*)$,

$$\hat{\underline{l}}^{\mathcal{U}}_n(t,s,\boldsymbol{\nu},z) = Y(t,s,\boldsymbol{\nu},z;c^*).$$

Proof. The proof is similar to the proof of Proposition 13 and is essentially obtained by using the Feynman-Kac formula. First, using Lemma 12, we have that

$$\begin{aligned}Y(t,s,\boldsymbol{\nu},z;c) &= E_{t,s,\boldsymbol{\nu},z}\Big(\hat{L}(h(S^n_T), Z_T)\Big) \\ &= E_{t,s,\boldsymbol{\nu},z}\Big(-\frac{1}{q} Z^q_T - H^n_T Z_T\Big) \\ &= \underbrace{E_{t,\boldsymbol{\nu},z}\Big(-\frac{1}{q} Z^q_T\Big)}_{:=Y_1(t,\boldsymbol{\nu},z)} - \underbrace{E_{t,s\boldsymbol{\nu},z}\Big(H^n_T Z_T\Big)}_{:=Y_2(t,s,\boldsymbol{\nu},z)}.\end{aligned}$$

Let focus on Y_1, using Feynman-Kac formula, we have that

$$\begin{aligned}0 =& \partial_t Y_1 - \partial_z Y_1\, rZ + \frac{1}{2}\partial_{zz} Y_1\, Z^2\Big(\frac{(\hat{\mu}-r)^2}{\hat{\nu}} + c^2\hat{\nu}\Big) + \sum_{i=1}^n \partial_{\nu_i} Y_1\,(-x_i\nu_i + \kappa(\theta - \hat{\nu})) \\ &+ \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \partial_{\nu_i\nu_j} Y_1\,\zeta^2\hat{\nu} + \sum_{i=1}^n \partial_{\nu_i z} Y_1\, Z\Big(-\rho\frac{(\hat{\mu}-r)}{\sqrt{\hat{\nu}}}\zeta\sqrt{\hat{\nu}} + \sqrt{1-\rho^2}\zeta\hat{\nu}c\Big), \\ Y_1(T,Z) =& -\frac{1}{q} Z^q.\end{aligned}$$

Suppose that Y_1 has the following form

$$Y_1 = -\frac{1}{q} Z^q_t \exp(C_t + \sum_{i=1}^n D^i_t \nu^i_t),$$

In this case, as we consider that $\hat{\mu}_t = r + A\hat{\nu}_t$, we have that

$$\begin{aligned}0 = Y_1\Big(&\partial_t C_t + \sum_{i=1}^n \partial_t D^i_t \nu_i + \sum_{i=1}^n D^i_t(-x_i\nu_i + \kappa(\theta - \hat{\nu})) + \zeta^2\hat{\nu}\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n D^i_t D^j_t \\ &- rq + \frac{1}{2}q(q-1)\,(A^2+c^2)\hat{\nu} + q\Big(-\rho A\hat{\nu}\zeta + \sqrt{1-\rho^2}\zeta\hat{\nu}c\Big)\sum_{i=1}^n D^i_t\Big),\end{aligned}$$

as $\hat{\nu} = \nu_0 + \sum_{i=1}^n w_i \nu_i$, we have

$$
\begin{aligned}
0 =& \Bigg(\partial_t C_t + \sum_{i=1}^n D_t^i \Bigg(\kappa(\theta - \nu_0) + q\Bigg(-\rho A \nu_0 \zeta + \sqrt{1-\rho^2} \zeta \nu_0 c \Bigg) \Bigg) - rq + \frac{1}{2} q(q-1)\,(A^2 + c^2)\nu_0 \\
&+ \frac{1}{2} \zeta^2 \nu_0 \sum_{i=1}^n \sum_{j=1}^n D_t^i D_t^j \Bigg) \\
&+ \sum_{i=1}^n \nu_i \,\Bigg(\partial_t D_t^i - x_i D_t^i - w_i \kappa \sum_{j=1}^n D_t^j + \zeta^2 w_i \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n D_t^j D_t^k \\
&+ \frac{1}{2} w_i q(q-1)\,(A^2 + c^2) + q w_i \Bigg(-\rho A \zeta + \sqrt{1-\rho^2} \zeta c \Bigg) \sum_{j=1}^n D_t^j \Bigg) \Bigg).
\end{aligned}
$$

We obtain that C_t and $(D_t^i)_{i=1,\ldots,n}$ solve Riccati type ODEs of the form

$$
\begin{aligned}
\partial_t C_t =& rq - \frac{1}{2} q(q-1)\,(A^2 + c^2)\nu_0 - \sum_{i=1}^n D_t^i \Bigg(\kappa(\theta - \nu_0) + q\Bigg(-\rho A \nu_0 \zeta + \sqrt{1-\rho^2} \zeta \nu_0 c \Bigg) \Bigg) && (37) \\
&- \frac{1}{2} \zeta^2 \nu_0 \sum_{i=1}^n \sum_{j=1}^n D_t^i D_t^j, && (38) \\
C_T =& 0
\end{aligned}
$$

and for $i = 1, ..., n$,

$$
\begin{aligned}
\partial_t D_t^i =& x_i D_t^i + w_i \sum_{j=1}^n D_t^j \Bigg(\kappa - q\Bigg(-\rho A \zeta + \sqrt{1-\rho^2} \zeta c \Bigg) \Bigg) - \zeta^2 w_i \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n D_t^j D_t^k \\
&- \frac{1}{2} w_i q(q-1)\,(A^2 + c^2), && (39) \\
D_T^i =& 0.
\end{aligned}
$$

Let now consider the second process Y_2. Using the Feynman-Kac formula, we obtain that Y_2 satisfies

$$
\begin{aligned}
0 =& \partial_t Y_2 - \partial_z Y_2\, rZ + \frac{1}{2} \partial_{zz} Y_2\, Z^2 \Bigg(\frac{(\hat{\mu} - r)^2}{\hat{\nu}} + c^2 \hat{\nu} \Bigg) + \sum_{i=1}^n \partial_{\nu_i} Y_2 \,\Bigg(-x_i \nu_i + \kappa(\theta - \hat{\nu}) \Bigg) \\
&+ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \partial_{\nu_i \nu_j} Y_2\, \zeta^2 \hat{\nu} + \sum_{i=1}^n \partial_{\nu_i z} Y_2\, Z \Bigg(-\rho \frac{(\hat{\mu} - r)}{\sqrt{\hat{\nu}}} \zeta \sqrt{\hat{\nu}} + \sqrt{1-\rho^2} \zeta \hat{\nu} c \Bigg), \\
&+ \partial_s Y_2\, \hat{\mu} s + \frac{1}{2} \partial_{ss} Y_2\, \hat{\nu} s + \sum_{i=1}^n \partial_{\nu_i s} Y_2\, s \hat{\nu} \zeta \rho - \partial_{sz} Y_2\, Z \frac{\hat{\mu} - r}{\sqrt{\hat{\nu}}} s \sqrt{\hat{\nu}}, \\
Y_2(T, s, z) =& h(s) z.
\end{aligned}
$$

Suppose that Y_2 has the form

$$
Y_2 = g(t, S_t^n, \boldsymbol{\nu}) Z_t.
$$

Therefore, we have that

$$
\begin{aligned}
0 = Z \Bigg(& \partial_t g - g\, r + \sum_{i=1}^n \partial_{\nu_i} g \,\Bigg(-x_i \nu_i + \kappa(\theta - \hat{\nu}) \Bigg) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \partial_{\nu_i \nu_j} g\, \zeta^2 \hat{\nu} \\
&+ \sum_{i=1}^n \partial_{\nu_i} g \,\Bigg(-\rho(\hat{\mu} - r)\zeta + \sqrt{1-\rho^2} \zeta^2 \hat{\nu} c \Bigg) + \partial_s g\, \hat{\mu} s + \frac{1}{2} \partial_{ss} g\, \hat{\nu} s \\
&+ \sum_{i=1}^n \partial_{\nu_i s} g\, s \hat{\nu} \zeta \rho - \partial_s g\, (\hat{\mu} - r) s \Bigg).
\end{aligned}
$$

The function $g(t,S_t^n,\boldsymbol{\nu})$ satisfies the following PDE

$$\begin{aligned}
0 &= \partial_t g - g\,r + \partial_s g\, rs + \sum_{i=1}^n \partial_{\nu_i} g\left(- x_i v_i + \kappa(\theta - \hat\nu) + \hat\nu\zeta(-\rho A + \sqrt{1-\rho^2}\zeta\, c)\right) \\
&+ \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \partial_{\nu_i\nu_j} g\, \zeta^2\hat\nu + \frac{1}{2}\partial_{ss} g\, \hat\nu s + \sum_{i=1}^n \partial_{\nu_i s} g\; s\hat\nu\sigma\rho, \\
g(T,s,\boldsymbol{\nu}) &= h(s).
\end{aligned}$$

We deduce, using once again the Feynman-Kac theorem (this time in the other sense), that

$$g(t,S_t^n,\hat\nu_t) = E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c)}(e^{-r(T-t)}H_T^n),$$

with under $\mathbb{Q}(c)-$measure,

$$dS_t^n = rS_t^n + \sqrt{\hat\nu_t}S_t^n dW_s^{\mathbb{Q}(c)}(t),$$
$$d\nu_t^i = \left(- x_i(\nu_t^i - \nu_0^i) + \kappa(\theta - \hat\nu) + \hat\nu\zeta(-\rho A + \sqrt{1-\rho^2}\zeta\, c)\right)dt + \zeta\sqrt{\hat\nu_t}dW_v^{\mathbb{Q}}(t),\ i = 1,...,n,$$

where $W_s^{\mathbb{Q}(c)}$ and $W_v^{\mathbb{Q}(c)}$ are standard brownian motions under $\mathbb{Q}(c) - measure$ with $d < W_s^{\mathbb{Q}(c)}, W_v^{\mathbb{Q}(c)} >_t = \rho\, dt$.
Finally, combining the different results, we obtain the annonced result

$$Y(t,s,\boldsymbol{\nu},z;c) = -\frac{1}{q}z^q \exp(C_t(c) + \sum_{i=1}^n D_t^i(c)\nu_t^i) - E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c)}(e^{-r(T-t)}H_T^n)\, z.$$

Furthermore, as by definition,

$$\underline{l}_n(t,s,\boldsymbol{\nu},v;c) = \sup_z \left\{ Y(t,s,\boldsymbol{\nu},z;c) + zv \right\},$$

using the first order condition, we obtain that the value at time t of $Z_t(c)$ satisfies

$$Z_t(c) = \exp\left(- \frac{1}{q-1}(C_t(c) + \sum_{i=1}^n D_t^i(c)\nu_t^i)\right) \times \left(\bar V_t^{\mathcal{U}} - E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c)}(e^{-r(T-t)}H_T^n)\right)^{\frac{1}{q-1}}$$

and then

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v;c) = -\frac{1}{q}Z_t^q(c)\exp(C_t(c) + \sum_{i=1}^n D_t^i(c)\nu_t^i) - E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c)}(e^{-r(T-t)}H_T^n)\, Z_t(c) + Z_t(c)\,\bar V_t^{\mathcal{U}}.$$

Thus, as

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) = \underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v;c^*),$$

we deduce the annonced result. Finally, as in the proof of Proposition 17 , we show that $c^* = c^*(z^*)$ and $z^*(c^*) = z^*$, we conclude that if $z = Z_t(c^*)$ then

$$\underline{\tilde l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) = Y(t,s,\boldsymbol{\nu},z;c^*).$$

□

Remark. Comparing to the optimal solution in the case of linear payoff (see Proposition 13), we observe that the form of the solution is the same, the only difference lies in the expression of the coefficients C_t and$(D_t^i)_{i=1,\dots,n}$. Moreover, in the classical Heston model, i.e. $n=1$, then C_t and D_t admit closed formulas.

Proposition 19. *Using the same assumptions as the Proposition 18 and denoting $g(t,S_t^n,\boldsymbol{\nu}_t) = E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c^*)}(e^{-r(T-t)}H_T^n)$, the approximate hedging ratio $\bar\xi_n^{\mathcal{U}}(t)$ defined by (32) is such that*

$$\begin{aligned}
\bar\xi_n^{\mathcal{U}}(t) &= \frac{1}{\hat\nu_t S_t^n}\left(g(t,S_t^n,\boldsymbol{\nu}_t) - \bar V_t^{\mathcal{U}}\right)\left(\frac{1}{p-1}(\hat\mu_t - r) - \rho\zeta\hat\nu_t\sum_{i=1}^n D_t^i(c^*)\right) \\
&+ \partial_s g(t,S_t^n,\boldsymbol{\nu}_t) + \frac{\rho\zeta}{S_t^n}\sum_{i=1}^n\left(\partial_{\nu_i} g(t,S_t^n,\boldsymbol{\nu}_t)\right),
\end{aligned} \tag{40}$$

where $\bar V_t^{\mathcal{U}}$ is the value at time t of the self-financing portfolio.

Proof. The proof is simple and relies on the definition of the approximate hedging ratio and on Proposition 18. By definition, we have that

$$\bar{\xi}_n^{\mathcal{U}}(t) = \frac{z\,\partial_{zz}\underline{l}_n^{\mathcal{U}}\,(\hat{\mu}-r)s - \partial_{sz}\underline{l}_n^{\mathcal{U}}\,\hat{\nu}s^2 - \rho\sum_{i=1}^n \partial_{\nu_i z}\underline{l}_n^{\mathcal{U}}\,\sqrt{\hat{\nu}}s\,\sigma(\hat{\nu})}{\hat{\nu}s^2},$$

with $z = z^* = z^*(c^*)$. With our assumptions, we know, by Proposition 18, that, for $z = z^*(c^*)$,

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) = -\frac{1}{q}z^q \exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i) - E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c^*)}(e^{-r(T-t)}H_T^n)\,z$$

thus, we can easily deduce the partial derivatives of the process $\underline{l}_n^{\mathcal{U}}(.)$ such that

$$\partial_z\underline{l}_n^{\mathcal{U}} = -z^{q-1}(c^*)\exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i) - g(t,S_t^n,\boldsymbol{\nu}_t),$$

and

$$\begin{aligned}
\partial_{zz}\underline{l}_n^{\mathcal{U}} &= -\,(q-1)z^{q-2}\exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i)\\
\partial_{zs}\underline{l}_n^{\mathcal{U}} &= -\,\partial_s g(t,S_t^n,\boldsymbol{\nu}_t)\\
\partial_{z\nu_i}\underline{l}_n^{\mathcal{U}} &= -\,z^{q-1}D_t^i(c^*)\exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i) - \partial_{\nu_i}g(t,S_t^n,\boldsymbol{\nu}_t),
\end{aligned}$$

therefore, the approximate hedging ratio is given by

$$\begin{aligned}
\bar{\xi}_n^{\mathcal{U}}(t) = \frac{1}{\hat{\nu}_t\Big(S_t^n\Big)^2}\Bigg(&-(q-1)Z_t^{q-1}\exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i)\,(\hat{\mu}_t - r)S_t^n + \partial_s g(t,S_t^n,\boldsymbol{\nu}_t)\,\hat{\nu}_t S_t^{n\,2}\\
&+\rho\zeta S_t^n\hat{\nu}_t\sum_{i=1}^n\Big(Z_t^{q-1}D_t^i(c^*)\exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i) + \partial_{\nu_i}g(t,S_t^n,\boldsymbol{\nu}_t)\Big)\Bigg).
\end{aligned} \tag{41}$$

The value at time t of the process Z_t is given by

$$Z_t(c^*) = \exp\Big(-\frac{1}{q-1}(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i)\Big)\times\Big(\bar{V}_t^{\mathcal{U}} - g(t,S_t^n,\boldsymbol{\nu}_t)\Big)^{\frac{1}{q-1}},$$

plugging $Z_t(c^*)$ in (41), the approximate hedging ratio becomes

$$\begin{aligned}
\bar{\xi}_n^{\mathcal{U}}(t) = \frac{1}{\hat{\nu}_t S_t^n}&\Big(g(t,S_t^n,\boldsymbol{\nu}_t) - \bar{V}_t^{\mathcal{U}}\Big)\Big(\frac{1}{p-1}(\hat{\mu}_t - r) - \rho\zeta\hat{\nu}_t\sum_{i=1}^n D_t^i(c^*)\Big)\\
&+\partial_s g(t,S_t^n,\boldsymbol{\nu}_t) + \frac{\rho\zeta}{S_t^n}\sum_{i=1}^n\Big(\partial_{\nu_i}g(t,S_t^n,\boldsymbol{\nu}_t)\Big).
\end{aligned}$$

That concludes the proof since we proved the different results announced. □

For the partial hedging problem in rough Heston model with power loss function, the lower bound $\underline{l}_n^{\mathcal{U}}(.)$ and the approximate hedging ratio $\bar{\xi}_n^{\mathcal{U}}(.)$ admit closed formulas. In particular, we observe that $\bar{\xi}_n^{\mathcal{U}}(.)$ is split into three parts: the first part is linked to a sharpe ratio, the second part corresponds to the delta and the third part is linked to a vega. Moreover, as the approximate rough Heston model allows a closed form for the characteristic function of the log-price, it enables, using a FFT (Fast Fourier Transform) pricing method, to efficiently compute the price but also the Greeks of vanilla options under the approximate rough Heston model. Thus, for vanilla options, the lower bound as well as the approximate hedging ratio can be quickly computed. Note that the characteristic function of the log-price is presented in the Appendix.

6 Numerical results

In this section, we illustrate the partial hedging method discussed in this paper for the rough Heston model. We first present stylized facts about the approximate rough Heston model, then we discuss results for hedging of linear payoff and finally we consider vanilla option hedging. For the different numerical results, inspired by [2], we have decided to consider the following parameters (under the real measure $\mathbb{P}$) for the rough Heston model:

$$r = 0.02,\ A = 1,\ S_0 = 100,\ \nu_0 = 0.04,\ \theta = 0.04,\ \lambda = 0.3,\ \zeta = 0.3\ and\ \rho = -0.7.$$

As already mentioned, rough volatility models, including the rough Heston model, allow to better model stylized facts observed on the financial markets. In particular, in option pricing, they allow to better represent the implied volatility smile as well as the ATM (at-the-money) skew. Using the characteristic function of the log-price in the approximate rough Heston, we can price vanilla options using a FFT method. From these prices, we can deduce the implied volatility for several strikes and time-to-maturity but also compute the ATM skew i.e. the derivative of the ATM implied volatility with respect to the log strikes. Figures 5 and 6[2] in Appendix represent respectively the implied volatility and the ATM skew. If we compare the Heston model to the approximate rough Heston model with $n = 20$, we observe that for the Heston model, the shape of the implied volatility smile does not change significantly as the time-to-maturity decreases, whereas for the approximate rough Heston model, the volatility smile becomes more and more pronounced as the time-to-maturity tends to 0. We also observe that the approximate rough volatility model with $n = 20$ captures the ATM skew while the Heston model does not capture this stylized fact at all. Therefore, as already shown in [1], we see that even with a reduced number of factors (here $n = 20$), the approximate rough Heston model allows to better capture stylized effects on the financial markets and is thus more appropriate to model stocks dynamics. For more analysis about the rough Heston and its Markov approximation, we refer among others to [1, 2, 14, 25].

6.1 Partial hedging of linear payoff

In this section, we consider the hedging of a linear payoff with a quadratic loss function. We take this toy case because we have shown that with a quadratic loss, the approximate partial hedging problem is solved with a closed formula. Thus, the objective of this paragraph is to check the convergence of the approximate solution when $n \to +\infty$ but also to verify that the lower and upper bounds of the optimal solution deduced using the dual control method are close enough to this optimal solution. The linear payoff considered for the following is

$$H_T^{linear} = \alpha + \beta S_T\ with\ \alpha = 0,\ \beta = 1,$$

and its replication price at time $t \in [0, T]$ is

$$H_t^{linear} = e^{-r(T-t)}\alpha + \beta S_t.$$

We will thus consider several initial values of the hedging portfolio such as

$$V_0 \leq H_0^{linear}.$$

We first look at the convergence of the approximate optimal solution l_n when $n \to +\infty$. Figure 1 and 2 present the evolution of l_n and ξ_0^n with respect to n for different values of H and with $V_0 = 0.8 \times H_0^{linear}$. We observe that for different Hurst coefficients, the approximate solution and initial hedging ratio converge and the convergence is faster the closer the Hurst coefficient is to 0.5. This is logical because the closer H is to 0.5, the less n factors are needed to model the rough volatility process accurately.

[2]Results are generated with real measure parameters

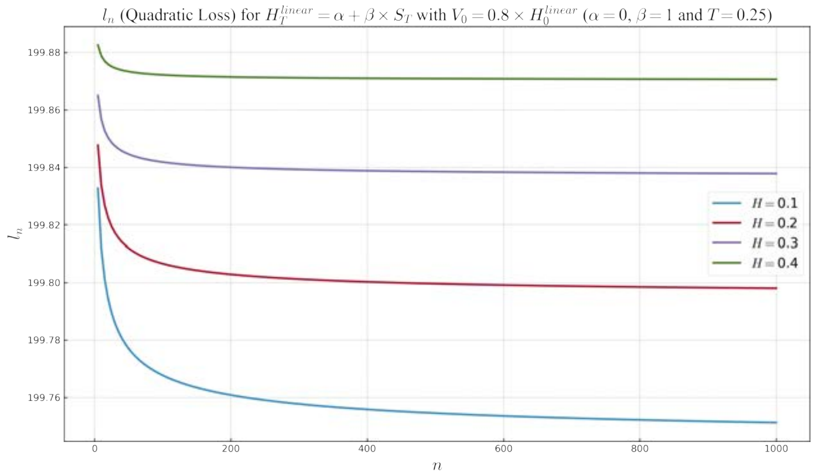

Figure 1: Convergence analysis of l_n for quadratic loss and linear payoff

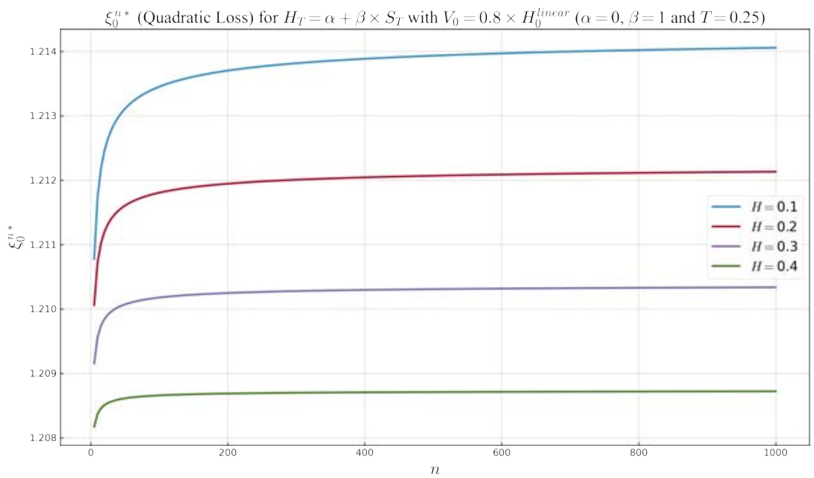

Figure 2: Convergence analysis of $\xi_0^{n\,*}$ for quadratic loss and linear payoff

We next compare the optimal solution with the upper and lower bounds deduced by the dual control method. As in the theoretical part, we consider dual controls of the form:

$$\gamma_t = c \times Z_t \times \zeta \sqrt{\hat{\nu}_t},\ c \in U = \mathbb{R}, \tag{42}$$

because it allows to keep closed forms under the approximate rough Heston model. In this case, the associated lower bound is given by (36) and the upper bound can be computed with the hedging ratio given by (40).

$Coef$	H	l_n	$\underline{l}_n^{\mathcal{U}}$	$\overline{l}_n^{\mathcal{U}}$	Abs. diff. $\lvert l_n - \underline{l}_n^{\mathcal{U}}\rvert$	Abs. diff. $\lvert \overline{l}_n^{\mathcal{U}} - l_n\rvert$	Abs. diff. $\lvert \overline{l}_n^{\mathcal{U}} - \underline{l}_n^{\mathcal{U}}\rvert$
0.75	0.1	312.1792	312.1786	312.1890	6.26×10^{-4}	9.85×10^{-3}	1.04×10^{-2}
0.75	0.5	312.3396	312.3371	312.3428	2.51×10^{-3}	3.23×10^{-3}	5.76×10^{-3}
0.5	0.1	1248.7169	1248.7145	1248.7213	1.66×10^{-3}	5.26×10^{-3}	6.87×10^{-3}
0.5	0.5	1249.3587	1249.3534	1249.3654	5.32×10^{-3}	6.71×10^{-3}	1.21×10^{-2}
0.25	0.1	2809.6131	2809.6078	2809.6218	5.33×10^{-3}	8.76×10^{-3}	1.44×10^{-2}
0.25	0.5	2811.0571	2811.0530	2811.0638	4.19×10^{-3}	6.74×10^{-3}	1.08×10^{-2}

Table 1: Comparison of the optimal solution, the lower bound and the upper bound for quadratic loss and linear payoff with $n = 20$, $V_0 = Coef \times H_0^{linear}$and $T = 0.25$

Table 1 compares the results for different values of V_0. Firstly, we observe that the optimal solution, the lower and upper bounds are very close, as we notice an absolute error that varies between $O(10^{-4})$ and $O(10^{-2})$, this allows to validate the relevance of the bounds deduced by using the dual control method. We also notice that the closer V_0 is to the replication price, the more the quadratic loss decreases and finally we observe that the quadratic loss is less when considering $H = 0.1$ compared to $H = 0.5$. This seems to indicate that for a linear payoff, the rougher is the volatility, the lower is the quadratic loss. The orders of the absolute errors between the bounds and the primal solutions are consistent with the order of the absolute errors made in applying the dual control method to portfolio optimization problems as in [31, 32]. This can be explained by the fact that the problem reduces to a portfolio optimization problem when considering partial hedging of linear payoff. Indeed, assuming for simplicity that $r = 0$, we can easily observe that our problem has the form of a portfolio optimization problem such that

$$l_n(t, s, \boldsymbol{\nu}, v) = \inf_{\xi_t^n \in \mathcal{R}_n} E_{t,s,\boldsymbol{\nu},v}\left(\frac{1}{2}\left(S_t^n - V_t + \int_t^T (1 - \xi_s^n) dS_s^n\right)^2\right).$$

Therefore, we observe that in the case of a linear payoff, there is no stochastic target to reach, which is not the case if we consider nonlinear payoffs like for example vanilla options.

6.2 Partial hedging of vanilla options

We now focus on more relevant payoff, namely vanilla options. As for the linear case, we take a quadratic loss function. Notice that we only consider Call options, but similar results can be deduced for Put options. For this type of payoff, we cannot derive an optimal solution to the partial hedging problem, but by using the dual control method, we can derive upper and lower bounds. To do this, for the same reasons as in the previous section, we consider dual controls of the form (42). Moreover, as the characteristic function of the log-price in the approximate rough Heston model is available, the bounds as well as the approximate hedge ratio are computed using a FFT method. For our numerical results, we decide to consider an initial portfolio value V_0 proportional to the Black-Scholes (BS) price of a Call option with as constant volatility, the mean-reverting level of the rough model. Thus, we consider that

$$V_0 = Coef \times BS_{call}(\sigma = \sqrt{\theta}).$$

Furthermore, as in [29], we benchmark the proposed partial hedging strategy with the optimal strategy assuming a constant volatility model of the BS type. Notice that for the BS model, the optimal hedge ratio denoted by ξ_t^{BS} associated with a quadratic loss has an explicit formula that is presented in the Appendix (see (50)) and has a similar form to $\bar{\xi}_n^{\mathcal{U}}(t)$ except that it does not depend on the volatility process. Figure 3 presents the initial approximate hedge ratio $\bar{\xi}_n^{\mathcal{U}}(0)$ as function of the number of factors n, for different values of H. As in the linear payoff case, we notice a convergence of the hedge ratios. The speed of convergence is higher, the higher is the Hurst coefficient, for the same reason as mentioned previously.

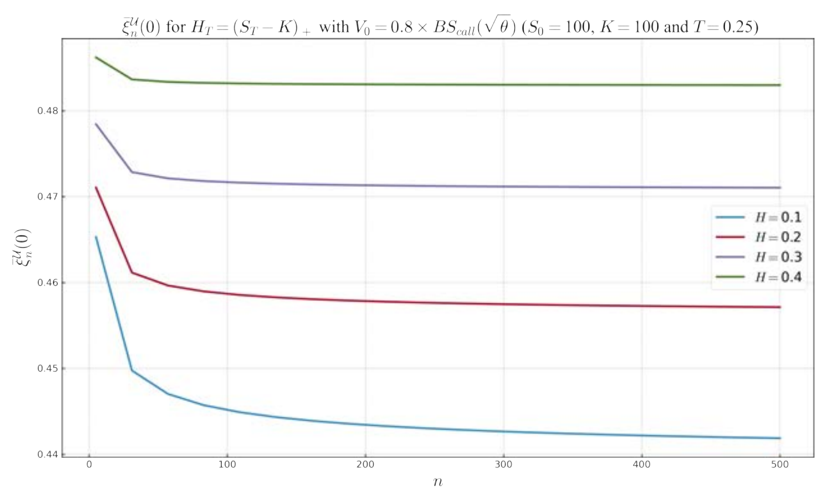

Figure 3: Convergence analysis for $\bar{\xi}_n^{\mathcal{U}}(0)$ for quadratic loss and ATM call

Figure 4 compares the sub-optimal strategy linked to the upper bound with the benchmark strategy for one simulation of $(S_t^n)_{0\leq t\leq T}$. We observe in this case that the sub-optimal strategy performs much better than the benchmark strategy.

Table 2 compares the results for different values of V_0. In contrast to the linear payoff, the gap between the upper and lower bounds is more pronounced. There are several reasons explaining this. As the vanilla options are nonlinear payoffs, the partial hedging problem cannot be reduced to a portfolio optimization problem. Therefore, we have a stochastic target leading to a more noisy problem. In this case, we notice that the approximate hedging ratio depends more on the choice of the sub-optimal dual control. In fact, in contrast with the linear case, we remark that the approximate hedging ratio (40) is affected by the choice of dual control, notably through the risk-neutral measure $\mathbb{Q}(c^*)$ used to compute the Greeks. The difference between the bounds can also be explained by the tracking error i.e. the error made by not continuously hedging the portfolio. Indeed, we observe in Table 3 that by increasing the frequency of portfolio rebalancing, the gap between the bounds decreases.

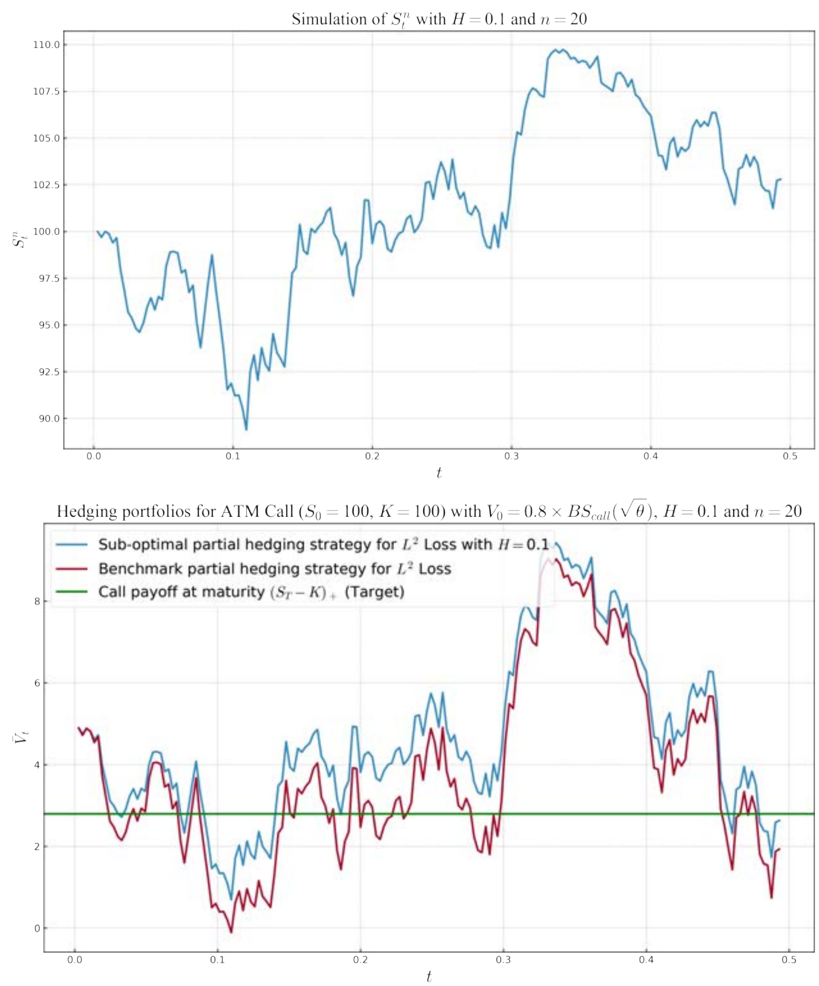

Figure 4: Simulation of hedging strategies with $V_0 = 0.8 \times BS_{call}(\sqrt{\theta})$

Benchmarking the sub-optimal strategy against the Black-Scholes strategy, we observe at Tables 2 and 3 that the hedging strategy associated with the upper bound outperforms the benchmark strategy, and the gap between the two strategies is much more pronounced when $H = 0.1$. Thus, the rougher is the volatility trajectory, the worse the benchmark strategy will perform compared to the sub-optimal strategy. As revealed by Table 4, this is explained by the fact that the smaller H is, the greater is the impact of the vega term on the approximate hedging ratio and therefore the bigger is the difference between the approximate hedging ratio $\bar{\xi}_n^{\mathcal{U}}(t)$ and the hedging ratio of the benchmark strategy ξ_t^{BS}.

$Coef$	H	$\underline{l}_n^{\mathcal{U}}$	$\bar{l}_n^{\mathcal{U}}$	$l_{BS(\sqrt{\theta})}$	Abs. Diff $\lvert \bar{l}_n^{\mathcal{U}} - \underline{l}_n^{\mathcal{U}} \rvert$	Abs. Diff. $\lvert l_{BS(\sqrt{\theta})} - \bar{l}_n^{\mathcal{U}} \rvert$
1	0.1	0.1769	0.6139	0.7893	0.4369	0.1753
1	0.5	0.0488	0.2449	0.2791	0.1962	0.0341
0.75	0.1	0.7781	0.9389	1.0985	0.1608	0.1596
0.75	0.5	0.6261	0.6851	0.7173	0.0588	0.0323
0.5	0.1	2.2267	2.3112	2.4252	0.0844	0.1141
0.5	0.5	2.2469	2.2704	2.3129	0.0235	0.0424
0.25	0.1	4.7591	4.7928	4.8965	0.0338	0.1036
0.25	0.5	4.9814	4.9902	5.0224	0.0088	0.0322

Table 2: Comparison of the lower bound, the upper bound and the benchmark ($l_{BS(\sqrt{\theta})}$ with $\sqrt{\theta} = 0.2$) for quadratic Loss, ATM Call and hedging once a day with $n = 20$, $V_0 = Coef \times BS_{Call}(\sigma = \sqrt{\theta})$ and $T = 0.25$

$Coef$	Hedging Freq.	$\underline{l}_n^{\mathcal{U}}$	$\bar{l}_n^{\mathcal{U}}$	$l_{BS(\sqrt{\theta})}$	Abs. Diff $\lvert \bar{l}_n^{\mathcal{U}} - \underline{l}_n^{\mathcal{U}} \rvert$	Abs. Diff. $\lvert l_{BS(\sqrt{\theta})} - \bar{l}_n^{\mathcal{U}} \rvert$
1	Once a day	0.1769	0.6139	0.7893	0.4369	0.1753
1	Twice a day	0.1769	0.5521	0.7206	0.3752	0.1685
0.75	Once a day	0.7781	0.9389	1.0985	0.1608	0.1596
0.75	Twice a day	0.7781	0.8774	1.0376	0.0993	0.1602
0.5	Once a day	2.2267	2.3112	2.4252	0.0844	0.1141
0.5	Twice a day	2.2267	2.2454	2.3701	0.0187	0.1246
0.25	Once a day	4.7591	4.7928	4.8965	0.0338	0.1036
0.25	Twice a day	4.7591	4.7614	4.8793	0.0023	0.1176

Table 3: Comparison of the lower bound, the upper bound and the benchmark ($l_{BS(\sqrt{\theta})}$ with $\sqrt{\theta} = 0.2$) for quadratic Loss, ATM Call for different hedging frenquencies with $H = 0.1$, $n = 20$, $V_0 = Coef \times BS_{Call}(\sigma = \sqrt{\theta})$ and $T = 0.25$

Model	$\xi_n^{\mathcal{U}}(0)$	$Vega\ term$
BS	0.5480	/
Heston	0.4917	-0.1103
Approx. rough Heston ($H = 0.3, n = 20$)	0.4704	-0.1571
Approx. rough Heston ($H = 0.1, n = 20$)	0.4457	-0.2093

Table 4: Impact of $Vega\ term$ on the hedging ratio with $Vega\ term := \frac{\rho\sigma}{S_0} \sum_{i=1}^{n} \partial_{\nu_i} E^{\mathbb{Q}(c^*)} \left(e^{-rT}(S_T - K)_+ \right)$ for ATM Call, $n = 20$, $V_0 = 0.8 \times BS_{call}(\sqrt{\theta})$ and $T = 0.25$

Finally, it is also interesting to study the impact of the correlation ρ on the hedging strategy. Figure 7 in Appendix presents the impact of the correlation $\rho \leq 0$ on the bounds. We observe that the loss decreases as ρ decreases because the closer the absolute value of the correlation is to 1, the more the volatility risk can be hedged by taking a hedging strategy on the underlying. We also observe that the relative difference between the upper bound and the benchmark loss widens as the correlation decreases. This can be explained by the fact that the closer the absolute value correlation is to 1, the larger is the difference between the approximate hedging ratio $\bar{\xi}_n^{\mathcal{U}}(t)$ and the Black-Scholes hedging ratio ξ_t^{BS}.

7 Conclusion

This paper discusses partial hedging strategies in rough volatility models. We formulate the problem as a stochastic control problem but, due to the non-Markovian nature of the rough volatility models, this problem is considerably difficult to solve.

Thanks to a Markov multifactor approximation of the volatility process, we introduce a Markov stochastic control problem. We show, using convergence results, that, instead of solving the original problem, we can solve the Markov problem with a small error. The optimal solution to this problem is characterized by a Hamilton-Jacobi-Bellman (HJB) equation. However, even by switching to the dual formulation of the problem, we need to solve a nonlinear PDE to obtain the optimal solution. Therefore, in general, we cannot derive an explicit form of the optimal solution.

In order to obtain explicit hedging strategies, we introduce a dual control method. We derive lower and upper bounds as well as sub-optimal hedging ratios for sub-optimal choices of dual control. Moreover, if the subset of admissible dual controls is large enough, we show that the discrepancies between bounds and the optimal solution are quite small. For a particular subset, explicit formulas for lower bound and sub-optimal hedging ratio are deduced in rough Heston model with power loss function. Furthermore, in rough Heston model, the sub-optimal hedging ratio exhibits a meaningful interpretation in term of Greeks and can be efficiently computed using a FFT method for hedging of vanilla options.

Numerical results show satisfying results especially for linear payoffs hedging since errors between bounds and optimal solution are of order $\mathcal{O}(10^{-3})$. For vanilla option hedging, the discrepancy between the bounds is slightly larger, yet remains acceptable. This can be explained by the fact that the nonlinear payoff hedging problem is noisier and that the sub-optimal choices of the dual control have more influence on the hedging strategies than in the linear case.

In terms of future research, several promising avenues can be explored within the context of hedging in rough volatility models. One potential direction involves investigating a backward stochastic differential equation (BSDE) approach, building upon previous work [3], to obtain the optimal solution of the Markov problem. Furthermore, considering a deep learning approach for solving the nonlinear partial differential equation arising from the HJB equation, as explored by [34], could provide valuable insights. A comparative study between the solutions obtained via these alternative methods and those derived from the dual control method discussed in this paper would be of great interest.

8 Appendix

Figures

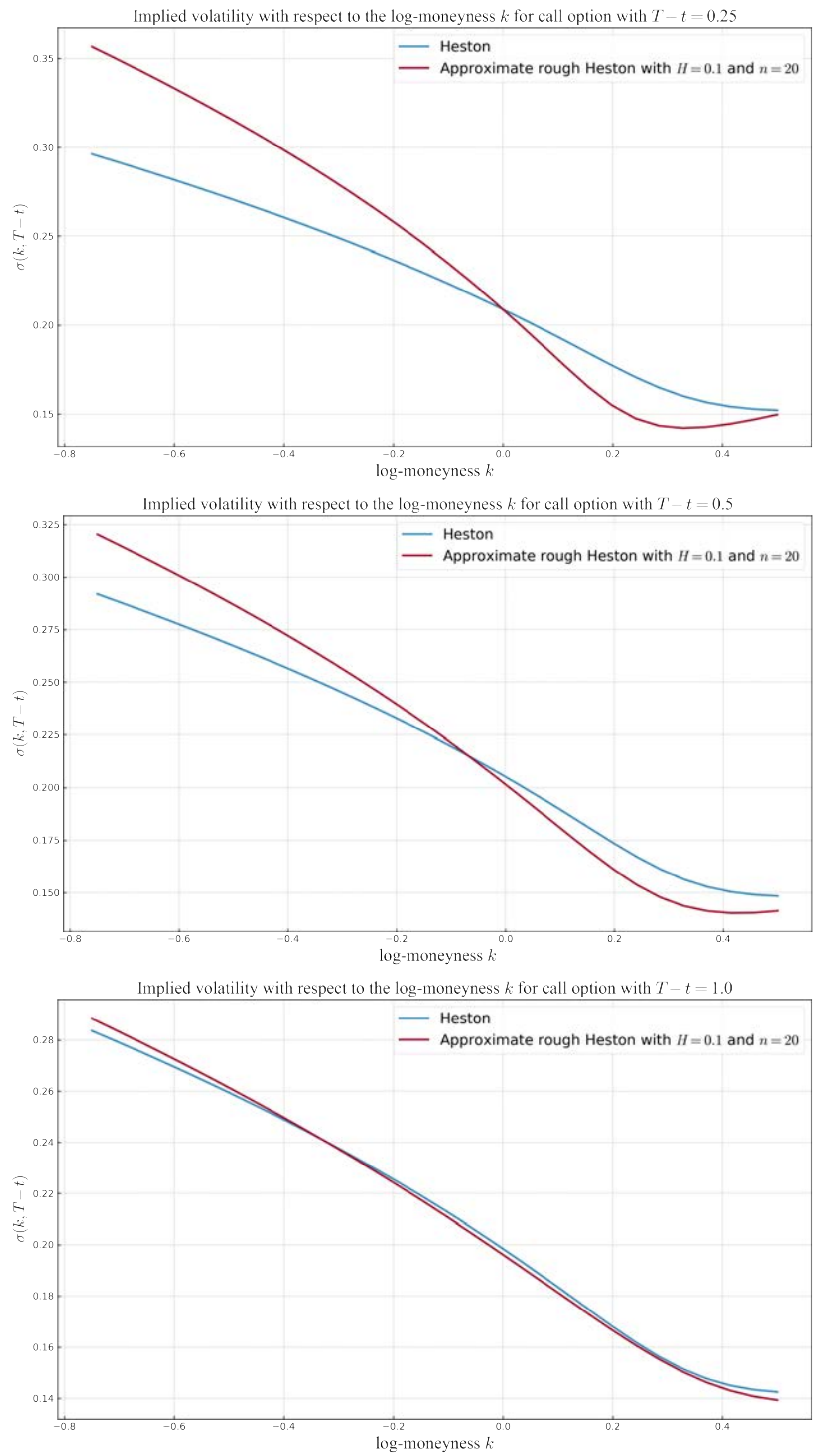

Figure 5: Vanilla implied volatility as a function of the the log-moneyness for different time-to-maturity

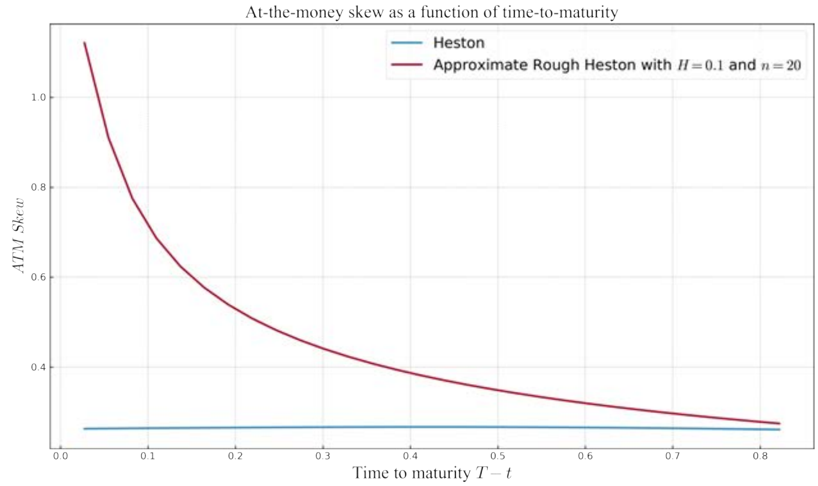

Figure 6: At-the-money skew asa function of time-to-maturity

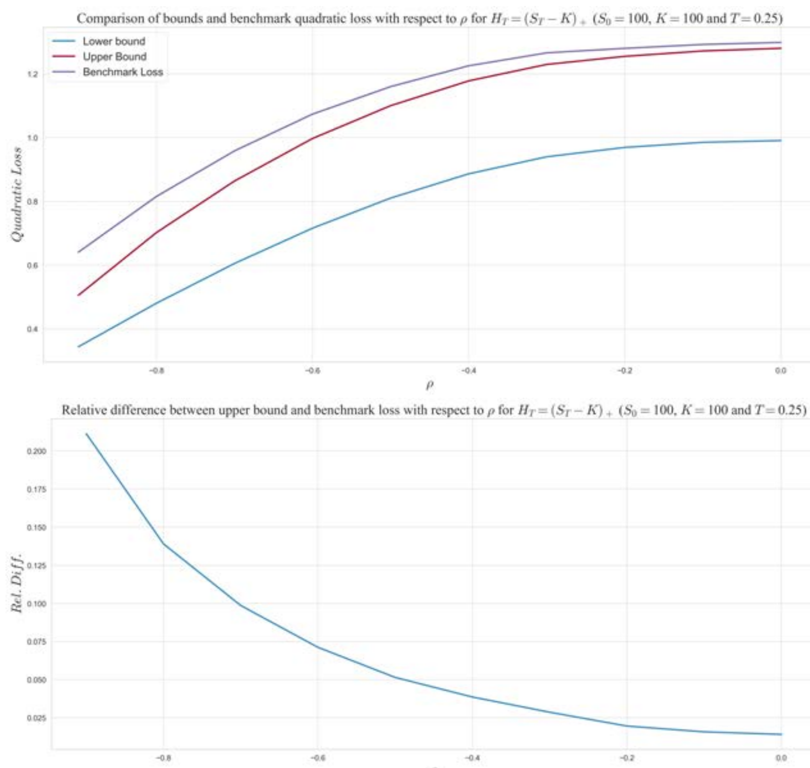

Figure 7: Impact of the correlation ρ for ATM Call with $n = 20$, $V_0 = 0.8 \times BS_{call}(\sqrt{\theta})$ and $T = 0.25$

Additional proofs.

Proof of Proposition 5.

Proof. First, based on their SDE and provide that $(S_t^n)_{0\leq t\leq T}$ and $(S_t)_{0\leq t\leq T}$ have the same initial value S_0 then it can be shown that the processes are solution of

$$\begin{aligned} S_t &= S_0 \times \exp\Big(\int_0^t (\mu_s - \frac{1}{2}\nu_s)ds + \int_0^t \sqrt{\nu_s}dW_v(s)\Big) \\ S_t^n &= S_0 \times \exp\Big(\int_0^t (\hat{\mu}_s - \frac{1}{2}\hat{\nu}_s)ds + \int_0^t \sqrt{\hat{\nu}_s}dW_v(s)\Big). \end{aligned}$$

By Theorem (4), monotone convergence implies that, for $t \in [0,T]$,

$$\lim_{n\to\infty} E\Big(\int_0^t (\sqrt{\hat{\nu}_s} - \sqrt{\nu_s})^2 ds\Big) = 0.$$

Therefore, by the definition of the stochastic Itô integral, we obtain that

$$\int_0^t \sqrt{\hat{\nu}_s}dW_v(s) \xrightarrow{L_2} \int_0^t \sqrt{\nu_s}dW_v(s)$$

But since L^2 convergence implies convergence in law and, by the Skorokhod representation theorem, that convergence in law implies almost sure convergence on a suitable probability space, we obtain that

$$\int_0^t \sqrt{\hat{\nu}_s}dW_v(s) \xrightarrow{a.s.} \int_0^t \sqrt{\nu_s}dW_v(s).$$

Moroever, using Theorem (4), we deduce that

$$\int_0^t (\hat{\mu}_s - \frac{1}{2}\hat{\nu}_s)ds \xrightarrow{a.s.} \int_0^t (\mu_s - \frac{1}{2}\nu_s)ds.$$

Finally as the exponential function is continuous, we conclude that

$$\exp\Big(\int_0^t (\hat{\mu}_s - \frac{1}{2}\hat{\nu}_s)ds + \int_0^t \sqrt{\hat{\nu}_s}dW_v(s)\Big) \xrightarrow{a.s.} \exp\Big(\int_0^t (\mu_s - \frac{1}{2}\nu_s)ds + \int_0^t \sqrt{\nu_s}dW_v(s)\Big)$$

and then

$$S_t^n \xrightarrow{a.s.} S_t,$$

as n goes to infinity.

The proof of the convergence of $(V_t^n)_{0\leq t\leq T}$ is similar. For sake of simplicity, we only consider the case where $r = 0$. In this case, as we assume $\forall t \in [0,T]$, $\xi_t^n = \xi_t$, we have that

$$\begin{aligned} V_t &= V_0 + \int_0^t \xi_s\mu_s S_s ds + \int_0^t \xi_s\sqrt{\nu_s}S_s dW_s(s) \\ V_t^n &= V_0 + \int_0^t \xi_s\hat{\mu}_s S_s^n ds + \int_0^t \xi_s\sqrt{\hat{\nu}_s}S_s^n dW_s(s) \end{aligned}$$

By Theorem (4) and Proposition (5), monotone convergence implies that, for $t \in [0,T]$,

$$\lim_{n\to\infty} E\Big(\int_0^t \xi_s^2(\sqrt{\hat{\nu}_s}S_s^n - \sqrt{\nu_s}S_s)^2 ds\Big) = 0.$$

Therefore, by the definition of the stochastic Itô integral, we obtain that

$$\int_0^t \xi_s\sqrt{\hat{\nu}_s}S_s^n dW_v(s) \xrightarrow{L_2} \int_0^t \xi_s\sqrt{\nu_s}S_s dW_v(s).$$

But since L^2 convergence implies convergence in law and, by the Skorokhod representation theorem, that convergence in law implies almost sure convergence on a suitable probability space, we obtain that

$$\int_0^t \xi_s \sqrt{\hat{\nu}_s} S_s^n dW_v(s) \xrightarrow{a.s.} \int_0^t \xi_s \sqrt{\nu_s} S_s dW_v(s).$$

Finally, as

$$\int_0^t \xi_s \hat{\mu}_s S_s^n ds \xrightarrow{a.s.} \int_0^t \xi_s \mu_s S_s ds,$$

we conclude that

$$V_t^n \xrightarrow{a.s.} V_t,$$

as n goes to infinity. $\square$

Proof of Proposition 9

Proof. To deduce the PDE of $\hat{l}_n(t,s,\nu,z)$ using the PDE of $l_n(t,s,\nu,v)$, we use the following relations between the primal and dual solution. We can remark that

$$\begin{aligned}
\partial_v l_n &= z,\ \partial_{vv} l_n = -\frac{1}{\partial_{zz}\hat{l}_n} \\
\partial_t l_n &= \partial_t \hat{l}_n,\ \partial_{sv} l_n = -\frac{\partial_{sz}\hat{l}_n}{\partial_{zz}\hat{l}_n} \\
\partial_s l_n &= \partial_s \hat{l}_n,\ \partial_{\nu_i v} l_n = -\frac{\partial_{\nu_i z}\hat{l}_n}{\partial_{zz}\hat{l}_n} \\
\partial_{\nu_i} l_n &= \partial_{\nu_i}\hat{l}_n,
\end{aligned}$$

and

$$\begin{aligned}
\partial_{ss} l_n &= \frac{\partial_{zz}\hat{l}_n \partial_{ss}\hat{l}_n - (\partial_{sz}\hat{l}_n)^2}{\partial_{zz}\hat{l}_n} \\
\partial_{\nu_i\nu_i} l_n &= \frac{\partial_{zz}\hat{l}_n \partial_{\nu_i\nu_i}\hat{l}_n - (\partial_{\nu_i z}\hat{l}_n)(\partial_{\nu_i z}\hat{l}_n)}{\partial_{zz}\hat{l}_n} \\
\partial_{\nu_i s} l_n &= \frac{\partial_{zz}\hat{l}_n \partial_{\nu_i s}\hat{l}_n - (\partial_{sz}\hat{l}_n)(\partial_{\nu_i z}\hat{l}_n)}{\partial_{zz}\hat{l}_n}.
\end{aligned}$$

Using now the PDE (19) satisfied by $l_n(.)$, we can easily derive that the PDE satisfied by $\hat{l}_n(.)$ is given by

$$\begin{aligned}
0 =& \partial_t \hat{l}_n + \mathcal{L}_{s,\boldsymbol{\nu}} \hat{l}_n - zr\partial_z \hat{l}_n + \frac{1}{2\hat{\nu}s^2} z^2\,(\hat{\mu}-r)^2 s^2 \partial_{zz}\hat{l}_n - z\,(\hat{\mu}-r)s\partial_{sz}\hat{l}_n \\
& - \frac{1}{\sqrt{\hat{\nu}}}\rho \sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n\, z\,(\hat{\mu}-r)\sigma(\hat{\nu}) - \frac{1}{2\partial_{zz}\hat{l}_n}\,\sigma(\hat{\nu})^2(1-\rho^2)\,\sum_{i=1}^n\sum_{j=1}^n \partial_{\nu_i z}\hat{l}_n \partial_{\nu_j z}\hat{l}_n.
\end{aligned}$$

Thus we just have shown that $\hat{l}_n$ satisfies the announed PDE. Concerning the optimal primal control $\xi^{n,*}$, we have that

$$\begin{aligned}
\xi_t^{n*} =& -\frac{\partial_v l_n\,(\hat{\mu}-r)S_t^n + \partial_{sv} l_n\,\hat{\nu}_t (S_t^n)^2 + \rho\sum_{i=1}^n \partial_{\nu_i v} l_n\,\sqrt{\hat{\nu}_t} S_t^n\,\sigma(\hat{\nu}_t)}{\partial_{vv} l_n\,\hat{\nu}_t (S_t^n)^2} \\
=& \frac{Z_t\,\partial_{zz}\hat{l}_n\,(\hat{\mu}-r)S_t^n - \partial_{sz}\hat{l}_n\,\hat{\nu}_t (S_t^n)^2 - \rho\sum_{i=1}^n \partial_{\nu_i z}\hat{l}_n\,\sqrt{\hat{\nu}_t} S_t^n\,\sigma(\hat{\nu}_t)}{\hat{\nu}_t (S_t^n)^2},
\end{aligned}$$

that concludes the proof. $\square$

Proof of Lemma 12.

Proof. The proof is simple and involves the first order condition since $L(h,v)$ is convex with respect to the variable v. By definition, the dual terminal value is such that

$$\hat{L}(h(s),z) := \inf_v \{L(h(s),v) - zv\},$$

in our case, it reduces to

$$\hat{L}(h(s),z) = \inf_v \{\frac{1}{p}(h(s)-v)^p - zv\}. \tag{43}$$

Using the first order condition, we have that the optimal v is such that

$$-(h(s)-v)^{p-1} - z = 0,$$

we easily deduce that

$$v = h(s) - (-z)^{\frac{1}{p-1}}.$$

Plugging this value in (43), we have that

$$\begin{aligned}
\hat{L}(h(s),z) &= \frac{1}{p}(-z)^{\frac{p}{p-1}} - z(h(s) - (-z)^{\frac{1}{p-1}})\\
&= \frac{1}{p} z^{\frac{p}{p-1}} - z^{\frac{p}{p-1}} - z\,h(s)\\
&= -\frac{p-1}{p} z^{\frac{p}{p-1}} - z\,h(s)\\
&= -\frac{1}{q} z^q - z\,h(s).
\end{aligned}$$

□

Proof of Proposition 13

Proof. From (9), we know that $\hat{l}_n$ satisfies the PDE (20). Let consider the ansatz

$$\hat{l}_n(t,s,\nu,z) = \underbrace{-\frac{z^q}{q}\exp\left(C_t + \sum_{i=1}^n D_t^i \nu_t^i\right)}_{:=\hat{l}_n^1} - \left(\underbrace{e^{-r(T-t)}\alpha + \beta s}_{:=g(t,s)}\right) z, \tag{44}$$

with

$$\begin{aligned}
C_T &= 0,\\
D_T^i &= 0,\ i = 1,...,n\\
g(t,S_t) &= H_T^{linear}
\end{aligned}$$

such that

$$\hat{l}_n(T,s,\nu,z) = \hat{L}(h(s),z) = -\frac{1}{q}z^q - h(s)^{linear} z.$$

Plugging (44) in the PDE (20), we obtain

$$\begin{aligned}
0 = \hat{l}_n^1 \Bigg(&\partial_t C_t + \sum_{i=1}^n \partial_t D_t^i \nu_i + \sum_{i=1}^n D_t^i(-x_i\nu_i + \kappa(\theta - \hat{\nu})) + \zeta^2\hat{\nu}\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n D_t^i D_t^j\\
&- rq + \frac{1}{2}q(q-1)\,A^2\hat{\nu} - q\rho A\hat{\nu}\zeta\sum_{i=1}^n D_t^i - \frac{1}{2}(1-\rho^2)\zeta^2\hat{\nu}\frac{q}{q-1}\sum_{i=1}^n\sum_{j=1}^n D_t^i D_t^j\Bigg)\\
&- z\left(\partial_t g + \partial_s g\,\hat{\mu}s + \frac{1}{2}\partial_{ss}g\,\hat{\nu}s^2 - rg - (\hat{\mu}-r)s\partial_s g\right),
\end{aligned}$$

equivalently, we have

$$
\begin{aligned}
0 =& \hat{l}_n^1 \Bigg(\partial_t C_t + \sum_{i=1}^n D_t^i \Big(\kappa(\theta - \nu_0) - q\rho A \nu_0 \zeta \Big) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n D_t^i D_t^j \zeta^2 \nu_0 \Big(1 - (1-\rho^2) \frac{q}{q-1} \Big) - rq + \frac{1}{2} q(q-1)\, A^2 \nu_0 \Bigg) \\
&+ \hat{l}_n^1 \sum_{i=1}^n \nu_i \Bigg(\partial_t D_t^i - x_i D_t^i - w_i \Big(\kappa - q\rho A \zeta \Big) \sum_{j=1}^n D_t^j + \frac{1}{2} w_i \Big(\zeta^2 - (1-\rho^2) \zeta^2 \frac{q}{q-1} \Big) \sum_{j=1}^n \sum_{k=1}^n D_t^j D_t^k \\
&+ \frac{1}{2} w_i q(q-1)\, A^2 \Bigg) \\
&- z \left(\partial_t g + \partial_s g\, \mu s + \frac{1}{2} \partial_{ss} g\, \hat{\nu} s^2 - rg - (\hat{\mu} - r) s \partial_s g \right).
\end{aligned}
$$

Using the definition of $g(t,s)$,we observe that

$$
\partial_t g + \partial_s g\, \mu s + \frac{1}{2} \partial_{ss} g\, \nu s^2 - rg - (\hat{\mu} - r) s \partial_s g = 0.
$$

In this case, the PDE reduces to

$$
\begin{aligned}
0 =& \hat{l}_n^1 \Bigg(\partial_t C_t + \sum_{i=1}^n D_t^i \Big(\kappa(\theta - \nu_0) - q\rho A \nu_0 \zeta \Big) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n D_t^i D_t^j \zeta^2 \nu_0 \Big(1 - (1-\rho^2) \frac{q}{q-1} \Big) - rq + \frac{1}{2} q(q-1)\, A^2 \nu_0 \Bigg) \\
&+ \hat{l}_n^1 \sum_{i=1}^n \nu_i \Bigg(\partial_t D_t^i - x_i D_t^i - w_i \Big(\kappa - q\rho A \zeta \Big) \sum_{j=1}^n D_t^j + \frac{1}{2} w_i \Big(\zeta^2 - (1-\rho^2) \zeta^2 \frac{q}{q-1} \Big) \sum_{j=1}^n \sum_{k=1}^n D_t^j D_t^k \\
&+ \frac{1}{2} w_i q(q-1)\, A^2 \Bigg).
\end{aligned}
$$

We obtain that C_t and $(D_t^i)_{i=1,\dots,n}$ solve Riccati type ODEs of the form

$$
\begin{aligned}
\partial_t C_t =& rq - \frac{1}{2} q(q-1)\, A^2 \nu_0 - \sum_{i=1}^n D_t^i \Big(\kappa(\theta - \nu_0) - q\rho A \nu_0 \zeta \Big) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n D_t^i D_t^j \zeta^2 \nu_0 \Big(1 - (1-\rho^2) \frac{q}{q-1} \Big), \\
C_T =& 0
\end{aligned} \tag{45}
$$

and for $i = 1, ..., n$,

$$
\begin{aligned}
\partial_t D_t^i =& x_i D_t^i + w_i \sum_{j=1}^n D_t^j \Big(\kappa - q \Big(-\rho A \zeta + (1-\rho^2) \zeta^2 c \Big) \Big) - \frac{1}{2} w_i \Big(\zeta^2 - (1-\rho^2) \zeta^2 \frac{q}{q-1} \Big) \sum_{j=1}^n \sum_{k=1}^n D_t^j D_t^k \\
&- \frac{1}{2} w_i q(q-1)\, A^2, \\
D_T^i =& 0.
\end{aligned} \tag{46}
$$

Moreover, we know that the value at time t of the process Z_t should satisfy

$$
\partial_z \hat{l}_n + V_t = 0,
$$

we easily deduce that

$$
V_t^n = g(t, S_t^n) + Z_t^{q-1} \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i),
$$

and

$$
Z_t = \exp\Big(-\frac{1}{q-1} (C_t + \sum_{i=1}^n D_t^i \nu_t^i) \Big) \times \Big(V_t - g(t, S_t^n) \Big)^{\frac{1}{q-1}}. \tag{47}
$$

Therefore, using the relation between the primal and the dual, we conclude that the primal solution is given by

$$
l_n(t, s, \boldsymbol{\nu}, v) = -\frac{Z_t^q}{q} \exp\left(C_t + \sum_{i=1}^n D_t^i \nu_t^i \right) - \left(e^{-r(T-t)} \alpha + \beta S_t^n \right) Z_t + Z_t\, V_t^n,
$$

with Z_t satisfying (47). Let now focus on the expression of the optimal hedging ratio. For that, we need to compute partial derivatives of $\hat{l}_n$. Using the closed form of $\hat{l}_n$, we deduce that

$$\partial_z \hat{l}_n = -Z_t^{q-1} \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i) - g(t, S_t^n),$$

and

$$\begin{aligned}
\partial_{zz} \hat{l}_n &= -(q-1) Z_t^{q-2} \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i) \\
\partial_{zs} \hat{l}_n &= -\beta \\
\partial_{z\nu_i} \hat{l}_n &= -Z_t^{q-1} D_t^i \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i),
\end{aligned}$$

Plugging now these values in the expression of the optimal hedging ratio leads to

$$\begin{aligned}
\xi_t^{n\,*} = &\frac{1}{\hat{\nu}_t \Big(S_t^n\Big)^2} \Big(-(q-1) Z_t^{q-1} \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i)\, (\hat{\mu}_t - r) S_t^n + \beta\, \hat{\nu}_t S_t^{n\,2} \\
&+ \rho\zeta S_t^n \hat{\nu}_t \sum_{i=1}^n \Big(Z_t^{q-1} D_t^i \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i) \Big) \Big).
\end{aligned} \tag{48}$$

Finally, using the value of Z_t (47), we conclude that the optimal hedging ratio is given by

$$\xi_t^{n\,*} = \frac{1}{\hat{\nu}_t S_t^n} \Big(e^{-r(T-t)}\alpha + \beta S_t^n - V_t^n \Big) \Big(\frac{1}{p-1} (\hat{\mu}_t - r) - \rho\zeta \hat{\nu}_t \sum_{i=1}^n D_t^i \Big) + \beta.$$

That concludes the proof as we have shown all the stated results. □

Characteristic function of the log-price in approximate rough Heston

Remember that the approximate rough Heston volatility model is given by $\hat{\nu}_t = \nu_0 + \sum_{i=1}^n w_i \nu_t^i$ where $(\nu_t^1, ..., \nu_t^n)$ *is solution of the* n *dimensionnal SDE defined by*

$$\begin{aligned}
\nu_t^i &= -\int_0^t x_i \nu_s^i ds + \int_0^t \kappa(\theta - \hat{\nu}_s) ds + \int_0^t \zeta \sqrt{\hat{\nu}_s} dW_v(s),\ i = 1, ..., n, \\
\nu_0^i &= 0,
\end{aligned} \tag{49}$$

Proposition. *The characteristic function of the log-price in approximate rough Heston is given by, for* $t \in [0, T]$,

$$\phi_t(T, x) := E_{t,s,\boldsymbol{\nu}} \Big(\exp(i\, x\, \log(S_T^n)) \Big) = \exp\Big(C_t + \sum_{i=1}^n D_t^i \nu_t^i + i\, x\, \log(S_t^n) \Big),$$

with for $0 \le t \le T$, C_t *and* $(D_t^i)_{i=1,...,n}$ *solve Riccati type ODEs of the form*

$$\begin{aligned}
\partial_t C_t &= -r\, i\, x + \nu_0 \Big(\frac{1}{2} x^2 - (A - \frac{1}{2})\, i\, x \Big) - \sum_{k=1}^n D_t^k (\rho\zeta\nu_0\, i\, x + \kappa(\theta - \nu_0)) - \frac{1}{2} \zeta^2 \nu_0 \sum_{k=1}^n \sum_{l=1}^n D_t^k D_t^l, \\
C_T &= 0
\end{aligned}$$

and for $k = 1, ..., n$,

$$\begin{aligned}
\partial_t D_t^k =& x_k D_t^k + \omega_k \Big(\frac{1}{2} x^2 - (A - \frac{1}{2})\, i\, x \Big) - w_k \sum_{j=1}^n D_t^j \Big(\rho\zeta\, i\, x - \kappa \Big) - \frac{1}{2} \zeta^2 w_k \sum_{j=1}^n \sum_{l=1}^n D_t^j D_t^l \\
D_T^k =& 0.
\end{aligned}$$

Proof. The proof is a direct application of Feynman-Kac theorem to $E_{t,s,\boldsymbol{\nu}} \Big(\exp(ix\, \log(S_T^n)) \Big)$ with SDEs (9) and (49). □

Partial Hedging under Black and Scholes

Proposition. *Suppose that the risky asset* $(S_t)_{0\leq t\leq T}$ *has the following Black-Scholes dynamic under the real measure* $\mathbb{P}$

$$dS_t = \mu S_t dt + \sigma S_t dW_S(t),$$

the optimal hedging ratio associated to a power loss $(\xi_t^{BS})_{0\leq t\leq T}$ *is given by*

$$\xi_t^{BS} := \frac{1}{p-1}\left(E^{\mathbb{Q}}(H_T|\mathcal{F}_t) - V_t\right)\frac{(\mu - r)}{\sigma S_t} + \partial_S E^{\mathbb{Q}}(H_T|\mathcal{F}_t), \tag{50}$$

with the "classical" risk-neutral measure $\mathbb{Q}$ *such that under this measure*

$$dS_t = rS_t dt + \sigma S_t dW_S^{\mathbb{Q}}(t).$$

Proof. It is sufficient to solve the dual problem and use the closed form of the dual solution to get the form the optimal hedge ratio (50). □

References

[1] Abi Jaber, E., 2019. Lifting the Heston model. Quantitative Finance, 19(12), pp.1995-2013.

[2] Abi Jaber, E. and El Euch, O., 2019. Multifactor approximation of rough volatility models. SIAM Journal on Financial Mathematics, 10(2), pp.309-349.

[3] Abi Jaber, E., Miller, E. and Pham, H., 2021. Markowitz portfolio selection for multivariate affine and quadratic Volterra models. SIAM Journal on Financial Mathematics, 12(1), pp.369-409.

[4] Aichinger, F. and Desmettre, S., 2023. Utility maximization in multivariate Volterra models. SIAM Journal on Financial Mathematics, 14(1), pp.52-98.

[5] Alfonsi, A. and Kebaier, A., 2021. Approximation of Stochastic Volterra Equations with kernels of completely monotone type. arXiv preprint arXiv:2102.13505.

[6] Bäuerle, N. and Desmettre, S., 2020. Portfolio optimization in fractional and rough Heston models. SIAM Journal on Financial Mathematics, 11(1), pp.240-273.

[7] Bayer, C., Friz, P. and Gatheral, J., 2016. Pricing under rough volatility. Quantitative Finance, 16(6), pp.887-904.

[8] Bayer, C. and Breneis, S., 2023. Markovian approximations of stochastic Volterra equations with the fractional kernel. Quantitative Finance, 23(1), pp.53-70.

[9] Biagini, F., Guasoni, P. and Pratelli, M., 2000. Mean-variance hedging for stochastic volatility models. Mathematical Finance, 10(2), pp.109-123.

[10] Cvitanic, J., 2000. Minimizing expected loss of hedging in incomplete and constrained markets. SIAM Journal on Control and Optimization, 38(4), pp.1050-1066.

[11] Dupret, J.L. and Hainaut, D., 2021. Portfolio insurance under rough volatility and Volterra processes. International Journal of Theoretical and Applied Finance, 24(06n07), p.2150036.

[12] Dupret, J.L., Barbarin, J. and Hainaut, D., 2023. Impact of rough stochastic volatility models on long-term life insurance pricing. European Actuarial Journal, 13(1), pp.235-275.

[13] El Euch, O. and Rosenbaum, M., 2018. Perfect hedging in rough Heston models. The Annals of Applied Probability, 28(6), pp.3813-3856.

[14] El Euch, O. and Rosenbaum, M., 2019. The characteristic function of rough Heston models. Mathematical Finance, 29(1), pp.3-38.

[15] Föllmer, H. and Leukert, P., 1999. Quantile hedging. Finance and Stochastics, 3, pp.251-273.

[16] Föllmer, H. and Leukert, P., 2000. Efficient hedging: cost versus shortfall risk. Finance and Stochastics, 4, pp.117-146.

[17] Fouque, J.P., Sircar, R. and Zariphopoulou, T., 2017. Portfolio optimization and stochastic volatility asymptotics. Mathematical Finance, 27(3), pp.704-745.

[18] Fouque, J.P. and Hu, R., 2018. Portfolio optimization under fast mean-reverting and rough fractional stochastic environment. Applied Mathematical Finance, 25(4), pp.361-388.

[19] Fouque, J.P. and Hu, R., 2018. Optimal portfolio under fast mean-reverting fractional stochastic environment. SIAM Journal on Financial Mathematics, 9(2), pp.564-601.

[20] Fukasawa, M., 2017. Short-time at-the-money skew and rough fractional volatility. Quantitative Finance, 17(2), pp.189-198.

[21] Fukasawa, M., Horvath, B. and Tankov, P., 2021. Hedging under rough volatility. arXiv preprint arXiv:2105.04073.

[22] Gatheral, J., Jaisson, T. and Rosenbaum, M., 2018. Volatility is rough. Quantitative finance, 18(6), pp.933-949.

[23] Garnier, J. and Sølna, K., 2020. Optimal hedging under fast-varying stochastic volatility. SIAM Journal on Financial Mathematics, 11(1), pp.274-325.

[24] Han, B. and Wong, H.Y., 2021. Mean–variance portfolio selection under Volterra Heston model. Applied Mathematics & Optimization, 84, pp.683-710.

[25] Hainaut, D., 2022. Continuous time processes for finance. Switching, self-exciting, fractional and other recent dynamics. Bocconi & Springer Series.

[26] Hainaut, D., 2022. Multivariate claim processes with rough intensities: Properties and estimation. Insurance: Mathematics and Economics, 107, pp.269-287.

[27] Hainaut, D., 2023. Pricing of spread and exchange options in a rough jump–diffusion market. Journal of Computational and Applied Mathematics, 419, p.114752.

[28] Horvath, B., Jacquier, A. and Tankov, P., 2020. Volatility options in rough volatility models. SIAM Journal on Financial Mathematics, 11(2), pp.437-469.

[29] Jonsson, M. and Sircar, K.R., 2002. Partial hedging in a stochastic volatility environment. Mathematical Finance, 12(4), pp.375-409.

[30] Kamizono, K., 2003. Partial hedging under transaction costs. SIAM journal on control and optimization, 42(5), pp.1545-1558.

[31] Ma, J., Li, W. and Zheng, H., 2020. Dual control Monte-Carlo method for tight bounds of value function under Heston stochastic volatility model. European Journal of Operational Research, 280(2), pp.428-440.

[32] Ma, J., Lu, Z. and Chen, D., 2023. Optimal reinsurance-investment with loss aversion under rough Heston model. Quantitative Finance, 23(1), pp.95-109.

[33] Mania, M., Tevzadze, R. and Toronjadze, T., 2008. Mean-variance hedging under partial information. SIAM Journal on Control and Optimization, 47(5), pp.2381-2409.

[34] Pham, H., Warin, X. and Germain, M., 2021. Neural networks-based backward scheme for fully nonlinear PDEs. SN Partial Differential Equations and Applications, 2(1), p.16.

[35] Touzi, N. and Tourin, A., 2013. Optimal stochastic control, stochastic target problems, and backward SDE (Vol. 29). New York: Springer.