Optimal Hedging of American Options in Discrete Time

Bruno Rémillard, Alexandre Hocquard, Hugues Langlois, and Nicolas Papageorgiou

Abstract In this article we study the price of an American style option based on hedging the underlying assets in discrete time. Like its European style analog, the value of the option is not given in general by an expectation with respect to an equivalent martingale measure. We provide the optimal solution that minimizes the hedging error variance. When the assets dynamics are Markovian or a component of a Markov process, the solution can be approximated easily by numerical methods already proposed for pricing American options. We proceed to a Monte Carlo experiment in which the hedging performance of the solution is evaluated. For assets returns that are either Gaussian or Variance Gamma, it is shown that the proposed solution results in lower root mean square hedging error than with traditional delta hedging.

Keywords Hedging • American option • Bermudan option • Risk • Martingale

MSC: 91G60, 91G20.

B. Rémillard (✉)
Department of Management Sciences and GERAD, HEC Montréal, 3000 chemin de la Côte Sainte-Catherine, Montréal (Québec), Canada H3T 2A7
e-mail: bruno.remillard@hec.ca

A. Hocquard
Brockhouse Cooper Asset Management, 1250 René-Lévesque Blvd. West, Montréal (Québec), Canada H3B 4W8
e-mail: ahocquard@pavilioncorp.com

H. Langlois
Desautels Faculty of Management, McGill University and Desjardins Asset Management, 1001 Sherbrooke Street West, Montréal (Québec), Canada H3A 1G5
e-mail: hugues.langlois@desjga.com

N. Papageorgiou
Department of Finance, HEC Montréal, 3000 chemin de la Côte Sainte-Catherine, Montréal (Québec), Canada H3T 2A7
e-mail: nicolas.papageorgiou@hec.ca

R.A. Carmona et al. (eds.), *Numerical Methods in Finance*, Springer Proceedings in Mathematics 12, DOI 10.1007/978-3-642-25746-9_5,

1 Introduction

The lack of realism of the assumptions underlying the Black-Scholes formula for pricing financial derivatives has engendered a large area of research attempting to address these shortcomings. Most of this literature has focused on the inadequacy of the normal distribution for modeling assets returns and the consequent poor performance of Black-Scholes option prices in matching observed market prices. Several richer structures for return dynamics have been proposed; GARCH volatility [14], low and high frequency volatility components [10, 16], stochastic volatility [9,18,19,21], jumps in returns [24], jumps in volatility [1,25], and regimes switching models [28, 29].

However, a crucial issue for the option writer is the well documented fact that replicating strategies derived from continuous time processes, but applied in discrete time, lead to non-zero expected hedging error, see Wilmott [32] for a detailed discussion. Indeed, an expected hedging error centered at zero requires in this case continuous trading of the hedging position in the underlying asset, and is thus impossible in reality.

In incomplete markets, there are an infinity of risk neutral probability measures and one has to be chosen in order to determine the derivative's price. It is illuminating to consider this choice from the perspective of an option writer attempting to hedge his market exposure. His utility is derived from the error realized by dynamically replicating the derivative's payoff. In other words, the equivalent martingale measure implied by observed option prices reflects the option writers' preference over possible hedging error outcomes. Few papers however have adopted explicitly this perspective. Pochart and Bouchaud [27] propose a non-parametric approach in which the value and hedging ratio functions are approximated with basis functions. This flexible method allows to choose from a wide variety of loss functions (for example, minimizing the expected hedging error shortfall or the hedging error value-at-risk).

An asymmetric loss function is intuitively appealing since dealers should care more about losses than if their replicating portfolio results in over-hedging. However, the optimal hedging strategies are hard to obtain in these cases and only numerical approximations have been proposed so far. The minimization of the variance of hedging errors significantly facilitates analytical treatment. While a quadratic loss function has also been considered in Bouchaud and Potters [4], Cornalba et al. [11], and Bouchaud [3], Schweizer [30] derives the optimal solution when the underlying source of risk is one-dimensional, and Rémillard and Rubenthaler [29] extend this result to the multi-dimensional case. This generalization is particularly useful since derivatives seldom depend on only one source of risk; the payoff function may depend on multiple assets, and the value is sensitive to time-varying volatility and interest rates.

We contribute by considering the special case when it is possible for the option holder to exercise early. Valuing an American option (or more precisely a Bermudan option since our analysis is cast in discrete time) is complicated by the need to determine the optimal exercise rule. The conventional no-arbitrage pricing approach for American style contingent claims involves taking the supremum of the expectation of the discounted payoff under a risk neutral measure over a set $\mathcal{T}$ of admissible stopping time

$$\sup_{\tau \in \mathcal{T}} E\left[e^{-r\tau} f\left(S_\tau\right)\right]$$

where τ is a stopping time, r the risk free rate, S the underlying asset and $f(\cdot)$ the payoff function. For more details, see Duffie [15] for an excellent textbook treatment of the theory of American contingent claim valuation.

One important aspect of our method is that the choice of a risk neutral measure is bypassed, the variance of hedging error is minimized directly under the objective measure. It therefore can be ascribed to the literature emphasizing the significance of agents' preferences on derivatives pricing. For example, Bates [2] underlines the importance of risks intermediation by option market makers on contingent claim valuation, and Garleanu et al. [17] provide an interesting empirical investigation of the effect of demand on derivatives prices.

Our approach is general enough to accommodate for any well-behaved payoff functions and most commonly used price processes, as long as it has finite variance. When the price is Markovian, or the component of a Markov process, the analytical solution can be implemented using numerical techniques found in the literature. One only needs an efficient method of estimating conditional expectations, for instance, the regression based techniques of Carriere [7], Tsitsiklis and Roy [31], Longstaff and Schwartz [22] or the stochastic mesh approach of Broadie and Glasserman [5]. For the numerical examples presented, we use the deterministic grid method of Papageorgiou et al. [26].

As an illustration, we apply the optimal solution to the case where assets returns follow general Lévy processes. These processes provide a rich modeling framework and have been used extensively over the past decade in the asset pricing literature, see among others Huang and Wu [20], Chan [8], and Carr and Wu [6].

Furthermore, we show the potential reduction in hedging error variance with Monte Carlo experiments in which a put option is hedged dynamically. First, we simulate paths for a geometric Brownian motion with constant volatility. In this case, the market is complete and the hedging performance can be directly compared to Black-Scholes delta hedging. It is found that the root mean square hedging error is strictly lower than delta hedging even as the number of hedging steps increases. Using $n = 250$, that is, hedging once a day for a put with 1 year to maturity, we obtain a RMSE of 0.1081 compared to 0.1151 for delta hedging. The second experiment uses the Variance Gamma model of Madan et al. [23]. Lower RMSE are also found in this incomplete market setup.

The next section presents the optimal solution to the variance minimization problem for a given stopping time, while Sect. 3 examines the choice of stopping time for an American option. Section 4 discusses in details the application of the optimal solution to general Lévy processes, and provides numerical examples of the hedging performance for Gaussian and Variance Gamma returns.

2 Optimal Hedging of American Options

Let $\mathbf{S}_k$ be a d-dimensional vector representing the value of the underlying assets at period k, and $\mathbb{F} = \{\mathcal{F}_k, k = 0, 1, \ldots, n\}$ be a filtration under which $\mathbf{S}$ is adapted with $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Set $\mathbf{\Delta}_k = \beta_k \mathbf{S}_k - \beta_{k-1}\mathbf{S}_{k-1}$, $k = 1, \ldots, n$ where the discount factors β_k are predictable, i.e. β_k is $\mathcal{F}_{k-1}$-measurable $\forall k \in \{1, \ldots, n\}$. Assume further that $\mathbf{S}$ is square integrable. Note that since β is a non-negative process bounded by 1, it is automatically square integrable.

Given a stopping time τ, the goal is to find an initial investment amount π_0 and a predictable investment strategy $\overrightarrow{\boldsymbol{\phi}} = (\phi_k)_{k=1}^n$ that minimize the expected quadratic hedging error

$$E\left[\left\{H_\tau\left(\pi_0, \overrightarrow{\boldsymbol{\phi}}\right)\right\}^2\right],$$

where

$$H_k = H_k\left(\pi_0, \overrightarrow{\boldsymbol{\phi}}\right) = \beta_k f_k - \pi_k,$$

$$\pi_k = \pi_0 + \sum_{j=1}^{k} \boldsymbol{\phi}_j^\top \mathbf{\Delta}_j, \quad k = 0, \ldots, n,$$

and f_k is a short notation for the payoff function evaluated at period k, $f(\mathbf{S}_k)$. We assume further that $f_k < \infty$ for all $k \in \{0, 1, \ldots, n\}$.

The proofs of all of the following results are given in Sects. 6, and 7.

Lemma 2.1. *Conditions* (1)–(2) *below are necessary and sufficient conditions for minimizing* $E\left\{H_\tau^2\left(v_0, \overrightarrow{\boldsymbol{\psi}}\right)\right\}$.

$$\mathbb{I}_{\{\tau \geq j\}} E\left\{(\beta_\tau f_\tau - \pi_\tau)\mathbf{\Delta}_j | \mathcal{F}_{j-1}\right\} = 0, \qquad j = 1, \ldots, n\,. \tag{1}$$

$$E\{(\beta_\tau f_\tau - \pi_\tau)\} = 0. \tag{2}$$

Before solving (1)–(2), one needs to introduce the following definitions. Set $\mathcal{P}_{n+1} = 1$, and for $k = n, \ldots, 1$, define

$$\mathcal{A}_k = E\left(\mathbf{\Delta}_k \mathbf{\Delta}_k^\top \mathcal{P}_{k+1} | \mathcal{F}_{k-1}\right),$$

$$\mathfrak{b}_k = \mathcal{A}_k^{-1} E\left(\mathbf{\Delta}_k \mathcal{P}_{k+1} | \mathcal{F}_{k-1}\right),$$

$$\mathfrak{a}_k = \mathbb{I}_{\{\tau \geq k\}} \mathcal{A}_k^{-1} E\left(\beta_\tau f_\tau \mathbf{\Delta}_k \mathcal{P}_{k+1} | \mathcal{F}_{k-1}\right),$$

$$\mathcal{P}_k = \prod_{j=k}^{\tau}\left(1 - \mathfrak{b}_j^\top \mathbf{\Delta}_j\right).$$

Further set $\mathfrak{c}_k = E(\mathcal{P}_{k+1}|\mathcal{F}_k)$, $k = 0, \ldots, n$. Under the assumption that $\mathcal{A}_k$ is invertible for all $k = 1, \ldots, n$, one can prove, as in Rémillard and Rubenthaler [29], that $\mathfrak{c}_k \in (0, 1]$.

Proposition 2.1. *Conditions* (1)–(2) *are met if and only if, for any* $j = 1, \ldots, n$

$$\mathbb{I}_{\{\tau \geq j\}} \boldsymbol{\phi}_j = \mathbb{I}_{\{\tau \geq j\}} \left\{ \mathfrak{a}_j - \pi_{j-1} \mathfrak{b}_j \right\}, \tag{3}$$

and π_0 *satisfies*

$$\pi_0 = I(\tau = 0) f_0 + \mathbb{I}(\tau \geq 1) E(\beta_\tau f_\tau \mathcal{P}_1)/E(\mathcal{P}_1). \tag{4}$$

Furthermore, on $\{\tau \geq j\}$,

$$E\left\{(\beta_\tau f_\tau - \pi_\tau)|\mathcal{F}_{j-1}\right\} = E\left\{(\beta_\tau f_\tau - \pi_j)\mathcal{P}_{j+1}|\mathcal{F}_{j-1}\right\}, \tag{5}$$

for all $j = 1, \ldots, n$.

Equation 3 provides the optimal hedging strategy. Finally, one can also define the "value" process of the option viz.

$$\beta_k \mathcal{C}_k = E(\beta_\tau f_\tau \mathcal{P}_{k+1}|\mathcal{F}_k)/\mathfrak{c}_k, \quad k = 0, \ldots, n.$$

Note that $\mathcal{C}_k$ can be interpreted as the amount of money to invest at the risk-free rate at time k to minimize the hedging error associated with the exercising strategy τ. In particular, the contingent claim's initial value is given by $\pi_0 = \mathcal{C}_0$.

The following relations will be needed.

Proposition 2.2. *For all* $k = 0, \ldots, n-1$,

$$\mathbb{I}(\tau \geq k)\beta_k \mathcal{C}_k = \mathbb{I}(\tau = k)\beta_k f_k + \mathbb{I}(\tau > k) E(\beta_\tau f_\tau \mathcal{P}_{k+1}|\mathcal{F}_k)/\mathfrak{c}_k,$$

and

$$\mathbb{I}(\tau > k)\mathfrak{c}_k \beta_k \mathcal{C}_k = \mathbb{I}(\tau > k) E\{\beta_{k+1} \mathcal{C}_{k+1} \mathcal{P}_{k+1}|\mathcal{F}_k\}.$$

3 Choosing a Stopping Time Strategy

Having found the optimal hedging strategy for a given stopping time, one could be tempted to minimize $E\left[\left\{H_\tau\left(\pi_0, \overrightarrow{\boldsymbol{\phi}}\right)\right\}^2\right]$ over all stopping times. Unfortunately, one would end up with the trivial solution $\tau \equiv 0$ leading to a zero hedging error.

The heuristic solution we propose is to define an optimal stopping time for the option holder as in the traditional case when working under a risk neutral measure.

More precisely, set $P_{n+1} = 1$, and for $k = n, \ldots, 1$, define

$$A_k = E\left(\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top P_{k+1} | \mathcal{F}_{k-1}\right),$$

$$\mathbf{b}_k = A_k^{-1} E\left(\boldsymbol{\Delta}_k P_{k+1} | \mathcal{F}_{k-1}\right),$$

$$P_k = \prod_{j=k}^{n} \left(1 - \mathbf{b}_j^\top \boldsymbol{\Delta}_j\right),$$

$$c_k = E(P_{k+1} | \mathcal{F}_k),$$

where the notation P, A, $\mathbf{b}$, and c is used to contrast with the stopping time case discussed previously. As in Rémillard and Rubenthaler [29], the "value" $\mathcal{V}_k$ at period k of a European style option with payoff $\mathcal{V}_n$ at maturity n is given by

$$\beta_k \mathcal{V}_k = \frac{E(\beta_n \mathcal{V}_n P_{k+1} | \mathcal{F}_k)}{c_k}, \qquad k = 0, \ldots, n.$$

$\mathcal{V}_k$ corresponds to the initial investment at the risk-free rate when forming the portfolio which minimizes the expected quadratic hedging error at maturity.

By analogy with the solution of the Snell problem, define $Z_n = f_n$ and

$$\beta_k Z_k = \max\left\{\beta_k f_k, \frac{E(\beta_{k+1} Z_{k+1} P_{k+1} | \mathcal{F}_k)}{c_k}\right\},$$

for $k = n-1, \ldots, 0$. If $\beta_k \mathbf{S}_k$ is a martingale, then $E(\boldsymbol{\Delta}_k | \mathcal{F}_{k-1}) = 0$ and one can easily check that $P \equiv 1$ and $\mathbf{b}_k \equiv 0$. Hence

$$\beta_k Z_k = \max\left\{\beta_k f_k, E(\beta_{k+1} Z_{k+1} | \mathcal{F}_k)\right\},$$

which is the usual formula for the value of an American option. Note that in this case, we also have $\mathcal{P}_k \equiv 1$ and $\mathfrak{b}_k \equiv 0$.

Using these insights, we choose the following stopping strategy:

$$\tau^* = \min\{j \geq 0; Z_j = f_j\}.$$

Clearly $\tau^* \leq n$ since $Z_n = f_n$ by construction, and τ^* is a stopping time.

Note that $\{\tau^* \geq k\} \cap \{\tau^* \leq k\} = \{Z_k = f_k\} \cap \{\tau^* \geq k\}$ while $\{\tau^* > k\} = \{Z_k > f_k\} \cap \{\tau^* \geq k\}$.

3.1 Implementation of the Stopping Time Strategy

To be able to solve the problem numerically, one assumes that there exists a process $\mathbf{h}$ so that $\mathbf{Y}_k = (\mathbf{S}_k, \mathbf{h}_k)$ is Markov. We assume further that $f_k = f_k(\mathbf{S}_k)$ for all $0 \leq k \leq n$. For simplicity, suppose that $\beta_k = (1+r)^{-k}$. It is then easy to check that $Z_k = g_k(\mathbf{Y}_k)$, with $g_n = f_n$, and

$$g_{k-1}(\mathbf{y}) = \max\left[f_{k-1}(\mathbf{y}), \frac{E\left[g_k(\mathbf{Y}_k)\{1-\mathbf{b}_k(\mathbf{y})^\top \boldsymbol{\Delta}_k\}c_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right]}{(1+r)c_{k-1}(\mathbf{y})}\right], \quad (6)$$

where $c_n \equiv 1$, and for $k = n, \ldots, 1$,

$$A_k(\mathbf{y}) = E\left\{\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top c_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right\},$$

$$\mathbf{b}_k(\mathbf{y}) = A_k^{-1}(\mathbf{y}) E\left\{\boldsymbol{\Delta}_k c_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right\},$$

$$c_{k-1}(\mathbf{y}) = E\left\{c_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right\} - \mathbf{b}_k(\mathbf{y})^\top A_k(\mathbf{y})\mathbf{b}_k(\mathbf{y}).$$

For more details on these equations, see Rémillard and Rubenthaler [29].

Recall that $\mathfrak{c}_k = E(\mathcal{P}_{k+1}|\mathcal{F}_k)$, for $k = 0, \ldots, n$, and $\mathfrak{c}_n \equiv 1$. It is easy to prove that

$$\mathcal{A}_k = \mathbb{I}(\tau^* < k)\mathbb{A}(0, \mathbf{Y}_{k-1}) + \mathbb{I}(\tau^* \geq k)\mathbb{A}(1, \mathbf{Y}_{k-1}),$$

$$\mathfrak{b}_k = \mathbb{I}(\tau^* < k)\mathbb{B}(0, \mathbf{Y}_{k-1}) + \mathbb{I}(\tau^* \geq k)\mathbb{B}(1, \mathbf{Y}_{k-1}),$$

$$\mathfrak{c}_k = \mathbb{I}(\tau^* \leq k) + \mathbb{I}(\tau^* > k)\gamma_k(\mathbf{Y}_k),$$

where

$$\mathbb{A}_k(0, \mathbf{y}) = E\left(\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top |\mathbf{Y}_{k-1}=\mathbf{y}\right),$$

$$\mathbb{A}_k(1, \mathbf{y}) = E\left\{\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \mathbb{I}_{\{f_k = Z_k\}}|\mathbf{Y}_{k-1}=\mathbf{y}\right\} + E\left\{\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \mathbb{I}_{\{f_k < Z_k\}}\gamma_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right\},$$

$$\mathbb{B}_k(0, \mathbf{y}) = \mathbb{A}_k^{-1}(0, \mathbf{Y}_{k-1}) E\left(\boldsymbol{\Delta}_k|\mathbf{Y}_{k-1}=\mathbf{y}\right),$$

$$\mathbb{B}_k(1, \mathbf{y}) = \mathbb{A}_k^{-1}(1, \mathbf{Y}_{k-1}) E\left\{\boldsymbol{\Delta}_k \mathbb{I}_{\{f_k = Z_k\}}|\mathbf{Y}_{k-1}=\mathbf{y}\right\} + \mathbb{A}_k^{-1}(1, \mathbf{Y}_{k-1}) E\left\{\boldsymbol{\Delta}_k \mathbb{I}_{\{f_k < Z_k\}}\gamma_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right\},$$

$$\gamma_{k-1}(\mathbf{y}) = E\left\{\mathbb{I}_{\{f_k = Z_k\}}|\mathbf{Y}_{k-1}=\mathbf{y}\right\} + E\left\{\mathbb{I}_{\{f_k < Z_k\}}\gamma_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right\} - \mathbb{B}_k(1, \mathbf{y})^\top \mathbb{A}_k(1, \mathbf{y})\mathbb{B}_k(1, \mathbf{y}).$$

One can now find expressions for $\mathcal{C}_k$ and $\mathfrak{a}_k$.

Proposition 3.1. $\mathbb{I}_{\{\tau^* > k\}}\mathcal{C}_k = C_k(\mathbf{Y}_k)$ *and* $\mathfrak{a}_k = \mathbb{I}_{\{\tau^* \geq k\}}\mathbf{a}_k(\mathbf{Y}_{k-1})$*, for some deterministic functions* C_k *and* $\mathbf{a}_k$*, where, for all* $k = n, \ldots, 1$*,*

$$C_{k-1}(\mathbf{y}) = \frac{E\left[\mathbb{I}_{\{f_k = Z_k\}} f_k \left\{1 - \mathbb{B}_k(1, \mathbf{y})^\top \boldsymbol{\Delta}_k\right\}|\mathbf{Y}_{k-1}=\mathbf{y}\right]}{(1+r)\gamma_{k-1}(\mathbf{y})} + \frac{E\left[\mathbb{I}_{\{f_k < Z_k\}} C_k(\mathbf{Y}_k)\left\{1 - \mathbb{B}_k(1, \mathbf{y})^\top \boldsymbol{\Delta}_k\right\}\gamma_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right]}{(1+r)\gamma_{k-1}(\mathbf{y})},$$

and

$$\mathbf{a}_k(\mathbf{y}) = (1+r)^{-k}\mathbb{A}_k^{-1}(1,\mathbf{y})E\left\{\mathbb{I}_{\{f_k=Z_k\}} f_k \boldsymbol{\Delta}_k|\mathbf{Y}_{k-1}=\mathbf{y}\right\}$$
$$+(1+r)^{-k}\mathbb{A}_k^{-1}(1,\mathbf{y})E\left\{\mathbb{I}_{\{f_k<Z_k\}}\boldsymbol{\Delta}_k\gamma_k(\mathbf{Y}_k)C_k(\mathbf{Y}_k)|\mathbf{Y}_{k-1}=\mathbf{y}\right\}.$$

In particular,

$$\mathbf{a}_n(\mathbf{y}) = (1+r)^{-n}\mathbb{A}_n^{-1}(1,\mathbf{y})E\left\{f_n(\mathbf{S}_n)\boldsymbol{\Delta}_n|\mathbf{Y}_{n-1}=\mathbf{y}\right\}.$$

Since all deterministic functions are defined as expectations, they can be approximated using the methodologies proposed in Del Moral et al. [12, 13] and Papageorgiou et al. [26]. The latter methodology will now be illustrated with some examples in the next section.

4 Examples of Application

In this section we show how our methodology can be applied to general Lévy processes, where the returns are independent and identically distributed. Optimal discrete time and continuous time hedging for European type options were tackled in Rémillard and Rubenthaler [29]. See also Rémillard et al. [28] for an implementation in discrete time.

4.1 Lévy Models

For these models, $\Delta_k = \tilde{S}_k - \tilde{S}_{k-1} = \tilde{S}_{k-1}\xi_k$, with $\xi_k = e^{R_k} - 1$, where the (excess) log-returns R_k are independent and identically distributed (iid for short). Set $\mu = E(\xi_k)$ and $B = E\left(\xi_k^2\right)$. ξ_k is the usual excess return at period k and these returns are assumed to be independent copies of ξ with law ν.

For simplicity, assume that $\tilde{f}_k = (1+r)^{-k} f_k(\tilde{S}_k)$ and that the process S is replaced by the discounted process $\tilde{S}$.

Note that c_k is independent of s and that $c_k = \lambda^{n-k}$, where $\lambda = 1 - \frac{\mu^2}{B} = \frac{\sigma^2}{\sigma^2+\mu^2} \in (0,1)$, and $\sigma^2 = \text{var}(\xi)$. Moreover,

$$A_k(s) = s^2 B c_k, \qquad b_k(s) = \frac{\mu}{sB}, \quad k = 1,\dots,n.$$

As a result, setting $b_0 = \frac{\mu}{B}$, one has

$$P_k = \prod_{j=k}^{n}\left(1 - b_0\xi_j\right).$$

Next, setting $\tilde{g}_k(s) = (1+r)^{-k} g_k(s)$, and using (6), one gets

$$\tilde{g}_{k-1}(s) = \max\left[\tilde{f}_{k-1}(s), \int \tilde{g}_k\{s(1+x)\}\frac{(1-b_0x)}{\lambda}\nu(dx)\right], \tag{7}$$

for $k = n, \ldots, 1$, with $\tilde{g}_n = \tilde{f}_n$.

Remark 4.1. Note that under a risk neutral measure, $b_0 = 0$ and $\tilde{g}_k$ given by (7) is the value of the option at period k if $\tau^* \geq k$.

Next, we have $\mathbb{A}_k(1,s) = s^2\tilde{\mathbb{A}}_k(1,s)$, and $\mathbb{B}_k(1,s) = \tilde{\mathbb{B}}_k(1,s)/s$, where

$$\begin{aligned}
\tilde{\mathbb{A}}_k(1,s) &= \int x^2 \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\} = \tilde{g}_k\{s(1+x)\}\}}\nu(dx) \\
&\quad + \int x^2 \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\} < \tilde{g}_k\{s(1+x)\}\}}\gamma_k\{s(1+x)\}\nu(dx), \\
\tilde{\mathbb{B}}_k(1,s) &= \frac{1}{\tilde{\mathbb{A}}_k(1,s)}\int x \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\} = \tilde{g}_k\{s(1+x)\}\}}\nu(dx) \\
&\quad + \frac{1}{\tilde{\mathbb{A}}_k(1,s)}\int x \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\} < \tilde{g}_k\{s(1+x)\}\}}\gamma_k\{s(1+x)\}\nu(dx), \\
\gamma_{k-1}(s) &= \int \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\} = \tilde{g}_k\{s(1+x)\}\}}\nu(dx) \\
&\quad + \int \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\} < \tilde{g}_k\{s(1+x)\}\}}\gamma_k\{s(1+x)\}\nu(dx) \\
&\quad - \tilde{\mathbb{B}}_k(1,s)^2\tilde{\mathbb{A}}_k(1,s).
\end{aligned}$$

Then, we obtain

$$\begin{aligned}
\tilde{C}_{k-1}(s) &= \frac{1}{\gamma_{k-1}(s)}\int \mathbb{I}_{\{f_k\{s(1+x)\} = \tilde{g}_k\{s(1+x)\}\}}\tilde{f}_k\{s(1+x)\} \\
&\quad \times \left\{1 - \tilde{\mathbb{B}}_k(1,s)x\right\}\nu(dx) \\
&\quad + \frac{1}{\gamma_{k-1}(s)}\int \mathbb{I}_{\{f_k\{s(1+x)\} < \tilde{g}_k\{s(1+x)\}\}}\tilde{C}_k\{s(1+x)\} \\
&\quad \times \left\{1 - \tilde{\mathbb{B}}_k(1,s)x\right\}\gamma_k\{s(1+x)\}\nu(dx)
\end{aligned}$$

and $a_k(s) = \tilde{a}_k(s)/s$, where

$$\tilde{a}_k(s) = \frac{1}{\tilde{\mathbb{A}}_k(1,s)} \int \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\}=\tilde{g}_k\{s(1+x)\}\}} \tilde{f}_k\{s(1+x)\} x \nu(dx)$$
$$+ \frac{1}{\tilde{\mathbb{A}}_k(1,s)} \int \mathbb{I}_{\{\tilde{f}_k\{s(1+x)\}<\tilde{g}_k\{s(1+x)\}\}} \tilde{C}_k\{s(1+x)\}$$
$$\times \gamma_k\{s(1+x)\} x \nu(dx).$$

4.1.1 Binomial Tree Model

For the well known binomial tree model,

$$\xi = \begin{cases} \tilde{u} = \frac{u}{1+r} - 1 & \text{with prob. } p, \\ \tilde{d} = \frac{d}{1+r} - 1 & \text{with prob. } 1-p. \end{cases}$$

We can verify that in this case

$$\int \tilde{g}_k\{s(1+x)\} \frac{(1-b_0 x)}{\lambda} \nu(dx) = q \tilde{g}_k\left(\frac{su}{1+r}\right) + (1-q)\tilde{g}_k\left(\frac{sd}{1+r}\right),$$

where $q = -\frac{\tilde{d}}{\tilde{u}-\tilde{d}} = \frac{1+r-d}{u-d}$. Hence, g_k is really the value of the American option in that setting, proving that τ^* is the well-known optimal stopping strategy.

It is also possible to check that $C_k = Z_k$ on $\{\tau^* \geq k\}$ and $H_{\tau^*} = 0$. The proof is given in Sect. 7.

4.1.2 Implementation

Here, $\xi_k = e^{R_k} - 1$, where the returns R_k are independent observations of R, with mean $\mu_p = \frac{T}{n}\left(\mu - r - \frac{\sigma^2}{2}\right)$ and variance $\sigma_p^2 = \frac{T}{n}\sigma^2$. In order to preserve the monotonicity and convexity properties of the functions $\tilde{g}_k$, it is suggested in Del Moral et al. [13] to generate $x_1, \ldots, x_N$ with the same law as $\xi = e^R - 1$ and set $\hat{b}_0 = \frac{\hat{\mu}}{\hat{B}}$ and $\hat{\lambda} = 1 - \hat{\mu}\hat{b}_0$, where

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^N x_i, \qquad \hat{B} = \frac{1}{N}\sum_{i=1}^N x_i^2.$$

On a grid $\{s_j; 1 \leq j \leq m\}$, with $0 < s_1 < \cdots < s_m$, approximate $\tilde{g}_{k-1}$ defined in (7) by

$$\hat{g}_{k-1}(s_j) = \max\left\{ \tilde{f}_{k-1}(s_j), \frac{1}{N}\sum_{i=1}^N \hat{g}_k\{s_j(1+x_i)\}\frac{(1-\hat{b}_0 x_i)}{\hat{\lambda}} \right\},$$

where $\hat{g}_k$ is linearly interpolated on the grid, except when $k = n$, where $\hat{g}_n = \tilde{f}_n$.

Next, $\tilde{\mathbb{A}}_k$, $\tilde{\mathbb{B}}_k$, γ_k, $\tilde{C}_k$ and $\tilde{a}_k$ are similarly approximated. More precisely, for all $j = 1, \ldots, m$,

$$\widehat{\tilde{\mathbb{A}}}_k(1, s_j) = \frac{1}{N} \sum_{i=1}^{N} x_i^2 \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} = \hat{g}_k\{s_j(1+x_i)\}\}} + \frac{1}{N} \sum_{i=1}^{N} x_i^2 \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} < \hat{g}_k\{s_j(1+x_i)\}\}} \hat{\gamma}_k\{s_j(1+x_i)\},$$

$$\hat{\mu}_k(s_j) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} = \hat{g}_k\{s_j(1+x_i)\}\}} + \frac{1}{N} \sum_{i=1}^{N} x_i \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} < \hat{g}_k\{s_j(1+x_i)\}\}} \hat{\gamma}_k\{s_j(1+x_i)\},$$

$$\widehat{\tilde{\mathbb{B}}}_k(1, s_j) = \hat{\mu}_k(s_j) / \widehat{\tilde{\mathbb{A}}}_k(1, s_j),$$

$$\hat{\gamma}_{k-1}(s_j) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} = \hat{g}_k\{s_j(1+x_i)\}\}} + \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} < \hat{g}_k\{s_j(1+x_i)\}\}} \hat{\gamma}_k\{s_j(1+x_i)\} - \hat{\mu}_k(s_j)^2 / \widehat{\tilde{\mathbb{A}}}_k(1, s_j),$$

$$\widehat{\tilde{a}}_k(s_j) = \frac{1}{\widehat{\tilde{\mathbb{A}}}_k(1, s_j)} \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} = \hat{g}_k\{s_j(1+x_i)\}\}} \tilde{f}_k\{s_j(1+x_i)\} x_i + \frac{1}{\widehat{\tilde{\mathbb{A}}}_k(1, s_j)} \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} < \hat{g}_k\{s_j(1+x_i)\}\}} \times \widehat{\tilde{C}}_k\{s_j(1+x_i)\} \hat{\gamma}_k\{s_j(1+x_i)\} x_i,$$

$$\widehat{\tilde{C}}_{k-1}(s_j) = \frac{1}{\hat{\gamma}_{k-1}(s_j)} \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} = \hat{g}_k\{s_j(1+x_i)\}\}} \tilde{f}_k\{s_j(1+x_i)\} + \frac{1}{\hat{\gamma}_{k-1}(s_j)} \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\{\tilde{f}_k\{s_j(1+x_i)\} < \hat{g}_k\{s_j(1+x_i)\}\}} \times \widehat{\tilde{C}}_k\{s_j(1+x_i)\} \hat{\gamma}_k\{s_j(1+x_i)\} - \frac{a_k(s_j)\hat{\mu}_k(s_j)}{\hat{\gamma}_{k-1}(s_j)}.$$

Example 4.1 (Black-Scholes model). A first experiment consists in comparing results of an American put evaluation under the Black-Scholes setting with parameters r, μ, σ, the time scale being expressed in years. It follows that the excess log-returns are Gaussian, with mean $\left(\mu - r - \frac{\sigma^2}{2}\right)\frac{T}{n}$ and variance $\sigma^2\frac{T}{n}$. In that case,

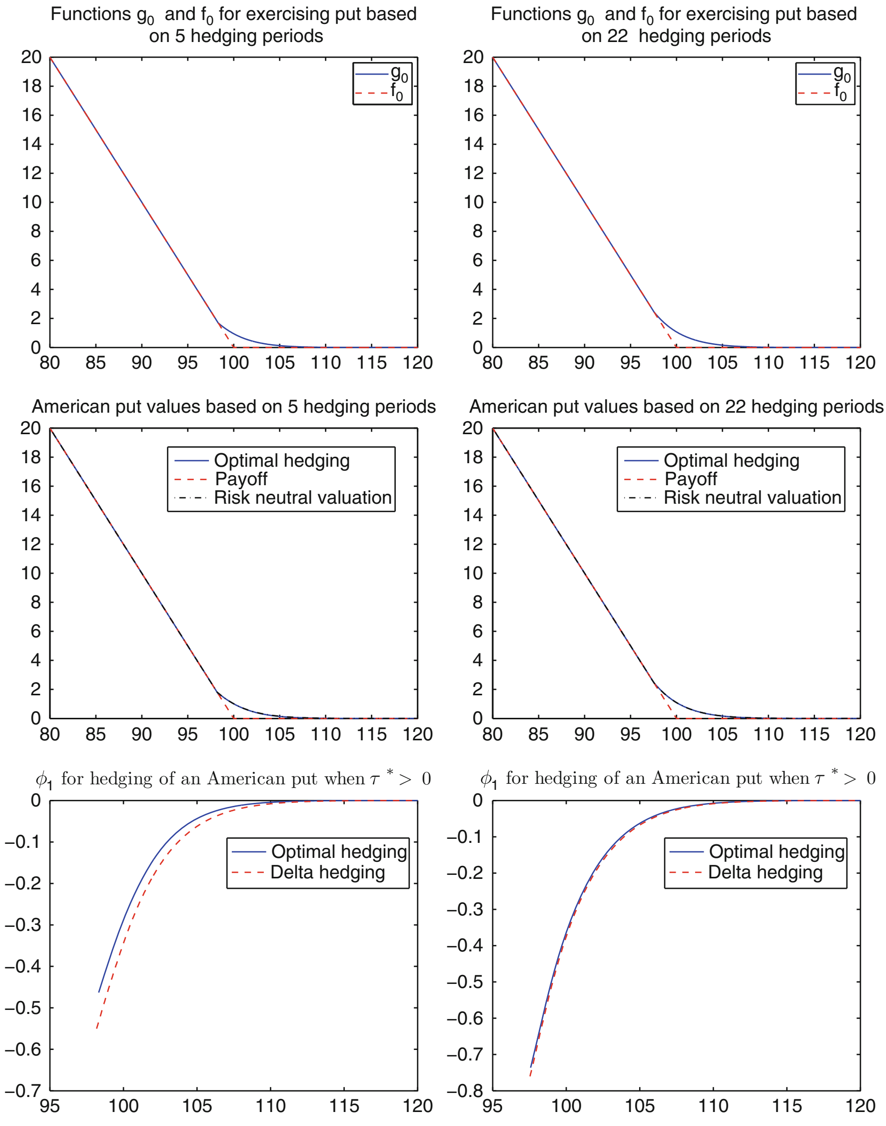

Fig. 1 Implementation results for a Black-Scholes model with $n = 5$ and $n = 22$ hedging periods, $N = 50,000$ simulated returns, on an equally spaced grid of $m = 2,001$ points, ranging from $S = 80$ to $S = 120$

since the model is complete, the risk neutral valuation can be used as a benchmark for the proposed methodology. The values of the parameters are $\mu = 0.09$, $\sigma = 0.06$, $T = 1$ year, $r = 0.05$, and $K = 100$.

Results for the exercising strategy, put values and number of shares at period 0 are given in Fig. 1. For $n = 5, 22, 250$ hedging periods, statistics of the hedging errors $H_\tau = e^{-r\tau} f_\tau - \pi_\tau$ are given in Table 1, corresponding to 10,000 scenarios. Estimation of their densities are pictured in Fig. 2. As expected, the chosen measure of accuracy given by the root mean square error (RMSE) is minimal for the optimal hedging. Also, because the model is complete, the limiting hedging error should be zero. The simulation results support that property since the RMSE is decreasing as n increases.

Example 4.2 (Variance Gamma model). As a second experiment, consider the simple Variance Gamma model of Madan et al. [23]. Here

$$R_k = \left(\mu - r - \beta - \frac{\sigma^2}{2}\right) \frac{T}{n} \beta \xi_k + \sigma Z_k \sqrt{\zeta_k},$$

with $Z_k \sim N(0,1)$, independent of ζ_k has a Gamma distribution with parameters $\left(\alpha \frac{T}{n}, 1/\alpha\right)$.

For sake of comparisons, the same values for r, μ, σ are used. One also set $\alpha = 1$ and $\beta = 0$, so the annual excess log-returns have a Laplace distribution.

Results for the exercising strategy, put values and number of shares at period 0 are given in Fig. 3. For $n = 5, 22, 250$ hedging periods, statistics of the hedging errors $H_\tau = e^{-r\tau} f_\tau - \pi_\tau$ are given in Table 2, corresponding to 10,000 scenarios. Estimation of their densities are pictured in Fig. 4. As expected, the chosen measure of accuracy given by the root mean square error (RMSE) is minimal for the optimal

Table 1 Statistics for hedging errors $H_\tau = e^{-r\tau} f_\tau - \pi_\tau$ with 10,000 scenarios for the Black-Scholes model

Statistic	$n = 5$ Hedging method		$n = 22$ Hedging method		$n = 250$ Hedging method	
	Optimal	Delta	Optimal	Delta	Optimal	Delta
Average	−0.0093	0.0514	−0.0014	0.0800	0.0176	0.0173
Median	−0.0303	0.0288	0.0021	0.0679	0.0068	0.0140
Volatility	0.6338	0.6997	0.3337	0.3570	0.1067	0.1138
Skewness	0.9994	0.2828	−0.1084	0.2235	1.3798	0.2940
Kurtosis	8.8976	4.7343	8.2623	4.5881	17.2701	5.2060
Minimum	−2.6930	−2.8857	−2.2118	−1.6005	−0.6897	−0.5562
Maximum	5.4555	4.7989	2.7628	2.0916	1.3398	0.7589
VaR(99%)	2.1448	2.0241	1.0064	1.0983	0.4076	0.3392
VaR(99.9%)	3.4508	2.9906	1.5167	1.4981	0.7864	0.5360
RMSE	0.6339	0.7016	0.3337	0.3659	0.1081	0.1151

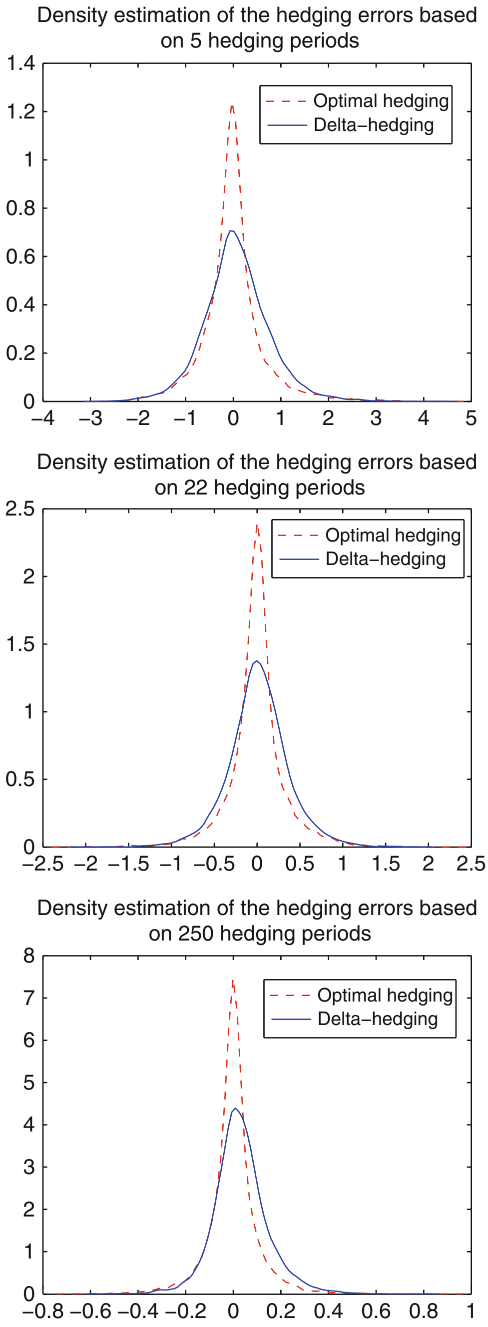

Fig. 2 Nonparametric estimation of the hedging errors in the Black-Scholes model, for $n = 5, 22, 250$ and 10,000 scenarios

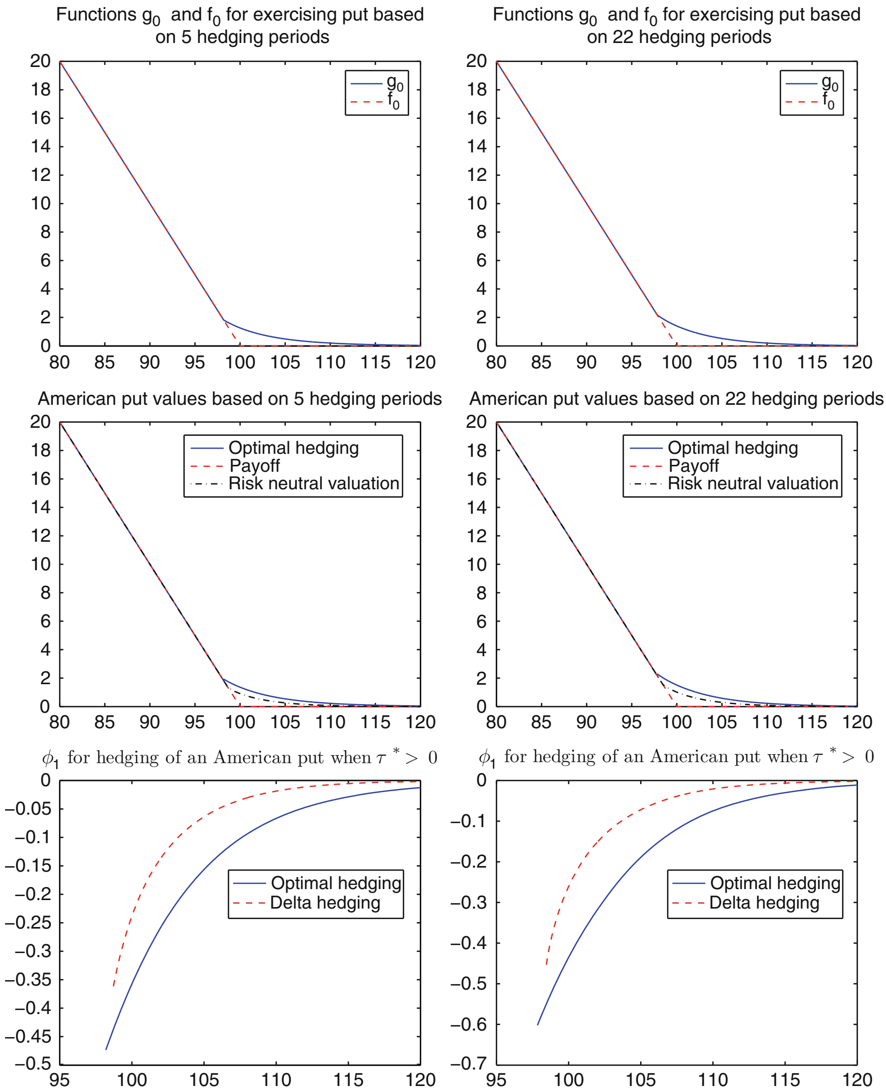

Fig. 3 Implementation results for a Variance Gamma model with $n = 5$ and $n = 22$ hedging periods, $N = 50,000$ simulated returns, on an equally spaced grid of $m = 2,001$ points, ranging from $S = 80$ to $S = 120$

hedging. Also, because the model is not complete, one cannot expect to achieve a very small RSME as in the Black-Scholes model. In fact, the results of the RSME in Table 2 are just slightly bigger than those of a European put option (not reported) given by 1.12 for the optimal hedging and 1.24 for delta hedging, and this relative comparison holds whatever n. Even if the optimal hedging prices are

Table 2 Statistics for hedging errors $H_\tau = e^{-r\tau} f_\tau - \pi_\tau$ with 10,000 scenarios for the Variance Gamma model

Statistic	$n = 5$ Hedging method		$n = 22$ Hedging method		$n = 250$ Hedging method	
	Optimal	Delta	Optimal	Delta	Optimal	Delta
Average	−0.0036	0.0883	0.0463	0.1580	−0.2765	0.3318
Median	−0.2429	−0.3714	−0.1777	−0.3845	−0.4428	−0.2605
Volatility	1.3152	1.4888	1.2697	1.5571	1.3148	1.5749
Skewness	3.4730	4.4748	3.3062	4.1763	1.9516	3.9836
Kurtosis	24.0348	31.2111	24.1017	26.1282	12.8497	24.3123
Minimum	−3.0890	−1.6807	−3.5971	−1.3954	−4.1502	−0.7568
Maximum	14.7487	19.1559	16.8558	19.0261	12.6207	18.3323
VaR(99%)	5.4306	6.8277	5.2051	7.4403	4.3733	7.7534
VaR(99.9%)	11.1911	13.3206	10.5964	12.4115	9.2105	13.2278
RMSE	1.3152	1.4914	1.2705	1.5651	1.3436	1.6095

larger than the delta hedging prices, the values of ϕ_1 are quite different, as shown by the bottom graphs. Remember that for most incomplete models, delta hedging is far from being optimal to evaluate the strategies ϕ_k as discussed in Rémillard and Rubenthaler [29].

5 Conclusion

In this paper we considered the case of a financial derivative with American exercise feature which is hedged in discrete time. Given our choice of a stopping rule, we provided the optimal solution that minimizes the variance of the hedging error, and the price associated with that replicating strategy.

The optimal solution is derived for the general case where assets' returns are model by a Lévy process. By augmenting if necessary the underlying assets' price vector such that it is Markovian, the optimal solution can be approximated easily using readily available numerical techniques for pricing Bermudan options previously proposed in the literature.

Finally, the hedging performance is evaluated with a Monte Carlo experiment where the underlying return dynamics is either Gaussian or Variance Gamma. In each case, it is shown that the root mean square hedging error is lower with our scheme than with traditional delta hedging. This result holds even as the number of hedging steps increases.

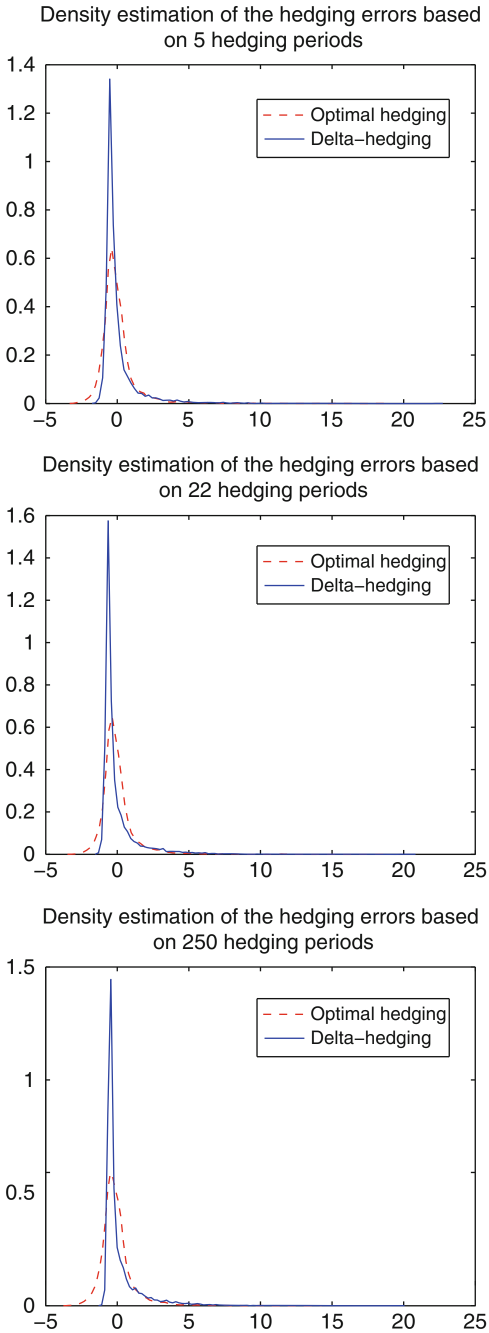

Fig. 4 Nonparametric estimation of the hedging errors in the Variance Gamma model for $n = 5, 22, 250$ and 10,000 scenarios

6 Proofs of the Main Results

6.1 *Proof of Lemma 2.1*

Suppose that $\overrightarrow{\boldsymbol{\phi}}$ is the optimal solution and for a fixed $j \in \{1,\dots,n\}$, let $\overrightarrow{\boldsymbol{\psi}}$ be the strategy defined by $\boldsymbol{\psi}_k = 0$, $k \neq j$, with $\boldsymbol{\psi}_j$ bounded $\mathcal{F}_{j-1}$-measurable. It follows that for any $\epsilon \in \mathbb{R}$,

$$
\begin{aligned}
L(\epsilon) &= \left\{H_\tau^2\left(\pi_0, \overrightarrow{\boldsymbol{\phi} + \epsilon \boldsymbol{\psi}}\right)\right\} \\
&= \sum_{k=0}^{j-1} E\left\{\mathbb{I}(\tau = k)\left(\beta_k f_k - \pi_k\right)^2\right\} \\
&\quad + \sum_{k=j}^{n} E\left\{\mathbb{I}(\tau = k)\left(\beta_k f_k - \pi_k - \epsilon \boldsymbol{\psi}_j^\top \boldsymbol{\Delta}_j\right)^2\right\} \\
&\geq L(0).
\end{aligned}
$$

As a result,

$$
\begin{aligned}
0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon} L(\epsilon)\bigg|_{\epsilon=0} &= -2 \sum_{k=j}^{n} E\left\{\mathbb{I}(\tau = k)\left(\beta_k f_k - \pi_k\right) \boldsymbol{\psi}_j^\top \boldsymbol{\Delta}_j\right\} \\
&= -2E\left\{\mathbb{I}(\tau \geq j)\left(\beta_\tau f_\tau - \pi_\tau\right) \boldsymbol{\psi}_j^\top \boldsymbol{\Delta}_j\right\}.
\end{aligned}
$$

Since the latter is true for any bounded $\mathcal{F}_{j-1}$-measurable function $\boldsymbol{\psi}_j$, and $\{\tau \geq j\} = \{\tau \leq j-1\}^c \in \mathcal{F}_{j-1}$, it follows that conditions (1)–(2) hold true.

On the other hand, if conditions (1)–(2) are met, then it is easy to check that for any other strategy $v_0, \overrightarrow{\boldsymbol{\psi}}$, setting $v_k = v_{k-1} + \boldsymbol{\psi}_k^\top \boldsymbol{\Delta}_k$, $k = 1,\dots,n$, one has

$$
\begin{aligned}
E\left\{H_\tau^2\left(v_0, \overrightarrow{\boldsymbol{\psi}}\right)\right\} &= E\left[\left\{\beta_\tau f_\tau - \pi_\tau + \pi_\tau - v_\tau\right\}^2\right] \\
&= E\left\{\left(\beta_\tau f_\tau - \pi_\tau\right)^2\right\} + E\left\{\left(\pi_\tau - v_\tau\right)^2\right\} \\
&\quad + \left(\pi_0 - v_0\right) E\left\{\left(\beta_\tau f_\tau - \pi_\tau\right)\right\} \\
&\quad + \sum_{j=1}^{n} E\left\{\mathbb{I}(\tau \geq j)\left(\beta_\tau f_\tau - \pi_\tau\right) \boldsymbol{\Delta}_j^\top \left(\boldsymbol{\phi}_j - \boldsymbol{\psi}_j\right)\right\} \\
&= E\left\{H_\tau^2\left(\pi_0, \overrightarrow{\boldsymbol{\phi}}\right)\right\} + E\left\{\left(\pi_\tau - v_\tau\right)^2\right\},
\end{aligned}
$$

since $\{\tau \geq j\} \in \mathcal{F}_{j-1}$. Hence the result. □

6.2 Proof of Proposition 2.1

We start by proving (3) for $j = n$ and then use reverse induction. Note that for all $j = 1, \ldots, n$, $\{\tau \geq j\} = \{\tau \leq j-1\}^c \in \mathcal{F}_{j-1}$, so (5) obviously holds for $j = n$ since $\mathcal{P}_{n+1} = 1$.

As a result,

$$
\begin{aligned}
0 &= E\left\{\mathbb{I}(\tau \geq n)(\beta_\tau f_\tau - \pi_\tau)\boldsymbol{\Delta}_n|\mathcal{F}_{n-1}\right\}\\
&= \mathbb{I}(\tau = n)\,E\left\{\left(\beta_n f_n - \pi_{n-1} - \boldsymbol{\phi}_n^\top\boldsymbol{\Delta}_n\right)\boldsymbol{\Delta}_n|\mathcal{F}_{n-1}\right\}\\
&= \mathbb{I}(\tau = n)\,\mathcal{A}_n\left(\mathfrak{a}_n - \pi_{n-1}\mathfrak{b}_n - \boldsymbol{\phi}_n\right).
\end{aligned}
$$

Hence, on $\{\tau \geq n\}$, $\boldsymbol{\phi}_n = \mathfrak{a}_n - \pi_{n-1}\mathfrak{b}_n$.

Suppose now that (3)–(5) hold for $k = j, \ldots, n$. We have to prove that (3)–(5) also hold for $k = j-1$.

First, by the induction hypothesis, on $\{\tau \geq j\}$, $\boldsymbol{\phi}_j = \mathfrak{a}_j - \pi_{j-1}\mathfrak{b}_j$, so

$$
\begin{aligned}
E\left\{\mathbb{I}(\tau \geq j)\boldsymbol{\phi}_j^\top\boldsymbol{\Delta}_j\mathcal{P}_{j+1}|\mathcal{F}_{j-1}\right\} &= \mathbb{I}(\tau \geq j)\,\boldsymbol{\phi}_j^\top\mathcal{A}_j\mathfrak{b}_j\\
&= \mathbb{I}(\tau \geq j)\,\mathfrak{a}_j^\top\mathcal{A}_j\mathfrak{b}_j - \pi_{j-1}\mathbb{I}(\tau \geq j)\,\mathfrak{b}_j^\top\mathcal{A}_j\mathfrak{b}_j\\
&= \mathbb{I}(\tau \geq j)\,E\left(\beta_\tau f_\tau\mathfrak{b}_j^\top\boldsymbol{\Delta}_j\mathcal{P}_{j+1}|\mathcal{F}_{j-1}\right)\\
&\quad -\pi_{j-1}\mathbb{I}(\tau \geq j)\,E\left(\mathfrak{b}_j^\top\boldsymbol{\Delta}_j\mathcal{P}_{j+1}|\mathcal{F}_{j-1}\right)\\
&= E\left\{\mathbb{I}(\tau \geq j)\left(\beta_\tau f_\tau - \pi_{j-1}\right)\mathfrak{b}_j^\top\boldsymbol{\Delta}_j\mathcal{P}_{j+1}|\mathcal{F}_{j-1}\right\}.
\end{aligned}
$$

Since $\{\tau \geq j-1\} \in \mathcal{F}_{j-2}$ and $\mathcal{P}_j = 1$ on $\{\tau = j-1\}$, one has

$$
\begin{aligned}
0 &= E\left\{\mathbb{I}(\tau \geq j-1)(\beta_\tau f_\tau - \pi_\tau)\boldsymbol{\Delta}_{j-1}|\mathcal{F}_{j-2}\right\}\\
&= E\left\{\mathbb{I}(\tau \geq j)(\beta_\tau f_\tau - \pi_\tau)\boldsymbol{\Delta}_{j-1}|\mathcal{F}_{j-2}\right\}\\
&\quad + E\left\{\mathbb{I}(\tau = j-1)\left(\beta_{j-1}f_{j-1} - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}|\mathcal{F}_{j-2}\right\}\\
&= E\left[E\left\{\mathbb{I}(\tau \geq j)(\beta_\tau f_\tau - \pi_\tau)\boldsymbol{\Delta}_{j-1}|\mathcal{F}_{j-1}\right\}|\mathcal{F}_{j-2}\right]\\
&\quad + E\left\{\mathbb{I}(\tau = j-1)\left(\beta_{j-1}f_{j-1} - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}|\mathcal{F}_{j-2}\right\}\\
&= E\left\{\mathbb{I}(\tau \geq j)\left(\beta_\tau f_\tau - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_{j+1}|\mathcal{F}_{j-2}\right\}\\
&\quad - E\left\{\mathbb{I}(\tau \geq j)\boldsymbol{\phi}_j^\top\boldsymbol{\Delta}_j\boldsymbol{\Delta}_{j-1}\mathcal{P}_{j+1}|\mathcal{F}_{j-2}\right\}\\
&\quad + E\left\{\mathbb{I}(\tau = j-1)\left(\beta_{j-1}f_{j-1} - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_j|\mathcal{F}_{j-2}\right\}\\
&= E\left\{\mathbb{I}(\tau \geq j)\left(\beta_\tau f_\tau - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_{j+1}|\mathcal{F}_{j-2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
&-E\left\{\mathbb{I}\left(\tau \geq j\right)\left(\beta_\tau f_\tau - \pi_{j-1}\right)\mathfrak{b}_j^\top \boldsymbol{\Delta}_j \mathcal{P}_{j+1}\boldsymbol{\Delta}_{j-1}|\mathcal{F}_{j-2}\right\}\\
&\quad +E\left\{\mathbb{I}\left(\tau = j-1\right)\left(\beta_{j-1} f_{j-1} - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_j|\mathcal{F}_{j-2}\right\}\\
= &\ E\left\{\mathbb{I}\left(\tau \geq j\right)\left(\beta_\tau f_\tau - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_j|\mathcal{F}_{j-2}\right\}\\
&\quad +E\left\{\mathbb{I}\left(\tau = j-1\right)\left(\beta_{j-1} f_{j-1} - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_j|\mathcal{F}_{j-2}\right\}\\
= &\ E\left\{\mathbb{I}\left(\tau \geq j-1\right)\left(\beta_\tau f_\tau - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_j|\mathcal{F}_{j-2}\right\},
\end{aligned}
$$

proving (5) for $j-1$.

Furthermore, since $\{\tau \geq j-1\} \in \mathcal{F}_{j-2}$, one has

$$
\begin{aligned}
0 = &\ E\left\{\mathbb{I}\left(\tau \geq j-1\right)\left(\beta_\tau f_\tau - \pi_{j-1}\right)\boldsymbol{\Delta}_{j-1}\mathcal{P}_j|\mathcal{F}_{j-2}\right\}\\
= &\ \mathbb{I}\left(\tau \geq j-1\right)E\left(\beta_\tau f_\tau \boldsymbol{\Delta}_{j-1}P_j|\mathcal{F}_{j-2}\right) - \mathbb{I}\left(\tau \geq j-1\right)\pi_{j-2}E\left(\boldsymbol{\Delta}_{j-1}\mathcal{P}_j|\mathcal{F}_{j-2}\right)\\
&\quad -\mathbb{I}\left(\tau \geq j-1\right)E\left(\boldsymbol{\Delta}_{j-1}\boldsymbol{\Delta}_{j-1}^\top\mathcal{P}_j|\mathcal{F}_{j-2}\right)\boldsymbol{\phi}_{j-1}\\
= &\ \mathbb{I}\left(\tau \geq j-1\right)\mathcal{A}_{j-1}\left(\mathfrak{a}_{j-1} - \pi_{j-2}\mathfrak{b}_{j-1} - \boldsymbol{\phi}_{j-1}\right).
\end{aligned}
$$

As a result one must have $\boldsymbol{\phi}_{j-1} = \mathfrak{a}_{j-1} - \pi_{j-2}\mathfrak{b}_{j-1}$ on $\{\tau \geq j-1\}$. Hence one may conclude that (3)–(5) hold for $j = 1, \ldots, n$.

In particular, taking $j = 1$ in (5), and using the condition $E\left(\beta_\tau f_\tau - \pi_\tau\right) = 0$, one gets

$$
\begin{aligned}
0 = &\ \mathbb{I}\left(\tau \geq 1\right)E\left(\beta_\tau f_\tau - \pi_\tau\right) = \mathbb{I}\left(\tau \geq 1\right)E\{\left(\beta_\tau f_\tau - \pi_1\right)\mathcal{P}_2\}\\
= &\ \mathbb{I}\left(\tau \geq 1\right)\left\{E\left(\beta_\tau f_\tau\right) - \pi_0 E\left(\mathcal{P}_2\right) - \boldsymbol{\phi}_1^\top \mathcal{A}_1 \mathfrak{b}_1\right\}\\
= &\ \mathbb{I}\left(\tau \geq 1\right)\left\{E\left(\beta_\tau f_\tau\right) - \pi_0 E\left(\mathcal{P}_2\right) - \left(\mathfrak{a}_1 - \pi_0\mathfrak{b}_1\right)^\top \mathcal{A}_1 \mathfrak{b}_1\right\}\\
= &\ \mathbb{I}\left(\tau \geq 1\right)\left\{E\left(\beta_\tau f_\tau \mathcal{P}_1\right) - \pi_0 E\left(\mathcal{P}_1\right)\right\}
\end{aligned}
$$

which completes the proof. □

6.3 *Proof of Proposition 2.2*

By definition of $\mathcal{C}_k$,

$$
\begin{aligned}
\mathbb{I}\left(\tau^* > k\right)\mathfrak{c}_k\beta_k\mathcal{C}_k = &\ \mathbb{I}\left(\tau^* > k\right)E\left(\beta_{\tau^*} f_{\tau^*}\mathcal{P}_{k+1}|\mathcal{F}_k\right)\\
= &\ E\{\mathbb{I}\left(\tau^* = k+1\right)\beta_{k+1} f_{k+1}\mathcal{P}_{k+1}|\mathcal{F}_k\}\\
&\quad +E\{\mathbb{I}\left(\tau^* > k+1\right)\beta_{\tau^*} f_{\tau^*}\left(1 - \mathfrak{b}_{k+1}^\top\boldsymbol{\Delta}_{k+1}\right)\mathcal{P}_{k+2}|\mathcal{F}_k\}\\
= &\ E\{\mathbb{I}\left(\tau^* = k+1\right)\mathcal{C}_{k+1}\mathcal{P}_{k+1}|\mathcal{F}_k\}
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \mathbb{I}\left(\tau^* > k+1\right) E\{\left(1 - \mathfrak{b}_{k+1}^\top \boldsymbol{\Delta}_{k+1}\right) \\
&\qquad \times E\left(\beta_{\tau^*} f_{\tau^*} \mathcal{P}_{k+2} | \mathcal{F}_{k+1}\right) | \mathcal{F}_k\} \\
&= E\{\mathbb{I}\left(\tau^* = k+1\right) \mathcal{C}_{k+1} \mathcal{P}_{k+1} | \mathcal{F}_k\} \\
&\quad + \mathbb{I}\left(\tau^* > k+1\right) E\{\left(1 - \mathfrak{b}_{k+1}^\top \boldsymbol{\Delta}_{k+1}\right) \beta_{k+1} \mathcal{C}_{k+1} \mathfrak{c}_{k+1} | \mathcal{F}_k\} \\
&= E\{\mathbb{I}\left(\tau^* = k+1\right) \beta_{k+1} \mathcal{C}_{k+1} E\left(\mathcal{P}_{k+1} | \mathcal{F}_{k+1}\right) | \mathcal{F}_k\} \\
&\quad + E\{\mathbb{I}\left(\tau^* > k+1\right) \beta_{k+1} \mathcal{C}_{k+1} E\left(\mathcal{P}_{k+1} | \mathcal{F}_{k+1}\right) | \mathcal{F}_k\} \\
&= \mathbb{I}\left(\tau^* > k\right) E\{\beta_{k+1} \mathcal{C}_{k+1} E\left(\mathcal{P}_{k+1} | \mathcal{F}_{k+1}\right) | \mathcal{F}_k\} \\
&= \mathbb{I}\left(\tau^* > k\right) E\{\beta_{k+1} \mathcal{C}_{k+1} \mathcal{P}_{k+1} | \mathcal{F}_k\}.
\end{aligned}
$$

□

6.4 Proof of Proposition 3.1

First, recall that by definition, $\mathcal{C}_n = f_n$ on $\{\tau^* = n\}$. Next,

$$
\mathbb{I}_{\{\tau^* > n-1\}} \mathcal{C}_{n-1} = \frac{1}{(1+r)\,\gamma_{n-1}\left(\mathbf{Y}_{n-1}\right)} E\left[f_n\left(\mathbf{S}_n\right) \{1 - \mathbb{B}_n\left(1, \mathbf{Y}_{n-1}\right)^\top \boldsymbol{\Delta}_n\} | \mathcal{F}_{n-1}\right],
$$

so one can set

$$
C_{n-1}\left(\mathbf{y}\right) = \frac{1}{(1+r)\,\gamma_{n-1}\left(\mathbf{y}\right)} E\left[f_n\left(S_n\right) \{1 - \mathbb{B}_n\left(1, \mathbf{y}\right)^\top \boldsymbol{\Delta}_n\} | \mathbf{Y}_{n-1} = \mathbf{y}\right].
$$

It follows that $\mathbb{A}_n\left(1, \mathbf{y}\right) = A_n\left(\mathbf{y}\right)$, $\mathbb{B}_n\left(1, \mathbf{y}\right) = b_n\left(\mathbf{y}\right)$, $\gamma_{n-1}\left(\mathbf{y}\right) = c_{n-1}\left(\mathbf{y}\right)$ and $g_{n-1}\left(\mathbf{y}\right) = C_{n-1}\left(\mathbf{y}\right)$.

Next, suppose $\mathbb{I}\left(\tau^* > j\right) \mathcal{C}_j = \mathbb{I}\left(\tau^* > j\right) C_j\left(\mathbf{Y}_j\right)$ for all $j > k$. Then

$$
\begin{aligned}
\mathbb{I}_{\{\tau^* > k\}} \mathfrak{c}_k \beta_k \mathcal{C}_k &= \mathbb{I}_{\{\tau^* > k\}} E\{\beta_{k+1} \mathcal{C}_{k+1} \mathcal{P}_{k+1} | \mathcal{F}_k\} \\
&= E\{\mathbb{I}_{\{\tau^* = k+1\}} \beta_{k+1} \mathcal{C}_{k+1} \mathcal{P}_{k+1} | \mathcal{F}_k\} \\
&\quad + E\{\mathbb{I}_{\{\tau^* > k+1\}} \beta_{k+1} \mathcal{C}_{k+1} \mathcal{P}_{k+1} | \mathcal{F}_k\} \\
&= \mathbb{I}_{\{\tau^* > k\}} E\Big[\mathbb{I}_{\{f_{k+1} = Z_{k+1}\}} \beta_{k+1} f_{k+1} \\
&\qquad \times \left\{1 - \mathbb{B}_{k+1}\left(1, \mathbf{Y}_k\right)^\top \boldsymbol{\Delta}_{k+1}\right\} | \mathcal{F}_k \Big] \\
&\qquad + \mathbb{I}_{\{\tau^* > k\}} E\Big[\mathbb{I}_{\{f_{k+1} < Z_{k+1}\}} \beta_{k+1} C_{k+1}\left(\mathbf{Y}_{k+1}\right) \\
&\qquad\quad \times \left\{1 - \mathbb{B}_{k+1}\left(1, \mathbf{Y}_k\right)^\top \boldsymbol{\Delta}_{k+1}\right\} \gamma_{k+1}\left(\mathbf{Y}_{k+1}\right) | \mathcal{F}_k \Big],
\end{aligned}
$$

proving that $\mathbb{I}(\tau^* > k)\,\mathcal{C}_k = \mathbb{I}(\tau^* > k)\,C_j(\mathbf{Y}_k)$. It follows that $\mathbb{I}(\tau^* > k)\,\mathcal{C}_k = \mathbb{I}(\tau^* > k)\,C_k(\mathbf{Y}_k)$ for all $k = 0, \ldots, n-1$.

Finally,

$$\begin{aligned}\mathfrak{a}_k &= \mathbb{I}_{\{\tau^* \geq k\}} \mathcal{A}_k^{-1} E\left(\beta_{\tau^*} f_{\tau^*} \mathbf{\Delta}_k \mathcal{P}_{k+1} | \mathcal{F}_{k-1}\right), \\ &= \mathbb{I}_{\{\tau^* \geq k\}} \mathbb{A}_k^{-1}(1, \mathbf{Y}_{k-1})\, E\left\{\mathbb{I}_{\{f_k = Z_k\}} \beta_k f_k \mathbf{\Delta}_k | \mathcal{F}_{k-1}\right\} \\ &\qquad + \mathbb{I}_{\{\tau^* \geq k\}} \mathbb{A}_k^{-1}(1, \mathbf{Y}_{k-1})\, E\left\{\mathbb{I}_{\{f_k < Z_k\}} \beta_k \mathbf{\Delta}_k \gamma_k(\mathbf{Y}_k)\, C_k(\mathbf{Y}_k) | \mathcal{F}_{k-1}\right\},\end{aligned}$$

proving that $\mathfrak{a}_k = \mathbb{I}_{\{\tau^* \geq k\}} \mathbf{a}_k(\mathbf{Y}_{k-1})$. □

7 Proof of the Perfect Hedging in the Binomial Tree Model

To prove the first statement, note that it is obviously true for $k = n$. Suppose now that it is true for k. One will show that it is also true for $k-1$. Since $\mathcal{C}_{k-1} = \tilde{f}_{k-1}$ on $\{\tau^* = k-1\}$, it suffices to show that $\tilde{C}_{k-1} = \tilde{g}_{k-1}$ on $\{\tau^* > k-1\}$.

First, $s = s_0 u^j d^{k-1-j} / (1+r)^k$, for some $j = 0, \ldots, k-1$. It follows that $\tilde{\mathbb{A}}_k(1, s) = p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0$ for some $\theta_0, \theta_1 \in (0, 1]$. In fact, $\theta_1 = \mathbb{I}_{\{\tilde{f}_k\left(\frac{su}{1+r}\right) = \tilde{g}_k\left(\frac{su}{1+r}\right)\}} + \mathbb{I}_{\{\tilde{f}_k\left(\frac{su}{1+r}\right) < \tilde{g}_k\left(\frac{su}{1+r}\right)\}} \gamma_k\left(\frac{su}{1+r}\right)$ and $\theta_0 = \mathbb{I}_{\{\tilde{f}_k\left(\frac{sd}{1+r}\right) = \tilde{g}_k\left(\frac{sd}{1+r}\right)\}} + \mathbb{I}_{\{\tilde{f}_k\left(\frac{sd}{1+r}\right) < \tilde{g}_k\left(\frac{sd}{1+r}\right)\}} \gamma_k\left(\frac{sd}{1+r}\right)$. Also $\tilde{\mathbb{B}}_k(1, s) = \frac{p\tilde{u}\theta_1 + (1-p)\tilde{d}\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\tilde{d}^2\theta_0}$ and

$$\begin{aligned}\gamma_{k-1}(s) &= p\theta_1 + (1-p)\,\theta_0 - \frac{\{p\tilde{u}\theta_1 + (1-p)\,\tilde{d}\theta_0\}^2}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0} \\ &= \frac{p(1-p)\,\theta_0\theta_1\left(\tilde{u} - \tilde{d}\right)^2}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0}.\end{aligned}$$

Note that

$$1 - \tilde{\mathbb{B}}_k(1, s)\,\tilde{u} = -\frac{(1-p)\,\theta_0\tilde{d}\left(\tilde{u} - \tilde{d}\right)}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0}$$

and

$$1 - \tilde{\mathbb{B}}_k(1, s)\,\tilde{d} = \frac{p\theta_1\tilde{u}\left(\tilde{u} - \tilde{d}\right)}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0},$$

so $p\theta_1\{1 - \tilde{\mathbb{B}}_k(1, s)\,\tilde{u}\}/\gamma_{k-1}(s) = q$ and $(1-p)\,\theta_0\{1 - \tilde{\mathbb{B}}_k(1, s)\,\tilde{d}\}/\gamma_{k-1}(s) = 1 - q$.

Next, by the induction hypothesis, $\tilde{C}_k = \tilde{g}_k$, so

$$\tilde{C}_{k-1}(s) = \frac{p\theta_1 \tilde{g}_k\left(\frac{su}{1+r}\right)\{1 - \tilde{\mathbb{B}}_k(1,s)\tilde{u}\}}{\gamma_{k-1}(s)} + \frac{(1-p)\theta_0 \tilde{g}_k\left(\frac{sd}{1+r}\right)\{1 - \tilde{\mathbb{B}}_k(1,s)\tilde{d}\}}{\gamma_{k-1}(s)}$$
$$= q\tilde{g}_k\left(\frac{su}{1+r}\right) + (1-q)\tilde{g}_k\left(\frac{sd}{1+r}\right) = \tilde{g}_{k-1}(s),$$

since $\tilde{f}_{k-1}(s) < q\tilde{g}_k\left(\frac{su}{1+r}\right) + (1-q)\tilde{g}_k\left(\frac{sd}{1+r}\right)$ on $\{\tau^* > k-1\}$.

To complete, the proof, note that $\pi_0 = \tilde{g}_0 = Z_0$. Suppose that on $\{\tau^* \geq k-1\}$, $\pi_{k-1} = Z_{k-1}$. One has to prove that $\pi_k = Z_k$ on $\{\tau^* \geq k\}$.

Next, on $\{\tau^* > k-1\}$, $Z_{k-1} = \tilde{g}_{k-1}(s)$ and

$$\tilde{a}_k(s) = \frac{p\tilde{g}_k\left(\frac{su}{1+r}\right)\tilde{u}\theta_1 + (1-p)\tilde{g}_k\left(\frac{sd}{1+r}\right)\tilde{d}\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\tilde{d}^2\theta_0}.$$

It follows that $\pi_k = \tilde{g}_{k-1}(s) + \{\tilde{a}_k - \tilde{\mathbb{B}}_k(1,s)\tilde{g}_{k-1}(s)\}\tilde{u}$ with probability p, and $\pi_k = \tilde{g}_{k-1}(s) + \{\tilde{a}_k - \tilde{\mathbb{B}}_k(1,s)\tilde{g}_{k-1}(s)\}\tilde{d}$ with probability $1-p$. So, on $\{\xi_k = \tilde{u}\}$,

$$\begin{aligned}
\pi_k &= \tilde{g}_{k-1}(s) + \{\tilde{a}_k - \tilde{\mathbb{B}}_k(1,s)\tilde{g}_{k-1}(s)\}\tilde{u} \\
&= \tilde{g}_{k-1}(s)\{1 - \tilde{\mathbb{B}}_k(1,s)\tilde{u}\} + \tilde{a}_k\tilde{u} \\
&= \{q\tilde{g}_k\left(\frac{su}{1+r}\right) + (1-q)\tilde{g}_k\left(\frac{sd}{1+r}\right)\}\{1 - \tilde{\mathbb{B}}_k(1,s)\tilde{u}\} \\
&\quad + \tilde{u}\frac{p\tilde{g}_k\left(\frac{su}{1+r}\right)\tilde{u}\theta_1 + (1-p)\tilde{g}_k\left(\frac{sd}{1+r}\right)\tilde{d}\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\tilde{d}^2\theta_0} \\
&= \tilde{g}_k\left(\frac{su}{1+r}\right)\left[q\{1 - \tilde{\mathbb{B}}_k(1,s)\tilde{u}\} + \frac{p\tilde{u}^2\theta_1}{p\tilde{u}^2\theta_1 + (1-p)\tilde{d}^2\theta_0}\right] \\
&\quad + \tilde{g}_k\left(\frac{sd}{1+r}\right)\left[(1-q)\{1 - \tilde{\mathbb{B}}_k(1,s)\tilde{u}\} + \frac{(1-p)\tilde{u}\tilde{d}\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\tilde{d}^2\theta_0}\right] \\
&= \tilde{g}_k\left(\frac{su}{1+r}\right)\left[-q\frac{\tilde{d}\left(\tilde{u} - \tilde{d}\right)(1-p)\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\tilde{d}^2\theta_0} + \frac{p\tilde{u}^2\theta_1}{p\tilde{u}^2\theta_1 + (1-p)\tilde{d}^2\theta_0}\right] \\
&\quad + \tilde{g}_k\left(\frac{sd}{1+r}\right)
\end{aligned}$$

$$\times \left[-(1-q)\, \frac{\tilde{d}\left(\tilde{u}-\tilde{d}\right)(1-p)\,\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0} + \frac{(1-p)\,\tilde{u}\tilde{d}\,\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0} \right]$$
$$= \tilde{g}_k\left(\frac{su}{1+r}\right).$$

Similarly, on $\{\xi_k = \tilde{d}\}$,

$$\begin{aligned}
\pi_k &= \tilde{g}_{k-1}(s) + \{\tilde{a}_k - \tilde{\mathbb{B}}_k(1,s)\,\tilde{g}_{k-1}(s)\}\tilde{d} \\
&= \tilde{g}_{k-1}(s)\{1 - \tilde{\mathbb{B}}_k(1,s)\,\tilde{d}\} + \tilde{a}_k\tilde{d} \\
&= \{q\tilde{g}_k\left(\frac{su}{1+r}\right) + (1-q)\,\tilde{g}_k\left(\frac{sd}{1+r}\right)\}\{1 - \tilde{\mathbb{B}}_k(1,s)\,\tilde{d}\} \\
&\quad + \tilde{d}\,\frac{p\tilde{g}_k\left(\frac{su}{1+r}\right)\tilde{u}\theta_1 + (1-p)\,\tilde{g}_k\left(\frac{sd}{1+r}\right)\tilde{d}\,\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0} \\
&= \tilde{g}_k\left(\frac{su}{1+r}\right)\left[q\{1 - \tilde{\mathbb{B}}_k(1,s)\,\tilde{d}\} + \frac{p\tilde{u}\tilde{d}\,\theta_1}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0}\right] \\
&\quad + \tilde{g}_k\left(\frac{sd}{1+r}\right)\left[(1-q)\{1 - \tilde{\mathbb{B}}_k(1,s)\,\tilde{d}\} + \frac{(1-p)\,\tilde{d}^2\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0}\right] \\
&= \tilde{g}_k\left(\frac{su}{1+r}\right)\left[q\,\frac{p\tilde{u}\left(\tilde{u}-\tilde{d}\right)\theta_1}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0} + \frac{p\tilde{u}\tilde{d}\,\theta_1}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0}\right] \\
&\quad + \tilde{g}_k\left(\frac{sd}{1+r}\right)\left[\frac{(1-q)\,p\tilde{u}\left(\tilde{u}-\tilde{d}\right)\theta_1}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0} + \frac{(1-p)\,\tilde{d}^2\theta_0}{p\tilde{u}^2\theta_1 + (1-p)\,\tilde{d}^2\theta_0}\right] \\
&= \tilde{g}_k\left(\frac{sd}{1+r}\right).
\end{aligned}$$

Therefore, $\pi_k = Z_k$ on $\{\tau^* \geq k\}$, proving that $H_{\tau^*} = 0$. □

References

1. Bates, D. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *Review of Financial Studies*, 9:69–108.
2. Bates, D. (2003). Empirical option pricing: A retrospection. *Journal of Econometrics*, 116:387–404.
3. Bouchaud, J.-P. (2000). Elements for a Theory of Financial Risks. *Physica A: Statistical Mechanics and its Applications*, 285:18–28.

4. Bouchaud, J.-P. and Potters, M. (2002). Back to basics: historical option pricing revisited. *Philosophical Transactions: Mathematical, Physical & Engineering Sciences*, 357:2019–2028.
5. Broadie, M. and Glasserman, P. (2004). A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance*, 7:35–72.
6. Carr, P. and Wu, L. (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 17:113–141.
7. Carriere, J. (1996). Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: Mathematics and Economics*, 19:19–30.
8. Chan, T. (1999). Pricing contingent claims on stocks driven by Levy processes. *The Annals of Applied Probability*, 9:504–528.
9. Christoffersen, P., Jacobs, K., and Mimoumi, K. (2010). Volatility dynamics for the S&P 500: Evidence from realized volatility, daily returns and option prices. *Review of Financial Studies*, 23:3141–3189.
10. Christoffersen, P., Jacobs, K., Ornthanalai, C., and Wang, Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, 90:272–297.
11. Cornalba, L., Bouchaud, J.-P., and Potters, M. (2002). Option pricing and hedging with temporal correlations. *International Journal of Theoretical and Applied Finance*, 5:307–320.
12. Del Moral, P., Rémillard, B., and Rubenthaler, S. (2006). Monte Carlo approximations of American options. Technical report, GERAD.
13. Del Moral, P., Rémillard, B., and Rubenthaler, S. (2011). Monte Carlo Approximations of American Options that Preserve Monotonicity and Convexity. In Carmona, R., Moral, P. D., Hu, P., and Oudjane, N., editors, *American Styles Options Pricing*, page In press. Springer.
14. Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance*, 5:13–32.
15. Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton University Press.
16. Engle, R. and Lee, G. (1999). *Cointegration, causality, and forecasting: a festschrift in honor of Clive W.J. Granger*, chapter A permanent and transitory component model of stock return volatility, pages 475–497. Oxford University Press, New York.
17. Garleanu, N., Pedersen, L. H., and Poteshman, A. M. (2009). Demand-based option pricing. *Review of Financial Studies*, 22:4259–4299.
18. Heston, S. (1993). A closed-form solution for options with stochastic volatility with application to bond and currency options. *Review of Financial Studies*, 6:327–343.
19. Heston, S. and Nandi, S. (2000). A closed-form GARCH option pricing model. *Review of Financial Studies*, 13:585–626.
20. Huang, J.-Z. and Wu, L. (2004). Specification analysis of option pricing models based on time-changed Lvy processes. *Journal of Finance*, 59:1405–1439.
21. Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42:281–300.
22. Longstaff, F. and Schwartz, E. (2001). Valuing American options by simulation: A simple least-square approach. *The Review of Financial Studies*, 14:113–147.
23. Madan, D. B., Carr, P., and Chang, E. C. (1998). The variance gamma process and option pricing.
24. Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.
25. Naik, V. (1993). Option valuation and hedging strategies with jumps in the volatility of asset returns. *Journal of Finance*, 48:1969–1984.
26. Papageorgiou, N., Rémillard, B., and Hocquard, A. (2008). Replicating the properties of hedge fund returns. *Journal of Alternative Invesments*, 11:8–38.
27. Pochart, B. and Bouchaud, J.-P. (2004). Option pricing and hedging with minimum local expected shortfall. *Quantitative Finance*, 4:607–618.
28. Rémillard, B., Hocquard, A., and Papageorgiou, N. A. (2010). Option pricing and dynamic discrete time hedging for regime-switching geometric random walks models. Technical report, HEC Montreal.
29. Rémillard, B. and Rubenthaler, S. (2009). Optimal hedging in discrete and continuous time. Technical report, GERAD.

30. Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, 20:1–32.
31. Tsitsiklis, J. and Roy, B. V. (1999). Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives. *IEEE Transactions on Automatic Control*, 44:1840–1851.
32. Wilmott, P. (2006). *Paul Wilmott on Quantitative Finance*, volume 3. John Wiley & Sons.