PARTIAL HEDGING IN ROUGH VOLATILITY MODELS

Edouard Motte, Donatien Hainaut

REPRINT | 2024 / 39

UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

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

Edouard Motte*, Donatien Hainaut†

March 22, 2024

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-martingales, stochastic control problems associated with 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 by 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. [26]. These models incorporate short-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 [10, 18, 24, 26, 33]. Moreover, [1, 6] show that Markovian approximation of rough volatility models can also effectively capture, with few parameters, implied volatility smile as well as at-the-money skew and recent extensive empirical results [5, 28, 40] show that Markovian counterparts of rough models perform well in capturing the SPX smiles and skew. The interest in rough processes also extends to other domains such as insurance, whether in terms of their impact on pricing and insurance portfolios [15, 16] or on claims modelling [31].

In this paper, we investigate the problem of hedging in rough volatility models. While previous research [17, 25] 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 does not exist for any given contingent claim, this is why we are interested in partial hedging strategies. Partial hedging strategies introduced by Föllmer and Leukert [19, 20], 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 [12, 14, 27, 34, 35, 38]. Notably, [34] extends the theory

*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)

of partial hedging to stochastic volatility environment by formulating the problem as a stochastic control problem. However, the 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 [4, 7, 9, 22, 23, 29, 37], 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 problems, [7, 22, 23] propose first order approximate solutions by relying on martingale distortion transformation of the value function, while [9, 37] 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, 8, 11, 31], 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 [1, 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 [21, 34, 36] 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. [21, 34] consider fast-mean reverting volatility models to propose asymptotic solutions, while [36] considers 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 [36] 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 filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ where $\mathbb{P}$ stands for the real measure, the filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ is the canonical filtration of a two-dimensional Brownian motion (W_S, B_ν) and 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

$$dS_t^0 = r\, S_t^0 dt,\ S_0^0 = 1,$$

and

$$dS_t = \mu_t S_t dt + \sqrt{\nu_t} S_t dW_S(t),\ S_0 = s_0 > 0,$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$, where $r \in \mathbb{R}$ is the risk-free interest rate, $\mu_t := r + A\nu_t,\ A \in \mathbb{R}_+$[1] and $(\nu_t)_{0\leq t\leq T}$ a rough volatility process. The rough volatility satisfies a stochastic Volterra equation of the form

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

[1] We assume that the excess of return of the risky asset is proportional to its variance.

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$, where $\nu_0 \in \mathbb{R}_+$, $b : \mathbb{R} \to \mathbb{R}$ Lipschitz continuous with $b(0) \geq 0$ and $\sigma : \mathbb{R} \to \mathbb{R}$ $\eta-$Hölder continuous with $\sigma(0) = 0$ and $\eta \in [1/2, 1]$, W_ν is a standard Brownian motion such that $W_\nu = \rho W_S + \sqrt{1-\rho^2} B_\nu$ for $\rho \in (-1, 1)$ and G a kernel assumed to be completely monotone[2]. The rough volatility model (1) is a general model that incorporates well known rough models such for example the rough Heston model introduced in [17] with volatility process given by

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

where $\kappa \in \mathbb{R}_+$ is the speed of mean reversion toward the level $\theta \in \mathbb{R}_+$ and $\zeta \in \mathbb{R}_+$ is the vol-of-vol parameter. 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], since functions $b(.)$ and $\sigma(.)$ are in particular continuous with linear growth, we can prove the existence of an unconstrained weak solution of the stochastic Volterra equation (1) when the fractional kernel satisfies (2). Moreover, the existence of a non-negative solution can be obtained by relying on Theorem B.4. in [2] .

Remark. If we consider the rough Heston model with $b(x) = \kappa(\theta - x)$ and $\sigma(x) = \zeta\sqrt{x}$, we can prove the uniqueness of a non-negative weak solution of the stochastic Volterra equation (1) when the kernel satisfies (2), see for instance [3].

In this financial market with rough volatility, we are concerned with the hedging of any contingent claim given by $\mathcal{F}_T-$measurable square-integrable random variable of the form

$$H_T = h(S_T),$$

with $h(.)$ a continuous function. El Euch and Rosenbaum [17] 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 possible for any contingent claim but we can still stay on the safe side by super-hedging the contingent claims. However, super-hedging strategies generally lead to super-hedging prices that are too high to be considered in practice. In cases 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 [19, 20]. It is this kind of strategy that we consider in the following. In this perspective, as usual for hedging problems[3], we consider a self-financing hedging portfolio denoted by $(V_t)_{0\leq t\leq T}$ 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

$$\begin{aligned} dV_t &= r\,(V_t - \xi_t S_t)dt + \xi_t\, dS_t, \\ V_0 &= v_0, \end{aligned}$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$. The hedging ratio process $(\xi_t)_{0\leq t\leq T}$ is admissible if $(\xi_t)_{0\leq t\leq T}$ is progressively measurable 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 $(\xi_t)_{0\leq t\leq T}$ 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, in our framework, perfect hedging is not possible for any given

[2]A Kernel $G(.)$ is completely monotone if it is infinitely differentiable on $(0, +\infty)$ such that $(-1)^j G^{(j)}(t) \geq 0$ for $t > 0$ and $j \geq 0$.

[3]For more details on how hedging problems can be tackled in complete and incomplete markets, we refer the reader to the book [13].

contingent claim, there exist payoffs[4] H_T for which

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

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

$$l(s_0, \nu_0, v_0) := \inf_{\xi_t \in \mathcal{R}} E_{s_0,\nu_0,v_0}\Big(L(h(S_T), V_T)\Big) \tag{3}$$

where $L(.)$ is a continuous proper convex loss function. Note that different choices can be made for the loss function $L(.)$. We can consider symmetric loss functions, like for example, power loss functions of the form

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

with the particular case $p = 2$ linked to mean-variance minimization problem. For other types of risk-averse traders, we can consider asymmetric functions, such as, for example, exponential or shortfall loss functions. 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 [9], 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 [9] 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 [1, 2, 8, 11, 31, 32], the starting point of the Markovian approximation is the representation of the kernel $G(t)$ in terms of a Laplace transform[5] such that

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

where λ is a measure on $\mathbb{R}_+$. For the 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 $x > 0$,

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

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

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

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

[4]Perfect hedging strategies exist for linear-form payoffs.

[5]This representation is possible since $G(.)$ is assumed to be completely monotone.

volatility process by defining a new stochastic process denoted by $(\hat{\nu}_t)_{0\leq t\leq T}$ as the unique strong continuous solution to the stochastic Volterra equation (1) with a kernel $\hat{G}(.)$ such that, for $t \in [0,T]$,

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

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ where $\nu_0 \in \mathbb{R}_+$, $b : \mathbb{R} \to \mathbb{R}$ Lipschitz continuous with $b(0) \geq 0$ and $\sigma : \mathbb{R} \to \mathbb{R}$ $\eta-$Hölder continuous with $\sigma(0) = 0$ and $\eta \in [1/2, 1]$. As for the rough volatility process, the existence of a non-negative weak solution of (6) can be obtained by relying on Theorem B.4 in [2]. Moreover, as, unlike the fractional kernel, the approximate kernel is smooth, the strong existence and uniqueness of $(\hat{\nu}_t)_{0\leq t\leq T}$ follow from Proposition B.3. in [2]. The following proposition states that the stochastic Volterra equation (6) can be reduced to a $n-$dimensional stochastic differential equation.

Proposition 1. *The solution of (6) is given by*

$$\hat{\nu}_t = \nu_0 + \sum_{i=1}^n w_i \nu_t^i \tag{7}$$

where $(\boldsymbol{\nu_t})_{0\leq t\leq T} := \Big((\nu_t^1, \nu_t^2, ..., \nu_t^n)\Big)_{0\leq t\leq T}$ *is solution of the n-dimensional SDE defined by*

$$\begin{aligned} \nu_t^i &= -\int_0^t x_i \nu_s^i ds + \int_0^t b(\hat{\nu}_s)ds + \int_0^t \sigma(\hat{\nu}_s)dW_\nu(s),\ i = 1, ..., n, \\ \nu_0^i &= 0, \end{aligned} \tag{8}$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$.

Proof. We refer the reader to the proof of Proposition 2.1. in [8]. □

Based on the approximate volatility process (6) and its Markov representation induced by SDEs (8), we can define the Markovian approximation of the stochastic control problem introduced in (3). First, we consider the approximate process $(S_t^n)_{0\leq t\leq T}$ 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

$$dS_t^n = \hat{\mu}_t S_t^n dt + \sqrt{\hat{\nu}_t} S_t^n dW_S(t), \tag{9}$$

but also by

$$dS_t^n = \hat{\mu}_t S_t^n dt + \sqrt{\nu_0 + \sum_{i=1}^n w_i \nu_t^i S_t^n dW_S(t)}, \tag{10}$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$, with $S_0^n = s_0$ and $\hat{\mu}_t = r + A\hat{\nu}_t = r + A\Big(\nu_0 + \sum_{i=1}^n w_i \nu_t^i\Big)$. 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

$$\begin{aligned} dV_t^n &= r(V_t^n - \xi_t^n S_t^n)dt + \xi_t^n dS_t^n, \\ V_0^n &= v_0, \end{aligned} \tag{11}$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$, 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

$$\pi_T^n := V_T^n - H_T^n, \tag{12}$$

with $H_T^n := h(S_T^n)$, a $\mathcal{F}_T-$measurable square-integrable random variable. As the market is incomplete, there exist payoffs for which

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

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 (9) 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

$$l_{n,\hat{\nu}}(s_0,\nu_0,v_0) := \inf_{\xi_t^n\in\mathcal{R}_n} E_{s_0,\nu_0,v_0}\bigg(L(h(S_T^n),V_T^n)\bigg). \tag{13}$$

Considering now the dynamic (10) 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

$$l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0) := \inf_{\xi_t^n\in\mathcal{R}_n} E_{s_0,\boldsymbol{\nu_0},v_0}\bigg(L(h(S_T^n),V_T^n)\bigg) \tag{14}$$

such that, for $\boldsymbol{\nu_0} = \boldsymbol{0^n}$,

$$l_{n,\hat{\nu}}(s_0,\nu_0,v_0) = l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0).$$

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 (14) 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 $l_{n,\boldsymbol{\nu}}(.)$ 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 $l_{n,\boldsymbol{\nu}}(.)$ and $l(.)$. The first step is to prove, the weak convergence of the value at time $t\in[0,T]$ of the approximate volatility process $\hat{\nu}_t$ to ν_t. To prove it, we rely on [1] and [2]. First, under specific assumptions on the weights $(w_i)_{i=1,\dots,n}$ and mean reversion terms $(x_i)_{i=1,\dots,n}$, we can prove that $\hat{G}$ converges to G in $L^2_{[0,T]}$.

Assumption 2. Fix $r_n > 1$ and suppose that the weights $(w_i)_{i=1,\dots,n}$ and mean reversion terms $(x_i)_{i=1,\dots,n}$ are given by

$$w_i := \frac{(r_n^{1-\alpha}-1)r_n^{(\alpha-1)(1+n/2)}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_n^{(1-\alpha)\,i},\; x_i := \frac{1-\alpha}{2-\alpha}\frac{r_n^{2-\alpha}-1}{r_n^{1-\alpha}-1} r_n^{i-1-n/2},\; i=1,\dots,n,$$

with $\alpha := H+1/2$ for $H\in(0,1/2)$ and $(r_n)_{n\geq 1}$ satisfying

$$r_n \downarrow 1,\; n\ln(r_n)\to\infty,$$

as n goes to infinity.

Remark. As stated in [1], Assumption 2 is satisfied if we consider $(r_n)_{n\geq1}$ of the form

$$r_n = 1+10n^{-0.9},\; n\geq 1.$$

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

Proposition 3. *(Lemma A.3 in [1]) Suppose that for all $n\geq 1$, $(w_i)_{i=1,\dots,n}$ and $(x_i)_{i=1,\dots,n}$ 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.*

$$||\hat{G}-G||_{L^2_{[0,T]}} \to 0,$$

as n goes to infinity.

Proof. We refer to [1] for the proof. □

Lemma 4. *Assume that $n\geq 1$, $(w_i)_{i=1,\dots,n}$ and $(x_i)_{i=1,\dots,n}$ satisfy Assumption 2. Let $\hat{G}(.)$ be the approximate fractional kernel defined by (5), then there exist positive constants δ and C such that*

$$\sup_{n\geq1}\bigg(\int_0^h|\hat{G}(s)|^2ds + \int_0^{T-h}|\hat{G}(h+s)-\hat{G}(s)|^2ds\bigg) \leq Ch^{2\delta},$$

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

Proof. As $(w_i)_{i=1,...,n}$ and $(x_i)_{i=1,...,n}$ satisfy Assumption 2, we can rewrite thoses parameters such that for $i = 1, ..., n$,

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

with $\eta_i^n = r_n^{i-n/2}$. As $\eta_0^n \neq 0$, the Assumption 3.1. in [2] is not satisfied and therefore the result of Lemma 5.2. in [2] cannot be directly applied. However, as mentioned in [1], the result of this lemma is deduced by adapting the proof of Lemma 5.2. in [2] using the same small adjustments highlighted in the proof of Lemma A.3. in [1]. □

Based on Proposition 3 and Lemma 4, we now establish a weak convergence of the value at time $t \in [0, T]$ of the approximate volatility process $\hat{\nu}_t$ toward ν_t.

Remark. In the subsequent results, the notation $\xrightarrow{\mathcal{L}}$ refers to weak convergence.

Theorem 5. *Assume that $n \geq 1$, $(w_i)_{i=1,...,n}$ and $(x_i)_{i=1,...,n}$ satisfy Assumption 2. Let $\left((\hat{\nu}_t)_{0\leq t\leq T}\right)_{n\geq 1}$ be a sequence of unique strong solutions to (6) and suppose that the stochastic Volterra equation (1) with fractional kernel admits a unique weak solution, then, $\forall t \in [0, T]$,*

$$\hat{\nu}_t = \nu_0 + \sum_{i=1}^{n} w_i \nu_t^i \xrightarrow{\mathcal{L}} \nu_t$$

as n goes to infinity.

Proof. The proof is similar to the proof of Theorem 3.5 in [2] and is an immediate consequence of Theorem 3.6. in [2] for one dimension. We need to check that the assumptions of the theorem are satisfied. By Lemma 4, we know that there exist positive constants δ and C such that

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

for any $t, h \geq 0$ with $t + h \leq T$. Thus, 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 fulfilled, 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 assumptions 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 limit ν is a solution of the Volterra equation (1). Thus, as we assume that the stochastic Volterra equation (1) with fractional kernel admits a unique weak solution, we deduce that

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

as n goes to infinity. □

Now that we establish a weak convergence for the value at time $t \in [0, T]$ of the volatility process, we can go a step further and show that that value of the approximate processes S_t^n and V_t^n converge weakly respectively toward S_t and V_t.

Proposition 6. *Assume that $n \geq 1$, $(w_i)_{i=1,\dots,n}$ and $(x_i)_{i=1,\dots,n}$ satisfy Assumption 2. Let $\Big((\hat{\nu}_t)_{0\leq t\leq T}\Big)_{n\geq 1}$ a sequence of unique strong solutions to (6) and suppose that the stochastic Volterra equation (1) with fractional kernel admits a unique weak solution. Consider the approximate processes $(S_t^n)_{0\leq t\leq T}$ and $(V_t^n)_{0\leq t\leq T}$ satisfying SDEs (9) and (11) such that $\forall t \in [0,T]$, $\xi_t^n = \xi_t$ a.s., then we have the following convergence results*

$$S_t^n \xrightarrow{\mathcal{L}} S_t,$$
$$V_t^n \xrightarrow{\mathcal{L}} V_t,$$

as n goes to infinity.

Proof. The proof is provided in Appendix A. □

With Proposition 6, we have shown that the value at time $t \in [0,T]$ of the approximate processes converge in law to the value at time $t \in [0,T]$ of 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 [9], we define $l^{\xi}(s_0,\nu_0,v_0)$ and $l_n^{\xi^n}(s_0,\nu_0,v_0)$ by

$$l^{\xi}(s_0,\nu_0,v_0) := E_{s_0,\nu_0,v_0}\Big(L(h(S_T),V_T(\xi))\Big)$$
$$l_n^{\xi^n}(s_0,\nu_0,v_0) := E_{s_0,\nu_0,v_0}\Big(L(h(S_T^n),V_T^n(\xi^n))\Big)$$

such that

$$l(s_0,\nu_0,v_0) = \inf_{\xi_t\in\mathcal{R}} l^{\xi}(s_0,\nu_0,v_0),$$
$$l_{n,\hat{\nu}}(s_0,\nu_0,v_0) = \inf_{\xi_t^n\in\mathcal{R}_n} l_n^{\xi^n}(s_0,\nu_0,v_0).$$

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

Lemma 7. *Assume that $n \geq 1$, $(w_i)_{i=1,\dots,n}$ and $(x_i)_{i=1,\dots,n}$ satisfy Assumption 2. Let $\Big((\hat{\nu}_t)_{0\leq t\leq T}\Big)_{n\geq 1}$ a sequence of unique strong solutions to (6) and suppose that the stochastic Volterra equation (1) with fractional kernel admits a unique weak solution. 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 $\Big(L(h(S_T^n),V_T^n)\Big)_{n\geq 1}$ is uniformly integrable, then*

$$l_n^{\xi}(s_0,\nu_0,v_0) \to l^{\xi}(s_0,\nu_0,v_0),$$

as n goes to infinity.

Proof. Using Proposition 6, since we assume that $\forall t \in [0,T]$, $\xi_t^n = \xi_t$ *a.s.*, we have that

$$S_t^n \xrightarrow{\mathcal{L}} S_t,$$
$$V_t^n \xrightarrow{\mathcal{L}} V_t,$$

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

$$L(h(S_T^n),V_T^n) \xrightarrow{\mathcal{L}} L(h(S_T),V_T),$$

and the uniform integrability of $\Big(L(h(S_T^n),V_T^n)\Big)_{n\geq 1}$ implies that

$$E\Big(L(h(S_T^n),V_T^n)\Big) \to E\Big(L(h(S_T),V_T)\Big),$$

as n goes to infinity. Thus we obtain that

$$l_n^{\xi}(s_0,\nu_0,v_0) \to l^{\xi}(s_0,\nu_0,v_0),$$

as n goes to infinity. □

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

Theorem 8. *Assume that $n \geq 1$, $(w_i)_{i=1,\ldots,n}$ and $(x_i)_{i=1,\ldots,n}$ satisfy Assumption 2. Let $\left((\hat{\nu}_t)_{0\leq t\leq T}\right)_{n\geq 1}$ a sequence of unique strong solutions to (6), suppose that the stochastic Volterra equation (1) with fractional kernel admits a unique weak solution and that the sequence $\left(L(h(S_T^n), V_T^n)\right)_{n\geq 1}$ is uniformly integrable. Let $(\xi_t^{n,*})_{0\leq t\leq T}$ be the optimal hedging ratio associated to the $n-$approximate stochastic control problem (14) with $\boldsymbol{\nu_0} = \boldsymbol{0^n}$. For every $\varepsilon > 0$, $\exists N \in \mathbb{N}$, such that $\forall n \geq N$,*

$$|l(s_0, \nu_0, v_0) - l^{\xi^{n,*}}(s_0, \nu_0, v_0)| < \varepsilon,$$

i.e.

$$\lim_{n\to+\infty} l^{\xi^{n,*}}(s_0, \nu_0, v_0) = l(s_0, \nu_0, v_0).$$

Moreover, for $\boldsymbol{\nu_0} = \boldsymbol{0^n}$,

$$\lim_{n\to+\infty} l_{n,\hat{\nu}}(s_0, \nu_0, v_0) = \lim_{n\to+\infty} l_{n,\boldsymbol{\nu}}(s_0, \boldsymbol{\nu_0}, v_0) = l(s_0, \nu_0, v_0).$$

Proof. Let fix $\varepsilon > 0$. Suppose that $(\xi_t^{n,*})_{0\leq t\leq T}$ is the optimal hedging ratio associated to the $n-$approximate stochastic control problem (14) with $\boldsymbol{\nu_0} = \boldsymbol{0^n}$, as $\forall n \geq 1$,

$$l_{n,\hat{\nu}}(s_0, \nu_0, v_0) = l_{n,\boldsymbol{\nu}}(s_0, \boldsymbol{\nu_0}, v_0),$$

for $\boldsymbol{\nu_0} = \boldsymbol{0^n}$, then we have that

$$l_{n,\hat{\nu}}(s_0, \nu_0, v_0) = l_n^{\xi^{n,*}}(s_0, \nu_0, v_0).$$

Using Lemma 7, we have that

$$\begin{aligned}\lim_{n\to\infty} l_{n,\hat{\nu}}(s_0, \nu_0, v_0) &= \lim_{n,m\to\infty} l_m^{\xi^{n,*}}(s_0, \nu_0, v_0)\\ &= \lim_{n\to\infty} l^{\xi^{n,*}}(s_0, \nu_0, v_0)\end{aligned}$$

or equivalently

$$\lim_{n\to\infty}\left(l_{n,\hat{\nu}}(s_0, \nu_0, v_0) - l^{\xi^{n,*}}(s_0, \nu_0, v_0)\right) = 0. \tag{15}$$

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

$$|l_{n,\hat{\nu}}(s_0, \nu_0, v_0) - l^{\xi^{n,*}}(s_0, \nu_0, v_0)| < \frac{\varepsilon}{2}. \tag{16}$$

Moreover, considering $\underline{l}(s_0, \nu_0, v_0)$ defined by

$$\underline{l}(s_0, \nu_0, v_0) := \lim_{n\to\infty} l_{n,\hat{\nu}}(s_0, \nu_0, v_0) = \lim_{n\to\infty} \inf_{\xi_t^n\in\mathcal{R}_n} l_n^{\xi^n}(s_0, \nu_0, v_0),$$

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

$$|\underline{l}(s_0, \nu_0, v_0) - l_{n,\hat{\nu}}(s_0, \nu_0, v_0)| \leq \frac{\varepsilon}{2} \tag{17}$$

Note that since

$$l(s_0, \nu_0, v_0) = \inf_{\xi_t\in\mathcal{R}} \lim_{n\to\infty} E_{s_0,\nu_0,v_0}\left(L(h(S_T^n), V_T^n)\right),$$

we have that

$$l(s_0, \nu_0, v_0) \geq \underline{l}(s_0, \nu_0, v_0).$$

By choosing $N := \max(N_1, N_2)$, we have that for $n \geq N$ inequalities (16) and (17) are satisfied. In this case, we have that

$$\begin{aligned}|l(s_0, \nu_0, v_0) - l^{\xi^{n,*}}(s_0, \nu_0, v_0)| &\leq |\underline{l}(s_0, \nu_0, v_0) - l^{\xi^{n,*}}(s_0, \nu_0, v_0)|\\ &\leq \underbrace{|\underline{l}(s_0, \nu_0, v_0) - l_{n,\hat{\nu}}(s_0, \nu_0, v_0)|}_{<\frac{\varepsilon}{2}} + \underbrace{|l_{n,\hat{\nu}}(s_0, \nu_0, v_0) - l^{\xi^{n,*}}(s_0, \nu_0, v_0)|}_{<\frac{\varepsilon}{2}}\\ &< \varepsilon.\end{aligned}$$

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

$$|l(s_0, \nu_0, v_0) - l^{\xi^{n,*}}(s_0, \nu_0, v_0)| < \varepsilon$$

or equivalently

$$\lim_{n \to +\infty} l^{\xi^{n,*}}(s_0, \nu_0, v_0) = l(s_0, \nu_0, v_0).$$

Moreover, using (15), we deduce that

$$\begin{aligned} \lim_{n \to +\infty} l_{n,\boldsymbol{\nu}}(s_0, \boldsymbol{\nu_0}, v_0) = \lim_{n \to +\infty} l_{n,\hat{\nu}}(s_0, \nu_0, v_0) &= \lim_{n \to +\infty} l^{\xi^{n,*}}(s_0, \nu_0, v_0) \\ &= l(s_0, \nu_0, v_0). \end{aligned}$$

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 $(\xi_t^{n,*})_{0 \leq t \leq T}$ associated to the $n-$approximate stochastic control problem (14) is $\varepsilon-$optimal for the original problem. On the other hand, we know that the solution of the approximate control problem $l_{n,\boldsymbol{\nu}}(.)$ 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 [34]. From [34], 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 [36], we propose a dual control method to obtain approximate solutions of the problem.

Assumption 9. *For the rest of the paper, we still assume that $b(.)$ is Lipschitz continuous with $b(0) \geq 0$ and $\sigma(.)$ is $\eta-$Hölder continuous with $\sigma(0) = 0$, $\eta \in [1/2, 1]$ but we additionally assume that these functions satisfy sufficient conditions for the approximate volatility process $(\hat{\nu}_t)_{0 \leq t \leq T} = (\nu_0 + \sum_{i=1}^n w_i \nu_t^i)_{0 \leq t \leq T}$ to remain strictly positive i.e. for $t \in [0, T]$,*

$$\hat{\nu}_t > 0,$$

almost surely. As shown in Appendix B, this assumption is satisfied in the approximate rough Heston case i.e. $b(x) = \kappa(\theta - x)$ and $\sigma(x) = \zeta\sqrt{x}$ if

$$2\kappa\theta > \zeta^2 \sum_{i=1}^n w_i, \ \nu_0 > 0,$$

such that for all $\varepsilon > 0$, $i \in \{1, ..., n\}$ and $t \in [0, T]$,

$$E\left(1_{\{t \leq \tau_\varepsilon\}} w_i \nu_t^i \hat{\nu}_t^{-(m+1)}\right) \leq E\left(1_{\{t \leq \tau_\varepsilon\}} \hat{\nu}_t^{-m}\right), \tag{18}$$

with $\tau_\varepsilon := \min\{t \geq 0 : \hat{\nu}_t \leq \varepsilon\}$ and $m := \frac{2\kappa\theta - \zeta^2 \sum_{i=1}^n w_i}{\zeta^2 \sum_{i=1}^n w_i}$.

For $n \geq 1$, the approximate partial hedging problem (14) involves solving a Markovian stochastic control problem with value function of the form[6]:

$$l_n(t, s, \boldsymbol{\nu}, v) := \inf_{\xi_t^n \in \mathcal{R}_n} E\left(L(h(S_T^n), V_T^n) | S_t^n = s, \boldsymbol{\nu_t} = \boldsymbol{\nu}, V_t^n = v\right), \tag{19}$$

[6]For $t = 0$, we have $l_n(t = 0, s_0, \boldsymbol{\nu}_0, v_0) = l_{n,\boldsymbol{\nu}}(s_0, \boldsymbol{\nu_0}, v_0)$.

for $(t, s, \boldsymbol{\nu}, v) \in [0,T] \times \underbrace{\mathbb{R}_+ \times \{\boldsymbol{\nu} \in \mathbb{R}^n : \nu_0 + \sum_{i=1}^n w_i \nu_i > 0\} \times \mathbb{R}}_{:=\mathbb{S}}$ and $\mathcal{R}_n$ 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 (19) is Markovian, we can solve it using the HJB equation. Assuming that $l_n(t, s, \boldsymbol{\nu}, v)$ is locally bounded on $[0,T) \times \mathbb{S}$ and the hamiltonian associated to the problem (19) is finite and continuous on $[0,T) \times \mathbb{S} \times \mathbb{R}^{2+n} \times \mathcal{S}_{2+n}$, classic results from dynamic programming (see Theorems 7.4. and 7.6. in [41]) imply that $l_n(t, s, \boldsymbol{\nu}, v)$ is the viscosity solution of the following HJB[7]:

$$\begin{aligned}
-\partial_t l_n - \inf_{\xi^n \in \mathbb{R}} \Big\{ & \partial_s l_n\, \hat{\mu} s + \sum_{i=1}^n \partial_{\nu_i} l_n \Big(- x_i \nu_i + b(\hat{\nu}) \Big) + \partial_v l_n \Big(rv + (\hat{\mu} - r)\, \xi^n s \Big) \\
& + \frac{1}{2} \partial_{ss} l_n\, \hat{\nu} s^2 + \frac{1}{2} \partial_{vv} l_n\, (\xi^n)^2\, \hat{\nu} s^2 + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \partial_{\nu_i \nu_j} l_n\, \sigma^2(\hat{\nu}) + \partial_{sv} l_n\, \xi^n\, \hat{\nu} s^2 \\
& + \rho \sum_{i=1}^n \partial_{\nu_i s} l_n\, \sqrt{\hat{\nu}} s\, \sigma(\hat{\nu}) + \rho \sum_{i=1}^n \partial_{\nu_i v} l_n\, \xi^n\, \sqrt{\hat{\nu}} s\, \sigma(\hat{\nu}) \Big\} = 0, \\
l_n(T, s, \boldsymbol{\nu}, v) = L\Big(& h(s), v \Big).
\end{aligned} \tag{20}$$

Proposition 10. *The primal optimal control $(\xi_t^{n,*})_{0\le t\le T}$ is given by*

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

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

$$\begin{aligned}
& \partial_t l_n + \mathcal{L}_{s,\boldsymbol{\nu}} l_n + \partial_v l_n\, rv - \frac{\Big(\partial_v l_n\, (\hat{\mu} - r) s + \partial_{sv} l_n\, \hat{\nu} s^2 + \rho \sum_{i=1}^n \partial_{\nu_i v} l_n\, \sqrt{\hat{\nu}} s\, \sigma(\hat{\nu}) \Big)^2}{2 \partial_{vv} l_n\, \hat{\nu} s^2} = 0, \\
& l_n(T, s, \boldsymbol{\nu}, v) = L\Big(h(s), v \Big).
\end{aligned} \tag{21}$$

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

Proof. To deduce the optimal control associated with the Markovian control problem, we solve the HJB equation (20). The HJB equation has a solution if the infimum is different from $-\infty$, it is the case if $\partial_{vv} l_n \ge 0$. In this case, assuming that $\partial_{vv} l_n \ge 0$, 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

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

as $\hat{\nu} > 0$, 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

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

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

□

The PDE satisfied by the optimal solution is nonlinear, therefore we cannot reduce l_n as an expectation using the 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 (19) and consider the concave dual $\hat{l}_n(.)$ of $l_n(.)$ with respect to the variable v as the additive inverse of the Legendre-Fenchel transform, such that

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

for $(t,s,\boldsymbol{\nu},z) \in [0,T]\times\mathbb{S}$. 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

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

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

$$\begin{aligned}0 =& \partial_t\hat{l}_n + \mathcal{L}_{s,\boldsymbol{\nu}}\hat{l}_n - zr\partial_z\hat{l}_n + \frac{1}{2\hat{\nu}s^2}z^2\,(\hat{\mu}-r)^2s^2\partial_{zz}\hat{l}_n - z\,(\hat{\mu}-r)s\partial_{sz}\hat{l}_n,\\ &-\frac{1}{\sqrt{\hat{\nu}}}\rho\sum_{i=1}^n\partial_{\nu_i z}\hat{l}_n\,z\,(\hat{\mu}-r)\sigma(\hat{\nu}) - \frac{1}{2\partial_{zz}\hat{l}_n}\,\sigma(\hat{\nu})^2(1-\rho^2)\,\sum_{i=1}^n\sum_{j=1}^n\partial_{\nu_i z}\hat{l}_n\partial_{\nu_j z}\hat{l}_n,\end{aligned} \tag{22}$$

with the associated terminal value $\hat{l}_n(T,s,\boldsymbol{\nu},z) = \hat{L}(h(s),z)$.

Proof. The proof is provided in Appendix A. □

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

$$dZ_t = -rZ_tdt - Z_t\frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}}dW_s(t) + \gamma_t dB_\nu(t),$$

on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{0\le t\le T},\mathbb{P})$. The dual control process $(\gamma_t)_{0\le t\le T}$ is admissible if $(\gamma_t)_{0\le t\le T}$ 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 12. *The dual solution $\hat{l}_n(t,s,\boldsymbol{\nu},z)$ is the solution of a stochastic control problem such that*

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

with $\mathcal{D}$ 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 $(\gamma_t^*)_{0\le t\le T}$ is given by

$$\gamma_t^* = -\sigma(\hat{\nu}_t)\sqrt{1-\rho^2}\sum_{i=1}^n\frac{\partial_{\nu_i z}\hat{l}_n}{\partial_{zz}\hat{l}_n}. \tag{23}$$

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

$$\hat{l}_n^{bis}(t,s,\nu,z) := \sup_{\gamma_t\in\mathcal{D}} E_{t,s,\boldsymbol{\nu},z}\bigg(\hat{L}(h(S_T^n),Z_T)\bigg). \tag{24}$$

We just have to prove that the HJB equation associated to $\hat{l}_n^{bis}$ matches the PDE (22). The HJB equation associated to the control problem (24) is given by

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

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, $\hat{l}_n^{bis}$ is the solution of the following PDE

$$0 = \partial_t \hat{l}_n^{bis} + \mathcal{L}_{s,\boldsymbol{\nu}} \hat{l}_n^{bis} - z\, r \partial_z \hat{l}_n^{bis} + \frac{1}{2} \partial_{zz} \hat{l}_n^{bis} z^2 \frac{(\hat{\mu} - r)^2}{\hat{\nu}^2} - \partial_{zs} \hat{l}_n^{bis}\, z (\hat{\mu} - r) s$$
$$- \sum_{i=1}^{n} \partial_{\nu_i z} \hat{l}_n^{bis}\, \rho\, z \frac{(\hat{\mu} - r)}{\sqrt{\hat{\nu}}} \sigma(\hat{\nu}) - \frac{1}{2\partial_{zz} \hat{l}_n^{bis}} \sigma^2(\hat{\nu})(1-\rho^2) \sum_{i=1}^n \sum_{j=1}^n \partial_{\nu_i z} \hat{l}_n^{bis} \partial_{\nu_j z} \hat{l}_n^{bis}, \tag{25}$$

We can observe that (25) is exactly the same PDE as (22) and as the two PDE's have the same terminal value, we can conclude by unicity 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 13. *By choosing $z(t, s, \boldsymbol{\nu}, v)$ solution of*

$$\partial_z \hat{l}_n(t, s, \boldsymbol{\nu}, z) + v = 0, \tag{26}$$

then

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

with $z = z(t, s, \boldsymbol{\nu}, v)$ the value at time t of the dual process $(Z_t)_{0 \leq t \leq T}$. Moreover, the optimal primal control $(\xi_t^{n,*})_{0 \leq t \leq T}$ can be expressed in term of dual solution such that, for $t \in [0, T]$,

$$\xi_t^{n,*} = \frac{z(t, S_t^n, \boldsymbol{\nu}_t, V_t^n)\, \partial_{zz} \hat{l}_n\, (\hat{\mu}_t - r) S_t^n - \partial_{sz} \hat{l}_n\, \hat{\nu}_t (S_t^n)^2 - \rho \sum_{i=1}^n \partial_{\nu_i z} \hat{l}_n \sqrt{\hat{\nu}_t} S_t^n\, \sigma(\hat{\nu}_t)}{\hat{\nu}_t (S_t^n)^2}, \tag{28}$$

with $z(t, s, \boldsymbol{\nu}, v)$ solution of (26).

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

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

Our goal is to prove that the dual of the dual is the primal. 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

$$\partial_z \hat{l}_n(t, s, \boldsymbol{\nu}, z) + v = 0, \tag{29}$$

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 (29). 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 11 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 (21). 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 continuous 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 by unicity 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

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

Finally, for $t \in [0,T]$, using the form of the optimal primal control $\xi_t^{n,*}$ and the expression of the primal value $l_n(.)$ as function of the dual value $\hat{l}_n(.)$, we deduce that

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

□

As the primal solution and the optimal primal control are functions of the dual solution, we can derive those expressions 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 (22) 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

$$H_T^{linear} := \alpha + \beta S_T, \tag{30}$$

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 the 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 14. *Suppose a power loss of the form $L(h,v) = \frac{1}{p}(h-v)^p$ with $p = 2k,\ k \in \mathbb{N}_0$, 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 A. □

Proposition 15. *Consider a linear payoff H_T^{linear} given by (30), a power loss of the form $L(h(s), v) = \frac{1}{p}(h(s) - v)^p$ with $p = 2k$, $k \in \mathbb{N}_0$, suppose that the volatility is modeled by a approximate rough Heston model i.e. $b(x) = \kappa(\theta - x)$ and $\sigma(x) = \zeta\sqrt{x}$ with $2\kappa\theta > \zeta^2 \sum_{i=1}^n w_i$ and $\nu_0 > 0$ such that (18) is satisfied. Therefore,*

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

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

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

with

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

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 A. 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

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

and plug it into PDE (22). Then using the form of $\hat{l}_n(.)$ and the relation (27) 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 (28) of $\xi_t^{n,*}$ and the closed form of the dual solution. □

Remark. Using the same assumptions as Proposition 15 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, see for instance Appendix A in [36].

For linear payoffs, the nonlinearity of the PDE satisfied by the dual solution[8] disappears because the Vega[9] is zero which is not the case for more general payoffs. As it is not possible to obtain explicit forms of the primal value function and the optimal hedging ratios for general payoffs, we are interested in deducing approximate solutions close enough to the optimal one. To approximate the solution of our problem in the general case, 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

[8] In the rough Heston case with a power loss function.

[9] Measure of option's price sensitivity to changes in the volatility of the risky asset.

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

$$
\begin{aligned}
l_n(t,s,\boldsymbol{\nu},v) &= \sup_z \left\{ \sup_{\gamma_t \in \mathcal{D}} E_{t,s,\boldsymbol{\nu},z}\left(\hat{L}(h(S_T^n), Z_T) \right) + zv \right\} \\
&\geq \sup_z \left\{ \sup_{\gamma_t \in \mathcal{U}} E_{t,s,\boldsymbol{\nu},z}\left(\hat{L}(h(S_T^n), Z_T) \right) + zv \right\}. \qquad (31)
\end{aligned}
$$

Based on inequality (31) and inspired by the dual control method stated in [36], 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

$$
Y(t,s,\boldsymbol{\nu},z;\gamma) := E_{t,s,\boldsymbol{\nu},z}\left(\hat{L}(\, h(S_T^n), Z_T(\gamma)\,) \right), \qquad (32)
$$

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

Theorem 16. *Let $\mathcal{U} \subseteq \mathcal{D}$ be a set of admissible 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 continuously differentiable, stricly concave and $z(t,s,\boldsymbol{\nu},v)$ is the solution of

$$
\partial_z \underline{\hat{l}}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},z) + v = 0. \qquad (33)
$$

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

$$
\bar{\xi}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) = \frac{z(t,s,\boldsymbol{\nu},v)\, \partial_{zz}\underline{\hat{l}}_n^{\mathcal{U}}\, (\hat{\mu}-r)s - \partial_{sz}\underline{\hat{l}}_n^{\mathcal{U}}\, \hat{\nu}s^2 - \rho \sum_{i=1}^n \partial_{\nu_i z}\underline{\hat{l}}_n^{\mathcal{U}}\, \sqrt{\hat{\nu}} s\, \sigma(\hat{\nu})}{\hat{\nu}s^2}, \qquad (34)
$$

with $z(t,s,\boldsymbol{\nu},v)$ solution of (33). We consider the approximate hedging process

$$
\xi_t^{n,\mathcal{U}} := \bar{\xi}_n^{\mathcal{U}}(t, S_t^n, \boldsymbol{\nu}_t, V_t^{n,\mathcal{U}}) \qquad (35)
$$

and the associated self-financing portfolio denoted by $(V_t^{n,\mathcal{U}})_{0\leq t\leq T}$ satisfying SDE (11) with $\xi_t^n = \xi_t^{n,\mathcal{U}}$. If we define

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

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,

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

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 [36]. 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 $(\xi_t^{n,\mathcal{U}})_{0\leq t\leq T}$ defined by (35) has the same form as the optimal hedging ratio $(\xi_t^{n,*})_{0\leq t\leq T}$ defined in (28) with the difference that we consider the subset of admissible dual control $\mathcal{U}$ instead of $\mathcal{D}$.

Theorem 16 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 [36] 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 obtain explicit formulas for the lower bound and the approximate hedging ratio. We discuss later the choice of the dual control subset.

Theorem 16 introduces a sub-optimal hedging strategy associated with the approximate hedging ratio $(\xi_t^{n,\mathcal{U}})_{0\leq t\leq T}$ 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:

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

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 17. *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 $\lim_{i\to+\infty} \mathcal{U}_i = \mathcal{D}$. For $n \geq 1$, suppose that the sequence of functions $\left(\underline{\hat{l}}_n^{\mathcal{U}_i}(.)\right)_{i\in\mathbb{N}}$ is twice continuously* *differentiable with second derivatives that converge uniformly in* $\mathbb{R}$ *and the sequence* $\left(L(h(S_T^n), V_T^{n,\mathcal{U}_i})\right)_{i\in\mathbb{N}}$ *is uniformly integrable, then* $\forall t \in [0,T]$,

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

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 $\mathcal{U}_i \subseteq \mathcal{U}_{i+1}$ and $\lim_{i\to+\infty}\mathcal{U}_i = \mathcal{D}$ and fix $n \geq 1$. First, we can show the convergence of $\hat{\underline{l}}_n^{\mathcal{U}_i}(t,s,\boldsymbol{\nu},v)$ 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 $\forall t \in [0,T]$,

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

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

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

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

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

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 $\left(\hat{\underline{l}}_n^{\mathcal{U}_i}(.)\right)_{i\in\mathbb{N}}$ is twice continuously differentiable with second derivatives that converge uniformly in $\mathbb{R}$, standard result in Analysis states that $\forall t \in [0,T]$,

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

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

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

Thus, since we have that

$$\xi_t^{n,\mathcal{U}_i} \xrightarrow{\mathcal{L}} \xi_t^{n,*},$$

as i goes to infinity, using similar arguments as for the proof of Proposition 6, we deduce that

$$V_T^{n,\mathcal{U}_i} \xrightarrow{\mathcal{L}} V_T^{n,*},$$

as i goes to infinity. Moreover, as $\Big(L(h(S_T^n), V_T^{n,\mathcal{U}_i})\Big)_{i\in\mathbb{N}}$ is uniformly integrable, using similar arguments as for the proof of Lemma 7, we obtain that, $\forall t \in [0,T]$,

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

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

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

That concludes the proof since we have proved 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

$$l(s_0,\nu_0,v_0) = \inf_{\xi_t\in\mathcal{R}} E_{s_0,\nu_0,v_0}\Big(L(h(S_T),V_T)\Big).$$

The proposed sub-optimal hedging strategy is $(\xi_t^{n,\mathcal{U}})_{0\le t\le T}$ associated with the approximate initial value function $\bar{l}_n^{\mathcal{U}}(s_0,\boldsymbol{\nu_0},v_0)$ defined by

$$\bar{l}_n^{\mathcal{U}}(s_0,\boldsymbol{\nu_0},v_0) := \bar{l}_n^{\mathcal{U}}(t=0,s_0,\boldsymbol{\nu_0},v_0),$$

with $\boldsymbol{\nu_0} = \boldsymbol{0^n}$. It 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 18. *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 $\lim_{i\to+\infty}\mathcal{U}_i = \mathcal{D}$. Suppose that the assumptions of Theorem 8 and Proposition 17 are satisfied. $\forall\varepsilon > 0$, $\exists N \in \mathbb{N}$, such that $\forall\, n \ge N$, $\exists M \in \mathbb{N}$ such that $\forall i \ge M$,*

$$|l(s_0,\nu_0,v_0) - \bar{l}_n^{\mathcal{U}_i}(s_0,\boldsymbol{\nu_0},v_0)| < \varepsilon$$

and

$$|l(s_0,\nu_0,v_0) - l^{\xi^{n,\mathcal{U}_i}}(s_0,\nu_0,v_0)| < \varepsilon,$$

with $\boldsymbol{\nu_0} = \boldsymbol{0^n}$. It means that the approximate hedging ratio $(\xi_t^{n,\mathcal{U}_i})_{0\le t\le T}$ associated to $\bar{l}_n^{\mathcal{U}_i}(s_0,\boldsymbol{\nu_0},v_0)$ is $\varepsilon-$optimal for the original problem.

Proof. The proof is almost direct. Fix $\varepsilon > 0$, from Theorem 8, we know that $\exists N \in \mathbb{N}$ such that $\forall n \ge N$,

$$|l(s_0,\nu_0,v_0) - l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0)| < \frac{\varepsilon}{2},$$

with $\boldsymbol{\nu_0} = \boldsymbol{0^n}$, and

$$|l(s_0,\nu_0,v_0) - l^{\xi^{n,*}}(s,\nu_0,v)| < \frac{\varepsilon}{2}.$$

Moreover, from Proposition 17, we have that $\forall n \ge 1$, $\exists M_1 \in \mathbb{N}$ such that $\forall i \ge M_1$,

$$|l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0) - \bar{l}_n^{\mathcal{U}_i}(s_0,\boldsymbol{\nu_0},v_0)| < \frac{\varepsilon}{2},$$

and since $\xi_t^{n,\mathcal{U}_i} \xrightarrow{\mathcal{L}} \xi_t^{n,*}$ as $i \to +\infty$, we deduce, using similar arguments as for the proof of Proposition 6 and Lemma 7, that $\forall n \ge 1$, $\exists M_2 \in \mathbb{N}$ such that $\forall i \ge M_2$,

$$|l^{\xi^{n,*}}(s_0,\nu_0,v_0) - l^{\xi^{n,\mathcal{U}_i}}(s_0,\nu_0,v_0)| < \frac{\varepsilon}{2}.$$

Therefore choosing $M := \max(M_1, M_2)$, for $\boldsymbol{\nu_0} = \boldsymbol{0^n}$, we know that $\forall n \geq N$, $\exists M \in \mathbb{N}$ such that $\forall i \geq M$,

$$\begin{aligned}|l(s_0,\nu_0,v_0) - \bar{l}_n^{\mathcal{U}_i}(s_0,\boldsymbol{\nu_0},v_0)| =& |l(s_0,\nu_0,v_0) - l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0) + l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0) - \bar{l}_n^{\mathcal{U}_i}(s_0,\boldsymbol{\nu_0},v_0)| \\ \leq& |l(s_0,\nu_0,v_0) - l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0)| + |l_{n,\boldsymbol{\nu}}(s_0,\boldsymbol{\nu_0},v_0) - \bar{l}_n^{\mathcal{U}_i}(s_0,\boldsymbol{\nu_0},v_0)| \\ <& \varepsilon,\end{aligned}$$

and

$$\begin{aligned}|l(s_0,\nu_0,v_0) - l^{\xi^{n,\mathcal{U}_i}}(s_0,\nu_0,v_0)| &= |l(s_0,\nu_0,v_0) - l^{\xi^{n,*}}(s_0,\nu_0,v_0) + l^{\xi^{n,*}}(s_0,\nu_0,v_0) - l^{\xi^{n,\mathcal{U}_i}}(s,\nu_0,v_0)| \\ &\leq |l(s_0,\nu_0,v_0) - l^{\xi^{n,*}}(s_0,\nu_0,v_0)| + |l^{\xi^{n,*}}(s_0,\nu_0,v_0) - l^{\xi^{n,\mathcal{U}_i}}(s_0,\nu_0,v_0)| \\ &< \varepsilon.\end{aligned}$$

□

The result of Proposition 18 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 sets. 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

$$\mathcal{U} = \left\{ (\gamma_t)_{0\leq t\leq T} = \Big(c \times Z_t \times \sigma(\hat{\nu}_t) \Big)_{0\leq t\leq T}, \ c \in U \subseteq \mathbb{R} \right\} \subseteq \mathcal{D}. \tag{36}$$

We notice that the chosen form of the dual controls belonging to $\mathcal{U}$ is similar to the form of the optimal dual control (23). This particular subset (36) 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

$$\begin{aligned}\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v) &= \sup_z \left\{ \sup_{\gamma_t \in \mathcal{U}} Y(t,s,\boldsymbol{\nu},z;\gamma) + zv \right\} \\ &= \sup_z \left\{ \max_{c\in U} Y(t,s,\boldsymbol{\nu},z;c) + zv \right\} &(37) \\ &= \max_{c\in U} \sup_z \left\{ Y(t,s,\boldsymbol{\nu},z;c) + zv \right\}. &(38)\end{aligned}$$

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

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

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 value at time $t \in [0,T]$ of the sub-optimal hedging portfolio $V_t^{n,\mathcal{U}}$ is easily interpreted as the price at time t of a modified payoff. This is in line with Föllmer and Leukert [19, 20] who present partial hedging strategies as perfect hedging of knock-out options.

Proposition 19. *If we consider a subset of admissible dual control $\mathcal{U}$ of the form (36), for $t \in [0,T]$, if $z = z(t,s,\boldsymbol{\nu},v)$ solution of (33) with $v = V_t^{n,\mathcal{U}}$, then*

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

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

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

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

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

on a filtred probability space $(\Omega, \mathcal{F}^{\mathbb{Q}(c)}, (\mathcal{F}_t^{\mathbb{Q}(c)})_{0\le t\le T}, \mathbb{Q}(c))$, where $\mathcal{F}_t^{\mathbb{Q}(c)}$ is the canonical filtration of a two-dimensional Brownian motion $(W_S^{\mathbb{Q}(c)}, B_\nu^{\mathbb{Q}(c)})$ and $W_\nu^{\mathbb{Q}(c)} = \rho\, W_S^{\mathbb{Q}(c)} + \sqrt{1-\rho^2}B_\nu^{\mathbb{Q}(c)}$. In particular, for a power loss function of the form $L(h(s),v) = \frac{1}{p}(h(s)-v)^p$ with $p = 2k$, $k \in \mathbb{N}_0$, the value sub-optimal hedging portfolio at time $t \in [0,T]$ is given by

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

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

Proof. From the Theorem 16, we know, using (33), that, at time $t \in [0,T]$, $z = z^* = z(t,s,\boldsymbol{\nu},v)$ with $z(t,s,\boldsymbol{\nu},v)$ solution of

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

Thus, by (37), 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 (38), 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 $t \in [0,T]$, $z = z^* = z(t,s,\boldsymbol{\nu},v) = z^*(c^*)$ and $c^*(z^*) = c^*$. Therefore, we obtain, at time $t \in [0,T]$ and for $z = z^*(c^*)$ with $v = V_t^{n,\mathcal{U}}$, the following relation

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

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

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

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

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

and therefore, for $t \in [0,T]$,

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

Since $\left(\frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}},\ c\,\sigma(\hat{\nu}_t) \right)_{0\le t\le T}$ 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

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

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

$$\begin{aligned} dS_t^n &= rS_t^n + \sqrt{\hat{\nu}_t} S_t^n dW_s^{\mathbb{Q}(c)}(t), \\ d\nu_t^i &= \Big(-x_i\nu_t^i + b(\hat{\nu}_t) + \sigma(\hat{\nu}_t)\Big(-\rho\frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}} + \sqrt{1-\rho^2} c\,\sigma(\hat{\nu}_t) \Big) \Big) dt + \sigma(\hat{\nu}_t) dW_\nu^{\mathbb{Q}(c)}(t),\ i = 1, ..., n, \\ dZ_t &= Z_t \bigg(\Big(-r + \frac{(\hat{\mu}_t - r)}{\hat{\nu}_t}^2 + c^2\sigma^2(\hat{\nu}_t) \Big) dt - \frac{(\hat{\mu}_t - r)}{\sqrt{\hat{\nu}_t}} dW_S^{\mathbb{Q}}(t) + c\,\sigma(\hat{\nu}_t) dB_\nu^{\mathbb{Q}(c)}(t) \bigg) \end{aligned}$$

where $W_S^{\mathbb{Q}(c)}$, $W_\nu^{\mathbb{Q}(c)}$ are standard Brownian motions under $\mathbb{Q}(c)-$measure with $d\langle W_S^{\mathbb{Q}(c)}, W_\nu^{\mathbb{Q}(c)} \rangle_t = \rho\, dt$ and $B_\nu^{\mathbb{Q}(c)}$ is a standard Brownian motion, independent from $W_S^{\mathbb{Q}(c)}$ such that $W_\nu^{\mathbb{Q}(c)} = \rho W_S^{\mathbb{Q}(c)} + \sqrt{1-\rho^2} B_\nu^{\mathbb{Q}(c)}$. In this case we have that

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

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

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

□

Still assuming that the subset of admissible dual controls $\mathcal{U}$ has the form (36), we next show that, for the rough Heston model, the lower bond as well as the approximate hedging ratio associated with a power loss function have explicit forms.

Proposition 20. *Consider a power loss of the form $L(h(s),v)=\frac{1}{p}(h(s)-v)^p$ with $p=2k$, $k\in\mathbb{N}_0$, suppose that the volatility is modeled by a approximate rough Heston model i.e. $b(x)=\kappa(\theta-x)$ and $\sigma(x)=\zeta\sqrt{x}$ with $2\kappa\theta>\zeta^2\sum_{i=1}^n w_i$ and $\nu_0>0$ such that (18) is satisfied. Moreover, assume that the subset of admissible dual control is given by (36). Therefore $Y(t,s,\boldsymbol{\nu},z;\gamma)$ defined by (32) is such that*

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

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

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

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

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

In this case, the lower bound satisfies

$$\underline{l}_n^{\mathcal{U}}(t,s,\boldsymbol{\nu},v)=-\frac{1}{q}\,z(t,s,\boldsymbol{\nu},v)^q\exp(C_t(c^*)+\sum_{i=1}^n D_t^i(c^*)\nu_t^i)+\,z(t,s,\boldsymbol{\nu},v)\,\Big(v-E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c^*)}(e^{-r(T-t)}H_T^n)\Big),\quad(39)$$

with

$$z(t,s,\boldsymbol{\nu},v)=\exp\bigg(-\frac{1}{q-1}(C_t(c^*)+\sum_{i=1}^n D_t^i(c^*)\nu^i)\bigg)\times\Big(v-E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c^*)}(e^{-r(T-t)}H_T^n))\Big)^{\frac{1}{q-1}}.\qquad(40)$$

Moreover, if $z=z(t,s,\boldsymbol{\nu},v)$,

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

Proof. The proof is similar to the proof of Proposition 15 and is essentially obtained by using the Feynman-Kac formula. First, using Lemma 14, we have that

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

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

$$\begin{aligned}0=&\partial_tY_1-\partial_zY_1\,rz+\frac{1}{2}\partial_{zz}Y_1\,z^2\Big(\frac{(\hat{\mu}-r)^2}{\hat{\nu}}+c^2\hat{\nu}\zeta^2\Big)+\sum_{i=1}^n\partial_{\nu_i}Y_1\,(-x_i\nu_i+\kappa(\theta-\hat{\nu}))\\&+\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n\partial_{\nu_i\nu_j}Y_1\,\zeta^2\hat{\nu}+\sum_{i=1}^n\partial_{\nu_iz}Y_1\,z\Big(-\rho\frac{(\hat{\mu}-r)}{\sqrt{\hat{\nu}}}\zeta\sqrt{\hat{\nu}}+\sqrt{1-\rho^2}\zeta^2\hat{\nu}c\Big),\\Y_1(T,z)&=-\frac{1}{q}z^q.\end{aligned}$$

Suppose that Y_1 has the following form

$$Y_1 = -\frac{1}{q} z^q \exp(C_t + \sum_{i=1}^n D_t^i \nu^i),$$

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

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

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

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

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

$$\begin{aligned}
\partial_t C_t = & rq - \frac{1}{2} q(q-1)\,(A^2 + c^2\zeta^2)\nu_0 - \sum_{i=1}^n D_t^i \Big(\kappa(\theta - \nu_0) + q\Big(-\rho A \nu_0 \zeta + \sqrt{1-\rho^2}\zeta^2 \nu_0 c\Big)\Big) && (41) \\
& - \frac{1}{2}\zeta^2 \nu_0 \sum_{i=1}^n \sum_{j=1}^n D_t^i D_t^j, && (42) \\
C_T = & 0
\end{aligned}$$

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

$$\begin{aligned}
\partial_t D_t^i = & x_i D_t^i + w_i \sum_{j=1}^n D_t^j \Big(\kappa - q\Big(-\rho A \zeta + \sqrt{1-\rho^2}\zeta^2 c\Big)\Big) - \zeta^2 w_i \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n D_t^j D_t^k \\
& - \frac{1}{2} w_i q(q-1)\,(A^2 + c^2\zeta^2), && (43) \\
D_T^i = & 0.
\end{aligned}$$

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

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

Suppose that Y_2 has the form

$$Y_2 = g(t,s,\boldsymbol{\nu})\, z.$$

Therefore, we have that

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

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

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

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

$$g(t,s,\boldsymbol{\nu}) = E^{\mathbb{Q}(c)}_{t,s,\boldsymbol{\nu}}(e^{-r(T-t)} H^n_T),$$

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

$$\begin{aligned}
dS^n_t &= rS^n_t + \sqrt{\hat{\nu}_t} S^n_t dW^{\mathbb{Q}(c)}_S(t), \\
d\nu^i_t &= \left(- x_i(\nu^i_t - \nu^i_0) + \kappa(\theta - \hat{\nu}) + \hat{\nu}\zeta(-\rho A + \sqrt{1-\rho^2}\zeta\, c)\right)dt + \zeta\sqrt{\hat{\nu}_t} dW^{\mathbb{Q}}_\nu(t),\ i = 1, ..., n,
\end{aligned}$$

where $W^{\mathbb{Q}(c)}_S$ and $W^{\mathbb{Q}(c)}_\nu$ are standard brownian motions under $\mathbb{Q}(c)-$measure with $d\langle W^{\mathbb{Q}(c)}_S, W^{\mathbb{Q}(c)}_\nu\rangle_t = \rho\, dt$. Finally, combining the different results, we obtain the annonced result

$$Y(t,s,\boldsymbol{\nu},z;c) = -\frac{1}{q} z^q \exp(C_t(c) + \sum_{i=1}^n D^i_t(c)\nu^i) - E^{\mathbb{Q}(c)}_{t,s,\boldsymbol{\nu}}(e^{-r(T-t)} H^n_T)\, z.$$

Furthermore, as by definition,

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

using the first order condition, we obtain that, given $c \in U$, the value at time t of $z(t,s,\boldsymbol{\nu},v;c)$ satisfies

$$z(t,s,\boldsymbol{\nu},v;c) = \exp\left(- \frac{1}{q-1}(C_t(c) + \sum_{i=1}^n D^i_t(c)\nu^i)\right) \times \left(v - E^{\mathbb{Q}(c)}_{t,s,\boldsymbol{\nu}}(e^{-r(T-t)} H^n_T)\right)^{\frac{1}{q-1}}$$

and then

$$\underline{l}_n(t,s,\boldsymbol{\nu},v;c) = -\frac{1}{q} z(t,s,\boldsymbol{\nu},v;c)^q \exp(C_t(c) + \sum_{i=1}^n D^i_t(c)\nu^i_t) + z(t,s,\boldsymbol{\nu},v;c)\left(v - E^{\mathbb{Q}(c)}_{t,s,\boldsymbol{\nu}}(e^{-r(T-t)} H^n_T)\right).$$

Thus, as

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

we deduce the annonced result. Finally, as in the proof of Proposition 19 , we show that $c^* = c^*(z^*)$ and $z(t,s,\boldsymbol{\nu},v;c^*) = z(t,s,\boldsymbol{\nu},v)$, we conclude that if $z = z(t,s,\boldsymbol{\nu},v)$ then

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

□

Remark. Compared to the optimal solution in the case of linear payoff (see Proposition 15), we observe that the form of the solution is the same, the only difference lies in the expression of the coefficients C_t and $(D_t^i)_{i=1,\dots,n}$. In the classical Heston model, i.e. $n=1$, then C_t and D_t admit closed formulas, see for instance Appendix A in [36].

Proposition 21. *Using the same assumptions as the Proposition 20 and denoting $g(t,s,\boldsymbol{\nu}) = E_{t,s,\boldsymbol{\nu}}^{\mathbb{Q}(c^*)}(e^{-r(T-t)}H_T^n)$, the approximate hedging ratio $\xi_t^{n,\mathcal{U}}$ defined by (35) is such that*

$$\xi_t^{n,\mathcal{U}} = \frac{1}{\hat{\nu}_t S_t^n}\Big(g(t,S_t^n,\boldsymbol{\nu}_t) - V_t^{n,\mathcal{U}}\Big)\Big(\frac{1}{p-1}(\hat{\mu}_t - r) - \rho\zeta\hat{\nu}_t\sum_{i=1}^n D_t^i(c^*)\Big)$$
$$+ \partial_s g(t,S_t^n,\boldsymbol{\nu}_t) + \frac{\rho\zeta}{S_t^n}\sum_{i=1}^n\Big(\partial_{\nu_i} g(t,S_t^n,\boldsymbol{\nu}_t)\Big), \tag{44}$$

where $V_t^{n,\mathcal{U}}$ is the value at time t of the self-financing portfolio associated to the sub-optimal hedging strategy.

Proof. The proof is simple and relies on the definition of the approximate hedging ratio and on Proposition 20. By definition, we have that

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

with $z = z(t,s,\boldsymbol{\nu},v)$. With our assumptions, we know, by Proposition 20, that, for $z = z(t,s,\boldsymbol{\nu},v)$,

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

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

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

and

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

therefore, the approximate hedging ratio process $\xi_t^{n,\mathcal{U}} = \bar{\xi}_n^{\mathcal{U}}(t,S_t^n,\boldsymbol{\nu_t},V_t^{n,\mathcal{U}})$ is given by

$$\xi_t^{n,\mathcal{U}} = \frac{1}{\hat{\nu}_t\Big(S_t^n\Big)^2}\Bigg(-(q-1)\,z(t,S_t^n,\boldsymbol{\nu_t},V_t^{n,\mathcal{U}})^{q-1}\exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i)\,(\hat{\mu}_t - r)S_t^n + \partial_s g(t,S_t^n,\boldsymbol{\nu}_t)\,\hat{\nu}_t\Big(S_t^n\Big)^2$$
$$+ \rho\zeta S_t^n\hat{\nu}_t\sum_{i=1}^n\Big(z(t,S_t^n,\boldsymbol{\nu_t},V_t^{n,\mathcal{U}})^{q-1}D_t^i(c^*)\exp(C_t(c^*) + \sum_{i=1}^n D_t^i(c^*)\nu_t^i) + \partial_{\nu_i} g(t,S_t^n,\boldsymbol{\nu}_t)\Big)\Bigg). \tag{45}$$

Using the form (40) of the function $z(t,s,\boldsymbol{\nu},v)$, the approximate hedging ratio becomes

$$\xi_t^{n,\mathcal{U}} = \frac{1}{\hat{\nu}_t S_t^n}\Big(g(t,S_t^n,\boldsymbol{\nu}_t) - V_t^{n,\mathcal{U}}\Big)\Big(\frac{1}{p-1}(\hat{\mu}_t - r) - \rho\zeta\hat{\nu}_t\sum_{i=1}^n D_t^i(c^*)\Big)$$
$$+ \partial_s g(t,S_t^n,\boldsymbol{\nu}_t) + \frac{\rho\zeta}{S_t^n}\sum_{i=1}^n\Big(\partial_{\nu_i} g(t,S_t^n,\boldsymbol{\nu}_t)\Big).$$

That concludes the proof since we proved the announced result. □

For the partial hedging problem in the rough Heston model with power loss function, the lower bound $\underline{l}_n^{\mathcal{U}}(.)$ and the approximate hedging ratio $\xi_t^{n,\mathcal{U}}$ admit explicit formulas. In particular, we observe that $\xi_t^{n,\mathcal{U}}$ 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 C.

6 Numerical results

In this section, we illustrate the partial hedging method discussed in this paper for the rough Heston model. Without loss of generality, we consider the following parameters (under the real measure $\mathbb{P}$) for the rough Heston model:

$$r = 0.02,\ A = 1,\ S_0 = 100,\ \nu_0 = 0.09,\ \theta = 0.09,\ \kappa = 2.95,\ \zeta = 0.3\ and\ \rho = -0.7. \tag{46}$$

Remark. With those parameters and $T = 0.25$, the Feller's condition for the approximate rough Heston model (see Appendix B) is satisfied for different Hurst indexes H and number of factors n. For example, if $H = 0.1$, the sufficient condition of strict positivity of the volatility is satisfied for all $n \leq 20$. Figure 1 shows the region for which the Feller's condition is satisfied as a function of the Hurst coefficient H and the number of factors n.

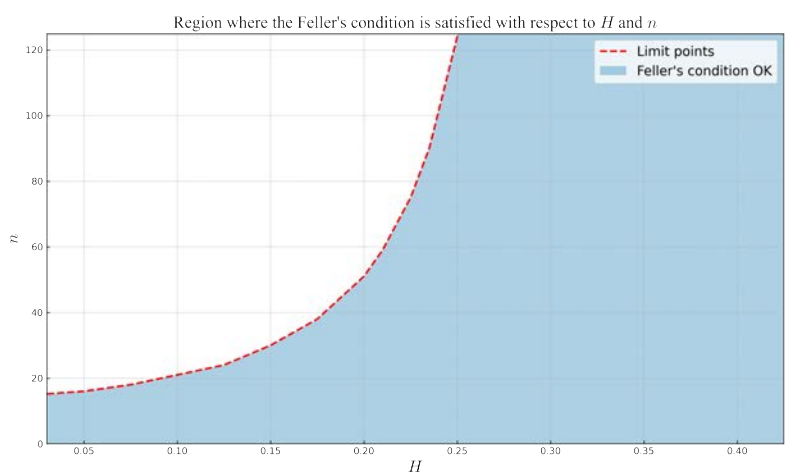

Figure 1: Region where the Feller's condition is satisfied with respect to H and n for the approximate rough Heston with parameters (46) and $T = 0.25$.

As already mentioned, rough volatility models and their Markovian approximation, including the (approximate) rough Heston model, allow to capture stylized facts observed on the financial markets. In particular, in option pricing, they allow to 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. As discussed in [1], the difference between the characteristics (implied volatility smile and ATM skew) of the rough Heston and its Markov approximation can be negligible even with a reduced number of factors (for example, $n = 20$). Thus, as $n = 20$ is sufficient for a relevant approximation of the rough Heston model, without loss of generality, we take $n = 20$ for most of the numerical results regarding partial hedging strategies. For more analysis and discussions about the rough Heston and its Markov approximation, we refer among others to [1, 2, 6, 18, 30].

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

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

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 $n \to +\infty$. Figures 2 and 3 present the evolution of l_n and $\xi_0^{n,*}$ with respect to n for different values of H[10] and with $V_0 = 0.8 \times H_0^{linear}$. We observe that, as expected from the theoretical results, the approximate solution and initial hedging ratio converge for different Hurst coefficients.

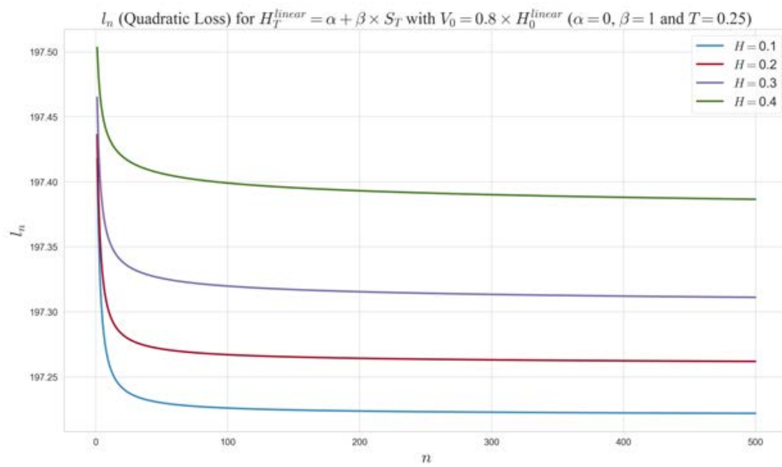

Figure 2: Convergence analysis of l_n for quadratic loss and linear payoff with $\alpha = 0$, $\beta = 1$ and $T = 0.25$

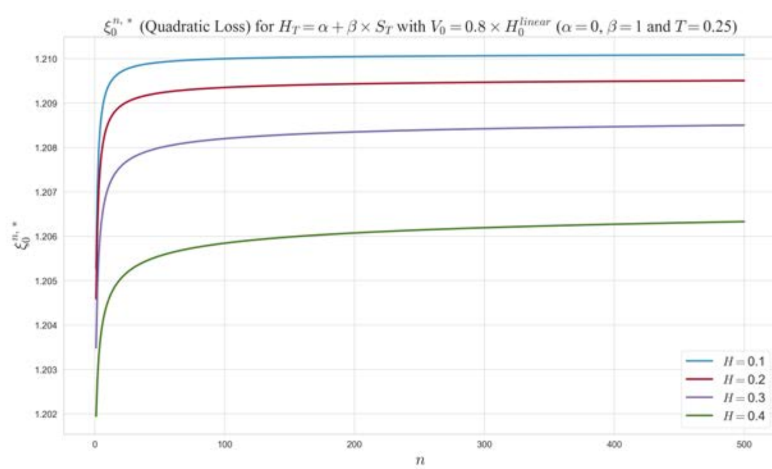

Figure 3: Convergence analysis of $\xi_0^{n,*}$ for quadratic loss and linear payoff with $\alpha = 0$, $\beta = 1$ and $T = 0.25$

[10] Results are generated to check the convergence but we have to keep in mind that the sufficient condition of strict positivity of the volatility is not satisfied for all n and H.

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:

$$\gamma_t = c \times Z_t \times \zeta\sqrt{\hat{\nu}_t},\ c \in U = \mathbb{R}, \tag{47}$$

because it allows to keep closed formulas under the approximate rough Heston model. In this case, the associated lower bound is given by (39) and the upper bound can be computed with the hedging ratio given by (44).

$Coef$	H	l_n	$\underline{l}_n^{\mathcal{U}}$	$\bar{l}_n^{\mathcal{U}}$	Abs. diff. $\lvert l_n - \underline{l}_n^{\mathcal{U}}\rvert$	Abs. diff. $\lvert\bar{l}_n^{\mathcal{U}} - l_n\rvert$	Abs. diff. $\lvert\bar{l}_n^{\mathcal{U}} - \underline{l}_n^{\mathcal{U}}\rvert$
0.75	0.1	308.1899	308.1888	308.1925	1.11×10^{-3}	2.64×10^{-3}	3.71×10^{-3}
0.75	0.5	308.3811	308.3759	308.3853	5.26×10^{-3}	4.28×10^{-3}	9.42×10^{-3}
0.5	0.1	1232.7597	1232.7553	1232.7693	4.41×10^{-3}	9.62×10^{-3}	1.43×10^{-2}
0.5	0.5	1233.5242	1233.5234	1233.5304	8.23×10^{-4}	6.21×10^{-3}	7.45×10^{-3}
0.25	0.1	2773.7095	2773.6996	2773.7263	9.91×10^{-3}	1.68×10^{-2}	2.67×10^{-2}
0.25	0.5	2775.4295	2775.4226	2775.4338	6.96×10^{-3}	4.32×10^{-3}	1.12×10^{-2}

Table 1: Comparison of the optimal solution, the lower bound and the upper bound for quadratic loss and linear payoff with $V_0 = Coef \times H_0^{linear}$, $\alpha = 0$, $\beta = 1$ and $T = 0.25$ (for $H = 0.1$: $n = 20$; for $H = 0.5$: $n = 1$, $w_1 = 1$ and $x_1 = 0$).

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 [36, 37]. 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

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

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 payoffs, 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 (47). Moreover, as the characteristic function of the log-price in the approximate rough Heston model is available, the lower bound 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 [34], 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 D (see (55)) and has a similar form to $\xi_t^{n,\mathcal{U}}$ except that it does not depend on the volatility process. Figure

4 presents the lower bound $\underline{l}_n^{\mathcal{U}}$ as a function of the number of factors n, for different values of H. As in the linear payoff case, we notice a convergence as the number of factors increases.

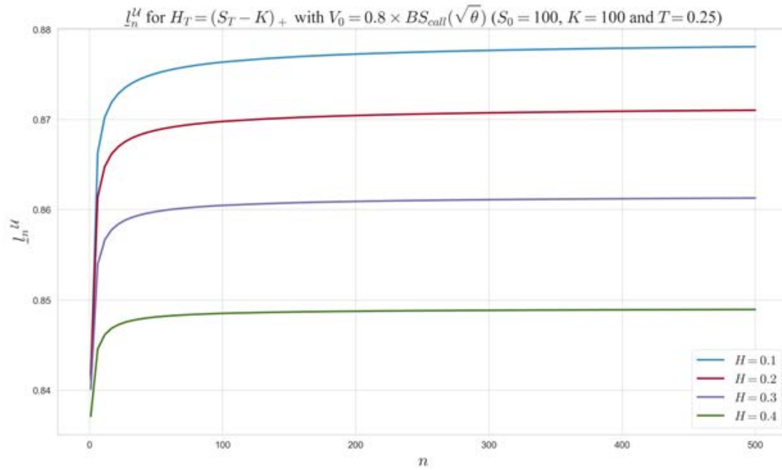

Figure 4: Convergence analysis for $\underline{l}_n^{\mathcal{U}}$ for quadratic loss and ATM call with $V_0 = 0.8 \times BS_{call}(\sigma = \sqrt{\theta})$ and $T = 0.25$.

Figure 5 compares the histogram of the simulated P&L generated by the sub-optimal strategy ($\xi_t^{n,\mathcal{U}}$) and the benchmark strategy (ξ_t^{BS}, $\sigma = \sqrt{\theta}$) with a daily hedging frequency. We observe that the sub-optimal strategy generates a more centred distribution with a thinner left tail and lower variance than the benchmark strategy. This is also confirmed by Table 3 which shows, for different strikes and hedging frequencies, that the sub-optimal strategy outperforms the benchmark strategy in terms of the variance of the P&L. Note that in terms of P&L expectation, the benchmark slightly outperforms the sub-optimal strategy, but the gap between the two strategies is not really significant, as revealed by Table 2.

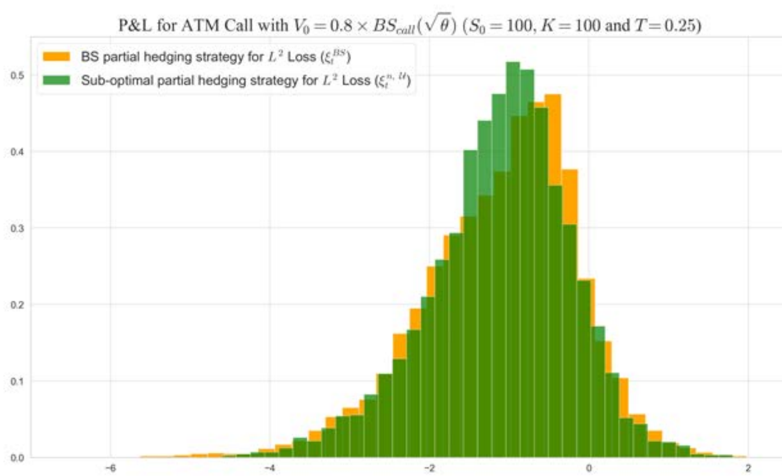

Figure 5: Histogram of the P&Ls associated to the sub-optimal strategy ($\xi_t^{n,\mathcal{U}}$) and the benchmark strategy (ξ_t^{BS}, $\sigma = \sqrt{\theta}$), ATM Call with $V_0 = 0.8 \times BS_{call}(\sigma = \sqrt{\theta})$, $H = 0.1$, $n = 20$, $T = 0.25$ and $\Delta_t = 1/365$.

Strike	$E(P\&L_T)$ with $\xi_t^{n,\mathcal{U}}$	$E(P\&L_T)$ with ξ_t^{BS}
S_0	-1.1376	-1.1284
$0.9 \times S_0$	-2.6539	-2.6413
$1.1 \times S_0$	-0.0801	-0.0757

Table 2: Expectation of the P&Ls associated to the sub-optimal strategy ($\xi_t^{n,\mathcal{U}}$) and the benchmark strategy (ξ_t^{BS}, $\sigma = \sqrt{\theta}$) with $V_0 = 0.8 \times BS_{call}(\sigma = \sqrt{\theta})$, $H = 0.1$, $n = 20$, $T = 0.25$ and $\Delta_t = 1/365$.

Strike	Δ_t	$\mathbb{V}(P\&L_T)$ with $\xi_t^{n,\mathcal{U}}$	$\mathbb{V}(P\&L_T)$ with ξ_t^{BS}	Rel. diff.
S_0	1/365	1.0157	1.1520	13.42%
S_0	1/730	0.8456	0.9901	17.08%
$0.9 \times S_0$	1/365	0.7931	1.0756	35.61%
$0.9 \times S_0$	1/730	0.6421	0.9553	48.78%
$1.1 \times S_0$	1/365	0.9086	1.0588	16.53%
$1.1 \times S_0$	1/730	0.7479	0.9031	20.76%

Table 3: Variance of the P&Ls associated to the sub-optimal strategy ($\xi_t^{n,\mathcal{U}}$) and the benchmark strategy (ξ_t^{BS}, $\sigma = \sqrt{\theta}$) with $V_0 = 0.8 \times BS_{call}(\sigma = \sqrt{\theta})$, $H = 0.1$, $n = 20$, and $T = 0.25$.

Table 4 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 (44) 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 5 that by increasing the frequency of portfolio rebalancing, the gap between the bounds decreases since the variance of the simulated P&L decreases as the hedging frequency increases (see also Table 3).

Benchmarking the sub-optimal strategy against the Black-Scholes strategy in terms of quadratic loss, we observe at Tables 4 and 5 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 performs compared to the sub-optimal strategy.

$Coef$	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	0.0927	0.4659	0.5678	0.3732	0.1019
1	0.5	0.0147	0.2601	0.2943	0.2454	0.0306
0.75	0.1	1.2929	1.5342	1.6358	0.2413	0.1016
0.75	0.5	1.2621	1.4210	1.4454	0.1589	0.0244
0.5	0.1	4.7731	4.9361	5.0357	0.1631	0.0996
0.5	0.5	4.7994	4.9231	4.9436	0.1237	0.0205
0.25	0.1	10.5916	10.6751	10.7555	0.0835	0.0804
0.25	0.5	10.6839	10.7528	10.7870	0.0689	0.0342

Table 4: Comparison of the lower bound, the upper bound and the benchmark ($l_{BS(\sqrt{\theta})}$ with $\sigma = \sqrt{\theta}$) for quadratic loss, ATM Call with $V_0 = Coef \times BS_{Call}(\sigma = \sqrt{\theta})$, $T = 0.25$, $\Delta_t = 1/365$ (for $H = 0.1$: $n = 20$; for $H = 0.5 : n = 1$, $w_1 = 1$ and $x_1 = 0$).

$Coef$	Δ_t	$\underline{l}_n^{\mathcal{U}}$	$\overline{l}_n^{\mathcal{U}}$	$l_{BS(\sqrt{\theta})}$	Abs. diff. $\lvert \overline{l}_n^{\mathcal{U}} - \underline{l}_n^{\mathcal{U}} \rvert$	Abs. diff. $\lvert l_{BS(\sqrt{\theta})} - \overline{l}_n^{\mathcal{U}} \rvert$
1	1/365	0.0927	0.4659	0.5678	0.3732	0.1019
1	1/730	0.0927	0.3773	0.4831	0.2846	0.1058
0.75	1/365	1.2929	1.5342	1.6358	0.2413	0.1016
0.75	1/730	1.2929	1.4562	1.5549	0.1633	0.0987
0.5	1/365	4.7731	4.9361	5.0357	0.1631	0.0996
0.5	1/730	4.7731	4.8513	4.9598	0.0712	0.1085
0.25	1/365	10.5916	10.6751	10.7555	0.0835	0.0804
0.25	1/730	10.5916	10.6333	10.7143	0.0418	0.0813

Table 5: Comparison of the lower bound, the upper bound and the benchmark ($l_{BS(\sqrt{\theta})}$ with $\sigma = \sqrt{\theta}$) 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 $T = 0.25$.

Finally, it is also interesting to study the impact of the correlation ρ on the hedging strategy. Figure 6 presents the impact of the correlation ρ on the bounds. We observe that the loss is reduced 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 $\xi_t^{n,\mathcal{U}}$ and the Black-Scholes hedging ratio ξ_t^{BS}.

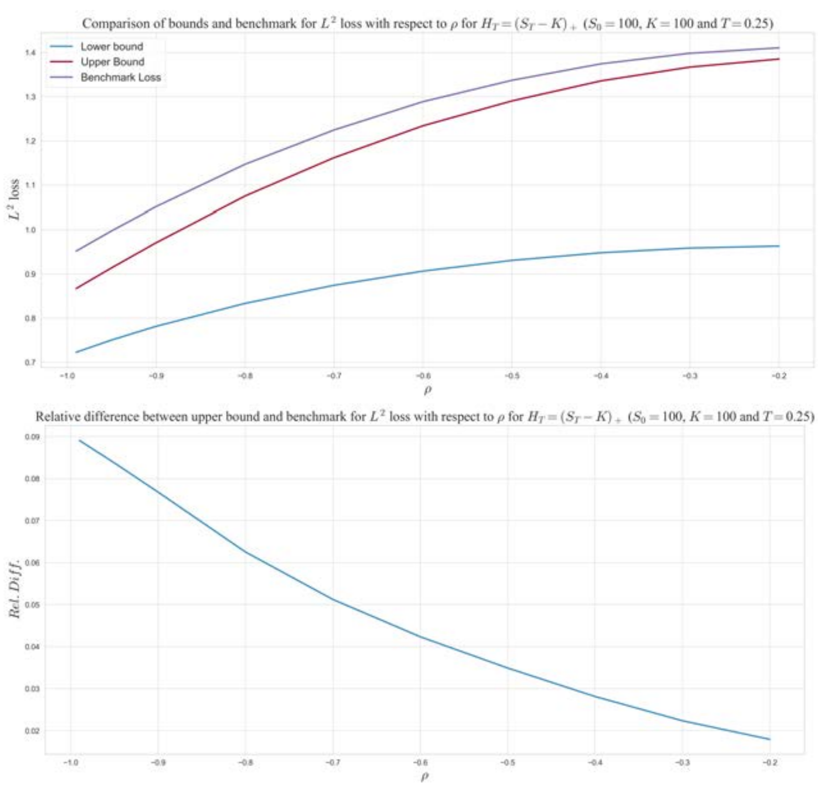

Figure 6: Impact of the correlation ρ for ATM Call with $H = 0.1$, $n = 20$, $V_0 = 0.8 \times BS_{call}(\sqrt{\theta})$, $\Delta_t = 1/365$ and $T = 0.25$.

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 [4], 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 [39], 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.

Appendix

A. Additional proofs

Proof of Proposition 6

Proof. First, based on their SDE and provided that $(S_t^n)_{0\le t\le T}$ and $(S_t)_{0\le t\le T}$ have the same initial value $S_0 = S_0^n = s_0$, it can be shown that the processes are solutions of

$$S_t = s_0 \times \exp\left(\int_0^t (\mu_s - \frac{1}{2}\nu_s)ds + \int_0^t \sqrt{\nu_s}dW_S(s)\right)$$
$$S_t^n = s_0 \times \exp\left(\int_0^t (\hat{\mu}_s - \frac{1}{2}\hat{\nu}_s)ds + \int_0^t \sqrt{\hat{\nu}_s}dW_S(s)\right),$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$. By Theorem (5) and uniform integrability of $(\hat{\nu}_t)_{n\ge 1}$, for $t \in [0,T]$,

$$E(\hat{\nu}_t) \to E(\nu_t),$$

as n goes to infinity. By applying Fubini's theorem, we obtain that

$$\lim_{n\to\infty} E\bigg(\int_0^t \hat{\nu}_s ds\bigg) = E\bigg(\int_0^t \nu_s ds\bigg).$$

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

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

But since L^2 convergence implies convergence in law, we obtain that

$$\int_0^t \sqrt{\hat{\nu}_s}dW_S(s) \xrightarrow{\mathcal{L}} \int_0^t \sqrt{\nu_s}dW_S(s).$$

Moroever, using Theorem (5), we deduce that

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

Finally as the exponential function is continuous, we conclude that

$$\exp\left(\int_0^t (\hat{\mu}_s - \frac{1}{2}\hat{\nu}_s)ds + \int_0^t \sqrt{\hat{\nu}_s}dW_S(s)\right) \xrightarrow{\mathcal{L}} \exp\left(\int_0^t (\mu_s - \frac{1}{2}\nu_s)ds + \int_0^t \sqrt{\nu_s}dW_S(s)\right)$$

and then

$$S_t^n \xrightarrow{\mathcal{L}} S_t,$$

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

$$V_t = v_0 + \int_0^t \xi_s\mu_s S_s ds + \int_0^t \xi_s\sqrt{\nu_s}S_s dW_S(s)$$
$$V_t^n = v_0 + \int_0^t \xi_s\hat{\mu}_s S_s^n ds + \int_0^t \xi_s\sqrt{\hat{\nu}_s}S_s^n dW_S(s),$$

on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$. By Theorem (5), and Fubini's theorem, for $t \in [0,T]$,

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

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

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

But since L^2 convergence implies convergence in law, we obtain that

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

Finally, as

$$\int_0^t \xi_s\hat{\mu}_s S_s^n ds \xrightarrow{\mathcal{L}} \int_0^t \xi_s\mu_s S_s ds,$$

we conclude that

$$V_t^n \xrightarrow{\mathcal{L}} V_t,$$

as n goes to infinity. □

Proof of Proposition 11

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

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

and

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

Using now the PDE (21) 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 and that concludes the proof. □

Proof of Lemma 14

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

$$\hat{L}(h(s),z) = \inf_v\{\frac{1}{p}(h(s)-v)^p - zv\}. \tag{48}$$

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 (48), 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 15

Proof. From (11), we know that $\hat{l}_n$ satisfies the PDE (22). Let consider the ansatz

$$\hat{l}_n(t,s,\boldsymbol{\nu},z) = \underbrace{-\frac{z^q}{q}\exp\left(C_t + \sum_{i=1}^{n}D_t^i\nu^i\right)}_{:=\hat{l}_n^1} - \Big(\underbrace{e^{-r(T-t)}\alpha + \beta s}_{:=g(t,s)}\Big)\,z, \tag{49}$$

with

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

such that

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

Plugging (49) in the PDE (22), 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

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

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

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

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

$$\partial_t C_t = rq - \frac{1}{2} q(q-1)\, A^2 \nu_0 - \sum_{i=1}^n D_t^i \Big(\kappa(\theta - \nu_0) - q\rho A \nu_0 \zeta \Big) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n D_t^i D_t^j \zeta^2 \nu_0 \Big(1 - (1-\rho^2) \frac{q}{q-1} \Big), \tag{50}$$

$$C_T = 0$$

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

$$\begin{aligned}\partial_t D_t^i =& x_i D_t^i + w_i \sum_{j=1}^n D_t^j \Big(\kappa + q\rho A\zeta\Big) - \frac{1}{2} w_i \Big(\zeta^2 - (1-\rho^2)\zeta^2 \frac{q}{q-1}\Big) \sum_{j=1}^n \sum_{k=1}^n D_t^j D_t^k \\ & - \frac{1}{2} w_i q(q-1)\, A^2, \\ D_T^i =& 0.\end{aligned} \tag{51}$$

Moreover, by (26), we know that the value of the dual process at time $t \in [0, T]$ can be expressed in function of $(t, s, \boldsymbol{\nu}, v)$ such that $z = z(t, s, \boldsymbol{\nu}, v)$ with $z(t, s, \boldsymbol{\nu}, v)$ satisfying

$$\partial_z \hat{l}_n + v = 0.$$

We easily deduce that

$$v = g(t, s) + z^{q-1} \exp(C_t + \sum_{i=1}^n D_t^i \nu^i),$$

and

$$z(t, s, \boldsymbol{\nu}, v) = \exp\Big(-\frac{1}{q-1}(C_t + \sum_{i=1}^n D_t^i \nu^i)\Big) \times \Big(v - g(t, s)\Big)^{\frac{1}{q-1}}. \tag{52}$$

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

$$l_n(t, s, \boldsymbol{\nu}, v) = -\frac{z(t, s, \boldsymbol{\nu}, v)^q}{q} \exp\Big(C_t + \sum_{i=1}^n D_t^i \nu^i\Big) - z(t, s, \boldsymbol{\nu}, v)\Big(e^{-r(T-t)}\alpha + \beta s - v\Big),$$

with $z(t, s, \boldsymbol{\nu}, v)$ satisfying (52). 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

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

and

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

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

$$\begin{aligned}\xi_t^{n,*} =& \frac{1}{\hat{\nu}_t \Big(S_t^n\Big)^2} \Bigg(-(q-1)\, z(t, S_t^n, \boldsymbol{\nu}_t, V_t^n)^{q-1} \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i)\, (\hat{\mu}_t - r) S_t^n + \beta\, \hat{\nu}_t \Big(S_t^n\Big)^2 \\ & + \rho\zeta S_t^n \hat{\nu}_t \sum_{i=1}^n \Big(z(t, S_t^n, \boldsymbol{\nu}_t, V_t^n)^{q-1} D_t^i \exp(C_t + \sum_{i=1}^n D_t^i \nu_t^i)\Big)\Bigg).\end{aligned} \tag{53}$$

Finally, using the fact that $Z_t = z(t, S_t^n, \boldsymbol{\nu}_t, V_t^n)$ with $z(t, s, \boldsymbol{\nu}, v)$ satisfying (52), 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. □

B. Feller condition in approximate rough Heston

Proposition. *Assume that the approximate volatility process $(\hat{\nu}_t)_{0\leq t\leq T}$ is modeled by a approximate rough Heston model with $b(x)=\kappa(\theta-x)$, $\sigma(x)=\zeta\sqrt{x}$ and suppose that $(w_i)_{i=1,\dots,n}$ and $(x_i)_{i=1,\dots,n}$ satisfy Assumption 2. If $2\kappa\theta>\zeta^2\sum_{i=1}^n w_i$ and $\nu_0>0$ such that for all $t\in[0,T]$, $i\in\{1,...,n\}$ and $\varepsilon>0$,*

$$E\left(1_{\{t\leq\tau_\varepsilon\}}w_i\nu_t^i\hat{\nu}_t^{-(m+1)}\right)\leq E\left(1_{\{t\leq\tau_\varepsilon\}}\hat{\nu}_t^{-m}\right),$$

with $\tau_\varepsilon:=\min\{t\geq 0:\hat{\nu}_t\leq\varepsilon\}$ and $m:=\frac{2\kappa\theta-\zeta^2\sum_{i=1}^n w_i}{\zeta^2\sum_{i=1}^n w_i}$. Then $\forall t\in[0,T]$,

$$\hat{\nu}_t>0,$$

almost surely.

Proof. Let $\varepsilon>0$ and, for $t\in[0,T]$, denote $\tau_\varepsilon:=\min\{t\geq 0:\hat{\nu}_t\leq\varepsilon\}$ and $\tau_\varepsilon\wedge t:=\min\{\tau_\varepsilon,t\}$. Define a strict positive constant m such that

$$m:=\frac{2\kappa\theta-\zeta^2\sum_{i=1}^n w_i}{\zeta^2\sum_{i=1}^n w_i}>0$$

and apply the Ito's lemma to the function $f(x)=x^{-m}$. In this case as

$$\hat{\nu}_t=\nu_0+\sum_{i=1}^n w_i\nu_t^i,$$

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

$$d\nu_t^i=\Big(-x_i\nu_t^i+\kappa(\theta-\hat{\nu}_t)\Big)dt+\zeta\sqrt{\hat{\nu}_t}dW_\nu(t),$$

we deduce that

$$\begin{aligned}
\hat{\nu}_{\tau_\varepsilon\wedge t}^{-m}=&\nu_0^{-m}-\int_0^{\tau_\varepsilon\wedge t}m\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+1)}\sum_{i=1}^n w_i\Big(-x_i\nu_{\tau_\varepsilon\wedge s}^i+\kappa(\theta-\hat{\nu}_{\tau_\varepsilon\wedge s})\Big)ds\\
&+\frac{1}{2}\int_0^{\tau_\varepsilon\wedge t}m(m+1)\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+2)}\zeta^2\hat{\nu}_{\tau_\varepsilon\wedge s}\left(\sum_{i=1}^n w_i\right)^2ds\\
&-\int_0^{\tau_\varepsilon\wedge t}m\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+1/2)}\zeta\sum_{i=1}^n w_i dW_\nu(s)\\
=&\nu_0^{-m}+m\kappa\sum_{i=1}^n w_i\int_0^{\tau_\varepsilon\wedge t}\hat{\nu}_{\tau_\varepsilon\wedge s}^{-m}ds\\
&-m\zeta\sum_{i=1}^n w_i\int_0^{\tau_\varepsilon\wedge t}\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+1/2)}dW_\nu(s)\\
&+\int_0^{\tau_\varepsilon\wedge t}\left(\frac{m(m+1)}{2}\zeta^2\sum_{i=1}^n w_i-m\kappa\theta\right)\sum_{i=1}^n w_i\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+1)}ds\\
&+\int_0^{\tau_\varepsilon\wedge t}m\sum_{i=1}^n w_i\Big(x_i\nu_{\tau_\varepsilon\wedge s}^i\Big)\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+1)}ds.
\end{aligned}$$

Using the definition of m, we have that

$$\begin{aligned}
\hat{\nu}_{\tau_\varepsilon\wedge t}^{-m}=&\nu_0^{-m}+m\kappa\sum_{i=1}^n w_i\int_0^{\tau_\varepsilon\wedge t}\hat{\nu}_{\tau_\varepsilon\wedge s}^{-m}ds\\
&-m\zeta\sum_{i=1}^n w_i\int_0^{\tau_\varepsilon\wedge t}\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+1/2)}dW_\nu(s)\\
&+\int_0^{\tau_\varepsilon\wedge t}m\sum_{i=1}^n w_i\Big(x_i\nu_{\tau_\varepsilon\wedge s}^i\Big)\hat{\nu}_{\tau_\varepsilon\wedge s}^{-(m+1)}ds.
\end{aligned}$$

Taking now the expectation, we obtain

$$\begin{aligned}
E\Big(\hat\nu_{\tau_\varepsilon\wedge t}^{-m}\Big) \le & \nu_0^{-m} + m\kappa\sum_{i=1}^n w_i\int_0^t E\Big(\hat\nu_{\tau_\varepsilon\wedge s}^{-m}\Big)ds \\
& + \int_0^t m\sum_{i=1}^n x_i E\Big(w_i 1_{\{s\le\tau_\varepsilon\}}\nu_s^i\hat\nu_s^{-(m+1)}\Big)ds \\
\le & \nu_0^{-m} + m\kappa\sum_{i=1}^n w_i\int_0^t E\Big(\hat\nu_{\tau_\varepsilon\wedge s}^{-m}\Big)ds \\
& + m\sum_{i=1}^n x_i\int_0^t E\Big(1_{\{s\le\tau_\varepsilon\}}\hat\nu_s^{-m}\Big)ds \\
\le & \nu_0^{-m} + m\Big(\kappa\sum_{i=1}^n w_i + \sum_{i=1}^n x_i\Big)\int_0^t E\Big(\hat\nu_{\tau_\varepsilon\wedge s}^{-m}\Big)ds,
\end{aligned}$$

Using the Grönwall's inequality, we deduce that

$$E\Big(\hat\nu_{\tau_\varepsilon\wedge t}^{-m}\Big) \le \nu_0^{-m}\exp\Big(m\Big(\kappa\sum_{i=1}^n w_i + \sum_{i=1}^n x_i\Big)\,t\Big).$$

Finally, using the Markov's inequality, we obtain that

$$\begin{aligned}
P(\tau_\varepsilon\le t) &= P(\hat\nu_{\tau_\varepsilon\wedge t}^{m}\le\varepsilon^m) \\
&= P(\hat\nu_{\tau_\varepsilon\wedge t}^{-m}\ge\varepsilon^{-m}) \\
&\le \varepsilon^m\nu_0^{-m}\exp\Big(m\Big(\kappa\sum_{i=1}^n w_i + \sum_{i=1}^n x_i\Big)\,t\Big),
\end{aligned}$$

with $m>0$. Therefore we conclude that for $t\in[0,T]$,

$$\lim_{\varepsilon\to 0}P(\tau_\varepsilon\le t) = 0$$

and $\forall t\in[0,T]$,

$$\hat\nu_t > 0,$$

almost surely. □

C. 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*

$$\begin{aligned}
\nu_t^i &= -\int_0^t x_i\nu_s^i ds + \int_0^t \kappa(\theta-\hat\nu_s)ds + \int_0^t \zeta\sqrt{\hat\nu_s}dW_\nu(s),\ i=1,...,n, \\
\nu_0^i &= 0,
\end{aligned} \tag{54}$$

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 $k = 1, ..., n$,

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

Proof. The proof is a direct application of Feynman-Kac theorem to $E_{t,s,\boldsymbol{\nu}}\left(\exp(ix\ \log(S_T^n))\right)$ with SDEs (10) and (54). □

D. Partial hedging in Black and Scholes framework

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*

$$\xi_t^{BS} := \frac{1}{p-1}\left(E^{\mathbb{Q}}(H_T|\mathcal{F}_t) - V_t\right)\frac{(\mu - r)}{\sigma S_t} + \partial_S E^{\mathbb{Q}}(H_T|\mathcal{F}_t), \quad (55)$$

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. In Black and Scholes framework, the dual solution satisfies the PDE (22) that is linear (the derivatives with respect to the volatility process disappear). Then, for power loss functions, we can easily deduce the explicit form of the dual solution and, using the link between the optimal hedging ratio and the dual solution, we derive the form the optimal hedging ratio (55). □

Acknowledgements

We thank the anonymous referees for their remarks and comments, which clearly helped to improve the quality of the paper. Edouard Motte was financially supported by the Fonds de la Recherche Scientifique (F.R.S. - FNRS) through a FRIA grant.

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. and El Euch, O., 2019. Markovian structure of the Volterra Heston model. Statistics & Probability Letters, 149, pp.63-72.

[4] 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.

[5] Abi Jaber, E. and Illand, C., 2022. Joint SPX-VIX calibration with Gaussian polynomial volatility models: deep pricing with quantization hints. arXiv preprint arXiv:2212.08297.

[6] Abi Jaber, E. and De Carvalho, N., 2023. Reconciling rough volatility with jumps. arXiv preprint arXiv:2303.07222.

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

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

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

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

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

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

[13] Bouchard, B. and Chassagneux, J.F., 2016. Fundamentals and advanced techniques in derivatives hedging. Springer.

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

[15] 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.

[16] 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.

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

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

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

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

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

[22] 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.

[23] 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.

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

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

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

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

[28] Guyon, J. and El Amrani, M., 2022. Does the Term-Structure of Equity At-the-Money Skew Really Follow a Power Law?. Available at SSRN 4174538.

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

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

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

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

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

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

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

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] Rømer, S.E., 2022. Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets. Quantitative Finance, 22(10), pp.1805-1838.

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