Variance-optimal hedging in general affine stochastic volatility models

Jan Kallsen∗ Arnd Pauwels†

Abstract

We consider variance-optimal hedging in general continuous-time affine stochastic volatility models. The optimal hedge and the associated hedging error are determined semi-explicitly in the case that the stock price follows a martingale. The integral representation of the solution opens the door to efficient numerical computation. The setup includes models with jumps in the stock price and in the activity process. It also allows for correlation between volatility and stock price movements. Concrete parametric models will be illustrated in a forthcoming paper.

Key words: variance-optimal hedging, Galtchouk-Kunita-Watanabe decomposition, stochastic volatility, affine processes, Laplace transform

Mathematics Subject Classification (2000): 91B28, 60H05, 60G48, 93E20

1 Introduction

Often observed so-called stylized features of stock returns include semi-heavy tails, volatility clustering, and the leverage effect (i.e. negative correlation between changes in volatility and stock prices). Stochastic volatility (SV) models account for these observations. Examples in the literature include the Heston model [14] and the Lévy-driven stochastic volatility models put forward in [2]. Other SV models are based on time-changed Lévy processes as in [6]. All these examples are affine in the sense of [10]. Further instances of such affine stochastic volatility models are discussed in [33] and [18].

Stochastic volatility typically leads to an incomplete market, i.e. perfect hedging strategies do not exist for many contingent claims. As a way out one may try to minimize the expected squared hedging error

\[ E \left[ (v + \vartheta \cdot S_T - H)^2 \right] \]

∗Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Christian-Albrechts-Platz 4, 24098 Kiel, Germany, (e-mail: kallsen@math.uni-kiel.de).
†MEAG AssetManagement GmbH, Abteilung Risikocontrolling, Oskar-von-Miller-Ring 18, 80333 München, Germany, (e-mail: apauwels@meag.com).
over all initial endowments \( v \in \mathbb{R} \) and all admissible hedging strategies \( \vartheta \). Here, \( H \) denotes the discounted payoff at time \( T \) of a European-style contingent claim and \( S \) the discounted price process of the underlying stock. The stochastic integral \( \vartheta \cdot S \) stands for the gains from dynamic trading in the stock according to strategy \( \vartheta \).

This problem and in particular its general structure have been extensively studied in the literature, cf. e.g. [29, 35, 1, 8] and in the references therein. If \( S \) is a martingale, the solution is determined in [11] based on the Galtchouk-Kunita-Watanabe decomposition. To be more specific, let

\[
V_t := E[H|\mathcal{F}_t]
\]

denote the martingale generated by \( H \). Then \( v^* := V_0 \) is the optimal initial endowment in the above hedging problem and the optimal hedging strategy can be written as

\[
\vartheta_t^* = \frac{d\langle V, S \rangle_t}{d\langle S, S \rangle_t}.
\] (1.1)

The corresponding hedging error equals

\[
E\left[ (v^* + \vartheta^* \cdot S_T - H)^2 \right] = E\left[ \langle V, V \rangle_T - \langle \vartheta^* \cdot S, \vartheta^* \cdot S \rangle_T \right].
\] (1.2)

Alternative representations of the solution are provided in [5] using the carré du champ operator and in [4] based on Malliavin derivatives. All these general characterisations do not immediately lead to numerical results in concrete models. E.g., it may not be obvious how to evaluate (1.1) and (1.2) because analytical expressions for \( V \) are typically not available.

Numerical approaches are discussed e.g. in [13, 9, 15]. The first reference uses PDE methods for specific continuous stochastic volatility models. [9] and similarly [16] consider a SV model involving jumps. A partial integro-differential equations is solved by finite-difference schemes in order to obtain the process \( V \) above. The hedging error is computed by Monte-Carlo simulation. This approach is applied to exotic contingent claims and it allows for options as hedging instruments.

In this paper we study variance-optimal hedging in a general affine stochastic volatility model. The objective is to determine semi-explicit expressions for the optimal hedging strategy and the hedging error which can be evaluated without implementing involved numerical schemes or computer-intensive Monte-Carlo simulations. They are obtained with the help of integral transform techniques which are used widely in option pricing (cf. e.g. [7] and [31]) and in [15] for the mean-variance hedging problem without stochastic volatility.

We focus on the case that \( S \) is a martingale. Firstly, this allows to cover a broader class of volatility structures than without this restriction, e.g. those involving a leverage term. Secondly, we believe that the excess drift of asset prices is of secondary importance for the hedging problem. Finally, quadratic hedging appears as an auxiliary problem in a first-order approximation to utility-based derivative pricing and hedging, cf. [24, 25, 26, 3, 19, 21]. Here the variance-optimal hedge must be determined under some equivalent martingale measure, i.e. \( S \) is by default a martingale under the relevant measure. For a treatment of the more involved non-martingale case we refer the reader to [23], which generalizes some of
the present results. In such context a measure change to the – generally signed – variance-optimal martingale measure plays im- or explicitly a major role. However, the setup in [23] allows for leverage only in very special cases. Moreover, it is written on a less rigorous mathematical level.

The structure of the paper is as follows. In Section 2 we introduce the general affine stochastic volatility model. Subsequently, we discuss an integral representations of the contingent claim that is to be hedged. Section 4 contains the solution to the hedging problem in this setup. For numerical results in concrete parametric models we refer to the forthcoming paper [20]. Proofs of the main results are to be found in the final section.

Unexplained notation is used as in [17]. Superscripts refer generally to components of a vector or vector-valued process rather than powers. The few exceptions should be obvious from the context. As a key tool we need the notion of semimartingale characteristics $(B, C, \nu)$. For a summary of important results we refer to the appendix.

$$h = (h_1, h_2)$$ generally denotes a componentwise truncation function on $\mathbb{R}^2$, i.e.

$$h(x_1, x_2) = (\hat{h}(x_1), \hat{h}(x_2)),$$

where $\hat{h} : \mathbb{R} \to \mathbb{R}$ is a one-dimensional truncation function, which can e.g. be chosen of the form

$$\hat{h}(x_k) = \begin{cases} 0 & \text{if } x_k = 0, \\ (1 \wedge |x_k|) \frac{x_k}{|x_k|} & \text{otherwise.} \end{cases} \quad (1.3)$$

All characteristics and Lévy-Khintchine triplets on $\mathbb{R}$ resp. $\mathbb{R}^2$ are expressed relative to truncation functions $\hat{h}$ resp. $h$. For ease of exposition we occasionally use the particular function (1.3) in proofs but the results hold for arbitrary $\hat{h}$.

## 2 The general affine stochastic volatility model

We denote the discounted price process of a univariate stock by $S = S_0 \exp(Z)$, i.e. $Z$ stands for logarithmic returns. Moreover we consider a positive activity process $y$ leading to randomly changing volatility in our general setup. We assume that the bivariate process $(y, Z)$ is an affine semimartingale in the sense of [10]. More specifically, we suppose that the characteristics $(B^{y,Z}, C^{y,Z}, \nu^{y,Z})$ of the $\mathbb{R}^+ \times \mathbb{R}$-valued semimartingale $(y, Z)$ are of the form

$$B^{y,Z}_t = \int_0^t (\beta_{(0)} + \beta_{(1)} y_s^-) \, ds,$$

$$C^{y,Z}_t = \int_0^t (\gamma_{(0)} + \gamma_{(1)} y_s^-) \, ds,$$

$$\nu^{y,Z}([0, t] \times G) = \int_0^t (\varphi_{(0)}(G) + \varphi_{(1)}(G)y_s^-) \, ds \quad (2.2)$$

for all $G \in \mathcal{B}^2$ and $t \in [0, T]$, where $(\beta_j, \gamma_j, \varphi_j), j = 0, 1$, are Lévy-Khintchine triplets on $\mathbb{R}^2$ which are admissible in the sense of [10] Definition 2.6 or [18] Definition 3.1, i.e.
\( \beta(j) \in \mathbb{R}^2, \gamma(j) \) is a symmetric, non-negative matrix in \( \mathbb{R}^{2 \times 2} \), and \( \varphi(j) \) is a \( \sigma \)-finite measure on \( \mathbb{R}^2 \setminus \{0\} \) satisfying \( \int (1 \wedge |x|^2) \varphi(j)(dx) < \infty \),

- \( \gamma_{1,0} = \gamma_{1,1} = \gamma_{2,0} = 0 \),
- \( \varphi_{(0)}((\mathbb{R}_+ \times \mathbb{R})^C) = \varphi_{(1)}((\mathbb{R}_+ \times \mathbb{R})^C) = 0 \), i.e. \( \varphi_{(0)}, \varphi_{(1)} \) are actually Lévy measures on \( \mathbb{R}_+ \times \mathbb{R} \),
- \( \int h_1(x) \varphi_{(0)}(dx) < \infty \) and \( \beta_{(0)}^1 - \int h_1(x) \varphi_{(0)}(dx) \geq 0 \),
- \( \int x_1 \varphi_{(1)}(dx) < \infty \). This additional condition prevents explosion in finite time, cf. [10, Lemma 9.2].

We refer to this setup as general affine stochastic volatility model. We provide a few examples of popular affine SV models in the literature.

**Example 2.1** The Heston [14] model can be written as

\[
\begin{align*}
    dZ_t &= (\mu + \delta y_t) dt + \sqrt{y_t} dW^1_t, \\
    dy_t &= (-\frac{\lambda}{\delta} + \sigma^2 \sqrt{y_t}) dt + \sigma^2 \sqrt{y_t} dW^2_t,
\end{align*}
\]

where \( \kappa \geq 0, \mu, \delta, \lambda, \sigma, \) are constants and \( W^1, W^2 \) Wiener processes with constant correlation \( \rho \). The bivariate process \((y, Z)\) is affine in the sense of (2.1–2.2) with

\[
\begin{align*}
    (\beta_{(0)}, \gamma_{(0)}, \varphi_{(0)}) &= \left( \begin{pmatrix} \kappa \\ \mu \end{pmatrix}, 0, 0 \right), \\
    (\beta_{(1)}, \gamma_{(1)}, \varphi_{(1)}) &= \left( \begin{pmatrix} -\lambda \\ \delta \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho \sigma \\ \rho \sigma & 1 \end{pmatrix}, 0 \right).
\end{align*}
\]

**Example 2.2** The so-called BNS model of [2] is of the form

\[
\begin{align*}
    dZ_t &= (\mu + \delta y_t) dt + \sqrt{y_t} dW_t + \rho dz_t, \\
    dy_t &= -\lambda y_t dt + dz_t,
\end{align*}
\]

where \( \mu, \delta, \lambda, \rho \) are constants and \( W \) is a Wiener process. \( z \) denotes a subordinator, i.e. an increasing Lévy process whose Lévy-Khintchine triplet we write as \((b^z, 0, F^z)\). The process \((y, Z)\) is affine in the sense of (2.1–2.2) with

\[
\begin{align*}
    \beta_{(0)} &= \begin{pmatrix} \mu + gb^z + \int_0^\infty \tilde{h}(g x) - \tilde{h}(x) F^z(dx) \\ g^2 \int_0^\infty \tilde{h}(g x) F^z(dx) \end{pmatrix}, & \gamma_{(0)} &= 0, \\
    \varphi_{(0)}(G) &= \int_0^\infty 1_G(x, g x) F^z(dx) \quad \forall G \in \mathcal{B}^z,
\end{align*}
\]

and

\[
\begin{align*}
    (\beta_{(1)}, \gamma_{(1)}, \varphi_{(1)}) &= \left( \begin{pmatrix} -\lambda \\ \delta \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, 0 \right),
\end{align*}
\]
In order to obtain more realistic autocorrelation patterns in volatility, Barndorff-Nielsen and Shephard replace the Lévy-driven Ornstein-Uhlenbeck process \( y_t \) by a linear combination of such processes. This extension has a multivariate affine structure according to [18, Section 4.3]. It can be treated along similar lines as the simpler BNS model above. We stick to the bivariate case in this paper in order not to confuse the reader with heavy notation.

Example 2.3 [6] considers stochastic volatility models which are based on time-changed Lévy processes, namely

\[
\begin{align*}
Z_t &= \mu t + X_{Y_t} + \varrho z_t, \\
dY_t &= y_t dt, \\
dz_t &= \lambda y_t dt + dz_t,
\end{align*}
\]

where \( \mu, \varrho, \lambda \) are constants and \( z, X \) denote independent Lévy processes with triplets \((b^z, 0, F^z)\) and \((b^X, c^X, F^X)\), respectively. \( z \) is supposed to be increasing. The model is of the form (2.1–2.2) with

\[
\begin{align*}
\beta(0) &= \left( \begin{array}{c} \mu + \varrho b^z + \int_0^\infty (\tilde{h}(\varrho x) - \varrho \tilde{h}(x))F^z(dx) \\
\gamma(0) &= 0, \\
\varphi(0)(G) &= \int_0^\infty 1_{G}(x, \varrho x)F^z(dx) \\
\forall G \in B^2,
\end{array} \right), \\
\beta(1) &= \left( \begin{array}{c} -\lambda b^X \\
\gamma(1) &= \left( \begin{array}{c} 0 \\
0 0 c^X \\
\varphi(1)(G) &= \int_\mathbb{R} 1_{G}(0, x)F^X(dx) \\
\forall G \in B^2,
\end{array} \right),
\right),
\end{align*}
\]

cf. [18, Section 4.4]. If we choose \( X \) as a Brownian motion with drift, we obtain the dynamics of the BNS model above.

Another possible choice of the time change rate \( y \) is a square-root process:

\[
\begin{align*}
Z_t &= \mu t + X_{Y_t} + \varrho (y_t - y_0), \\
dY_t &= y_t dt, \\
dz_t &= (\kappa - \lambda y_t) dt + \sigma \sqrt{y_t} dW_t,
\end{align*}
\]

where \( \kappa \geq 0, \mu, \varrho, \lambda, \sigma \) are constants, \( W \) denotes a Wiener process and \( X \) an independent Lévy process with triplet \((b^X, c^X, F^X)\). This model is of the form (2.1–2.2) with

\[
\begin{align*}
(\beta(0), \gamma(0), \varphi(0)) &= \left( \begin{array}{c} \kappa \\
\mu + \varrho \kappa, \\
0 0 \\
\end{array} \right), \\
\beta(1) &= \left( \begin{array}{c} b^X - \varrho \lambda \\
\gamma(1) &= \left( \begin{array}{c} \sigma^2 \\
\varrho \sigma^2 \\
\varrho^2 \sigma^2 + c^X \\
\varphi(1)(G) &= \int_\mathbb{R} 1_{G}(0, x)F^X(dx) \\
\forall G \in B^2,
\end{array} \right),
\right),
\end{align*}
\]

cf. [18, Section 4.4]. If we choose \( X \) as a Brownian motion with drift, we recover the dynamics of the Heston model — up to a rescaling of the volatility process \( y \).
The Lévy-Khintchine triplets \((\beta_j, \gamma_j, \varphi_j), j = 0, 1\) can be associated to corresponding Lévy exponents

\[
\psi_j(u) := u^\top \beta_j + \frac{1}{2} u^\top \gamma_j u + \int \left( e^{u^\top x} - 1 - u^\top h(x) \right) \varphi_j(dx). \tag{2.3}
\]

These functions \(\psi_j, j = 0, 1\) are defined on

\[
U_j := \left\{ u \in \mathbb{C}^2 : \int_{\{|x| \geq 1\}} \exp(\text{Re}(u^\top x)) \varphi_j(dx) < \infty \right\}.
\]

**Assumption 2.4** In order for \(S\) to be a locally square-integrable martingale, we assume that \((0, 2) \in U_0 \cap U_1\) and

\[
\psi_0(0, 1) = \psi_1(0, 1) = 0. \tag{2.4}
\]

Moreover, we suppose that

\[
\psi_0(0, 2) \neq 0 \quad \text{or} \quad \psi_1(0, 2) \neq 0.
\]

in order to avoid the degenerate case \(Z = 0\).

**Proposition 2.5** **Assumption 2.4** implies \(S \in \mathcal{H}^2_{\text{loc}}\). 

**Proof.** The differential characteristics \((b^Z, c^Z, F^Z)\) of \(Z\) can be easily calculated from (2.1–2.2) cf. Proposition A.2 or A.3. They are of the form

\[
b^Z_t = \beta^2_{(0)} + \beta^2_{(1)} y_t - \int (\tilde{h}(e^x - 1) - \tilde{h}(x)) F^Z_t(dx),
\]

\[
c^Z_t = \gamma^2_{(0)} + \gamma^2_{(1)} y_t - \int 1_G(x_2) \varphi_{(0)}(dx) + \int 1_G(x_2) \varphi_{(1)}(dx) y_t,
\]

\[
F^Z_t(G) = \int 1_G(x_2) \varphi_{(0)}(dx) + \int 1_G(x_2) \varphi_{(1)}(dx) y_t
\]

for all \(G \in \mathcal{B}\). Using Propositions A.2 and A.3 we obtain for the stochastic logarithm \(X := \mathcal{L}(S) = \frac{1}{S} \cdot S\):

\[
b^X_t = b^Z_t + \frac{1}{2} c^Z_t + \int \left( \tilde{h}(e^x) - 1 \right) F^Z_t(dx),
\]

\[
F^X_t(G) = \int 1_G(e^x - 1) F^Z_t(dx), \quad \forall G \in \mathcal{B}
\]

for the differential characteristics \((b^X, c^X, F^X)\) of \(X\). In particular, we have

\[
\int x^2 F^X_t(dx) = \int (e^{x^2} - 1)^2 \varphi_{(0)}(dx) + \left( \int (e^{x^2} - 1)^2 \varphi_{(1)}(dx) \right) y_t - .
\]

In view of Assumption 2.4 and [17] II.2.29] we have that \(X\) is a locally square-integrable semimartingale. Moreover, (2.4) implies

\[
b^X_t + \int (x - \tilde{h}(x)) F^X_t(dx) = \psi_{(0)}(0, 1) + \psi_{(1)}(0, 1) y_t = 0,
\]

in order to avoid the degenerate case \(Z = 0\). 

6
which implies that $X$ is actually a local martingale. Since $\mathcal{E}(X)_-$ is locally bounded, we have that $S = S_0 \mathcal{E}(X) = S_0 (1 + \mathcal{E}(X)_- \cdot X)$ is a locally square-integrable martingale as well. □

Due to the results of [10] the characteristic function of the bivariate affine process $(y, Z)$ is known explicitly.

**Proposition 2.6** The conditional characteristic function of $(y, Z)$ is of the form

$$E \left[ e^{u_1 y_t + u_2 Z_t} \; | \; \mathcal{F}_t \right] = \exp \left( \Psi_0(s, u_1, u_2) + \Psi_1(s, u_1, u_2) y_t + u_2 Z_t \right),$$

$u \in \mathbb{C}_- \times i\mathbb{R}$, where $\Psi_1 : \mathbb{R}_+ \times \mathbb{C}_- \times i\mathbb{R} \to \mathbb{C}_-$ solves the initial value problem

$$\frac{\partial}{\partial t} \Psi_1(t, u_1, u_2) = \psi_1(\Psi_1(t, u_1, u_2), u_2), \quad \Psi_1(0, u_1, u_2) = u_1 \quad (2.5)$$

and $\Psi_0 : \mathbb{R}_+ \times \mathbb{C}_- \times i\mathbb{R} \to \mathbb{C}$ is given by

$$\Psi_0(t, u_1, u_2) = \int_0^t \psi_0(\Psi_1(s, u_1, u_2), u_2) ds. \quad (2.6)$$

Here $\mathbb{C}_-$ is defined as $\mathbb{C}_- := \{ z \in \mathbb{C} : \text{Re}(z) \leq 0 \}$.

**Proof.** This is a special case of [18, Theorem 3.2]. □

### 3 European options

The hedging problem cannot be solved in closed form even for geometric Lévy processes without stochastic volatility. In order to obtain at least semiclosed solutions, we consider European-style claims whose discounted payoff at time $T$ is of the form $H = f(S_T)$. More specifically, we assume that the function $f : (0, \infty) \to \mathbb{R}$ can be written in integral form

$$f(s) = \int s^2 \Pi(dz)$$

with some finite complex measure $\Pi$ on a strip $S_f := \{ z \in \mathbb{C} : R' \leq \text{Re}(z) \leq R \}$, where $R', R \in \mathbb{R}$. The measure $\Pi$ is supposed to be symmetric in the sense that $\Pi(A) = \overline{\Pi(A)}$ for $A \in \mathcal{B}(\mathbb{C})$ and $\overline{A} := \{ \overline{z} \in \mathbb{C} : z \in A \}$. In most cases we can choose $R' = R$ and the measure $\Pi$ has a density, cf. [15]. E.g. we have

$$(s - K)^+ = \frac{1}{2\pi i} \int_{R-i\infty}^{R+i\infty} s \frac{K^{1-z}}{z(1-z)} dz$$

with $R > 1$ for a European call with strike $K$, which means that

$$\Pi(dz) = \frac{1}{2\pi i} \frac{K^{1-z}}{z(1-z)} dz$$

7
on $S_f := R + i\mathbb{R}$. The put $f(s) = (K - s)^+$ corresponds to the same formula but with $R < 0$. The integral representation of many other payoffs can be found in [15].

For the derivation of formulas in the next section some moment conditions are needed. We phrase them here in terms of analytical properties of the characteristic exponents in Proposition 2.6

**Assumption 3.1** We assume that the following conditions hold.

1. For some $\varepsilon > 0$ the mappings $(u_1, u_2) \mapsto \Psi_0(t, u_1, u_2), \Psi_1(t, u_1, u_2)$ have an analytic extension on
   \[ S := \{ u \in \mathbb{C}^2 : (\text{Re}(u_1), \text{Re}(u_2)) \in V_{\varepsilon}(0) \} \]
   for all $t \in [0, T]$, where
   \[ M_0 := \sup \{ 2\Psi_1(t, 0, r) : r \in [R' \land 0, R \lor 0], t \in [0, T] \} \]
   and
   \[ V_{\varepsilon}(a) := (\infty, (M_0 \lor 0) + \varepsilon) \times ((2R' \land 0) - \varepsilon, (2R \lor a) + \varepsilon) \]
   for $a \in \mathbb{R}_+$. These extensions are again denoted by $\Psi_0$ resp. $\Psi_1$.

2. The mappings $t \mapsto \Psi_0(t, u_1, u_2), t \mapsto \Psi_1(t, u_1, u_2)$ are continuous on $[0, T]$ for any $(u_1, u_2) \in S$.

3. $V_{\varepsilon}(2) \subset U_0 \cap U_1$. This is satisfied if the mappings $i\mathbb{R} \to \mathbb{C}$, $u \mapsto \psi_j(u)$ for $j = 0, 1$ have analytic extensions to $V_{\varepsilon}(2)$, in which case representation (2.3) holds for this extension. Note that this implies the integrability condition $(0, 2) \in U_0 \cap U_1$ in Assumption 2.4

**Remark 3.2**

1. From [36, Theorem III.13.XI] it follows that the analytic extensions $\Psi_0, \Psi_1$ solve (2.5), (2.6) as well. Moreover, $\Psi_0, \Psi_1, D_2 \Psi_0, D_2 \Psi_1$ are continuous on $[0, T] \times S$. For later use, we also note that the mapping $t \mapsto \Psi_1(t, 0, z)$ is twice continuously differentiable on $[0, T]$ for all $z \in S_f$.

2. From [10, Theorem 2.16(ii)] it follows with Assumption 3.1(1) that
   \[ E[e^{u_1y_{t,s} + u_2Z_{t,s}} \mid \mathcal{F}_t] = \exp(\Psi_0(s, u_1, u_2) + \Psi_1(s, u_1, u_2)y_t + u_2Z_t) \quad (3.1) \]
   for all $u \in S$. It is easy to see that $\Psi_j(t, u_1, u_2)$ is real-valued for $u \in \mathbb{R}_+^2 \cap S$, $j = 0, 1$.

3. By the first remark we have $\Psi_0(t, 0, z) = \int_0^t \psi_0(\Psi_1(s, 0, z), z)ds$ for $z \in S$. Equation (3.1) and Jensen’s inequality yield $\text{Re} (\Psi_1(t, 0, z)) \leq \Psi_1(t, 0, \text{Re}(z))$ and hence $(\Psi_1(t, 0, z), z) \in S$ for all $t \in [0, T]$ and all $z \in S_f$.

**Lemma 3.3** Under Assumption 3.1(1) $S_f^z$ is a square-integrable random variable for any $z \in S_f$. The same is true for $H$.  

8
PROOF. In view of Remark 3.2(2) we have
\[ E[|S_T^2|^2] \leq \exp \left( 2(|R'| \vee |R|) \log S_0 \right) E[e^{2R'T} + e^{2R'T}] < \infty. \]
Applying Hölder’s inequality we have
\[ E[|H|^2] \leq E \left[ \left( \int |S_T^2||\Pi|(dz) \right)^2 \right] \leq |\Pi|(S_f) E \left[ \int |S_T^2|^2|\Pi|(dz) \right] \]
\[ \leq |\Pi|(S_f)^2 \exp \left( 2(|R'| \vee |R|) \log S_0 \right) E[e^{2R'T} + e^{2R'T}] < \infty. \]
Here $|\Pi|$ indicates the total variation measure of $\Pi$ in the sense of [32, Section 6.1].

A key role in the hedging problem is played by the square-integrable martingale $V$ generated by $H$, i.e.
\[ V_t := E[H|\mathcal{F}_t], \quad t \in [0, T]. \]
We call it option price process because it could be used as such without introducing arbitrage to the market. But note that we do not assume $H$ to be traded, let alone with price process $V$.

**Proposition 3.4** Under Assumption 3.1(1) we have
\[ V_t = \int V(z)\Pi(dz), \quad t \in [0, T], \]
where the square-integrable martingale $V(z)$ for $z \in S_f$ is defined as $V(z)_t := E[S_T^z|\mathcal{F}_t]$ for $t \in [0, T]$.

**Proof.** For all $t \in [0, T]$ and $z \in S_f$ we have the estimate
\[ \int E[|V(z)_t||\Pi|(dz) \leq \int E[|V(z)_T||\Pi|(dz) \]
\[ \leq \int E[e^{R'e(z)\log S_T}]|\Pi|(dz) \]
\[ \leq E[e^{R'\log S_T} + e^{R\log S_T}]|\Pi|(S_f) < \infty \]
from Remark 3.2(2) and the finiteness of $\Pi$. An application of Fubini’s theorem yields
\[ E\left[ \int V(z)_t\Pi(dz)1_C \right] = E\left[ \int E[S_T^z|\mathcal{F}_t]1_C\Pi(dz) \right] \]
\[ = E\left[ \int S_T^z\Pi(dz)1_C \right] \]
\[ = E[V_t1_C] \]
for all $C \in \mathcal{F}_t$. This implies the assertion. \qed

The processes $V(z)$ will be determined in Theorem 4.1 as a by-product. Via (3.2) this leads to an integral representation of the option price process which is of interest in itself. It extends similar formulas in [6, 33] to the more general class of processes considered in this paper.
We turn now to the hedging problem itself. We generally assume that Assumptions 2.4, 3.1 hold. We define the set of admissible trading strategies as

\[ \Theta := \{ \vartheta \text{ predictable process} : E[|\vartheta|^2 \cdot \langle S, S \rangle_T] < \infty \}. \]

We call an initial capital \( v^* \in \mathbb{R} \) and an admissible strategy \( \vartheta^* \) variance-optimal (hedge) if they minimize

\[ E \left[ (v + \vartheta^* \cdot S_T - H)^2 \right] \]

over all such pairs \((v, \vartheta) \in \mathbb{R} \times \Theta\). The residue

\[ J_0 := E \left[ (v^* + \vartheta^* \cdot S_T - H)^2 \right] \]

is referred to as minimal hedging error. The following characterization of the variance-optimal hedge constitutes the first main result of this paper.

**Theorem 4.1** The variance-optimal initial capital \( v^* \) and the variance-optimal hedging strategy \( \vartheta^* \) are given by

\[
\begin{align*}
    v^* &= \int V(z)_0 \Pi(dz), \\
    \vartheta^*_t &= \int \frac{V(z)_t \kappa_0(t, z) + \kappa_1(t, z)y_t - \delta_0 - \delta_1y_t}{S_t} \Pi(dz),
\end{align*}
\]

where the process \( V(z) \) satisfies

\[ V(z)_t = S^*_t \exp \left( \Psi_0(T - t, 0, z) + \Psi_1(T - t, 0, z)y_t \right), \quad z \in S_f. \]

\[ \delta_0, \delta_1 \in \mathbb{R} \text{ and the functions } \kappa_0, \kappa_1 : [0, T] \times S_f \to \mathbb{C} \text{ are defined as} \]

\[
\begin{align*}
    \kappa_j(t, z) &:= \psi_j(\Psi_1(T - t, 0, z), z + 1) - \psi_j(\Psi_1(T - t, 0, z), z), \\
    \delta_j &:= \psi_j(0, 2), \quad j = 0, 1.
\end{align*}
\]

As is well known, the variance-optimal hedging strategy \( \vartheta^* \) given by Equation (4.1) also yields the solution to

\[ \min_{\vartheta \in \Theta} E \left[ (\tilde{v} + \vartheta \cdot S_T - H)^2 \right], \]

where \( \tilde{v} \in \mathbb{R} \) denotes a given initial capital instead of the optimizer from the previous theorem. Our second main result concerns the hedging error. This quantity gives an idea of the remaining risk. The seller of the option may take it into account in order to decide what risk premium to charge for the claim.
Theorem 4.2 The minimal hedging error is given by

\[ J_0 = \begin{cases} \int \int \int_0^T J_1(t, z_1, z_2)dt \Pi(dz_1)\Pi(dz_2), & \text{if } \delta_0 \neq 0, \delta_1 \neq 0, \\ \int \int \int_0^T J_2(t, z_1, z_2)dt \Pi(dz_1)\Pi(dz_2), & \text{if } \delta_0 = 0, \\ \int \int \int_0^T J_3(t, z_1, z_2)dt \Pi(dz_1)\Pi(dz_2), & \text{if } \delta_1 = 0. \end{cases} \]

The integrals over \( S_f \) have to be understood in the sense of the Cauchy principal value (cf. the following remark). The integrands \( J_k : [0, T] \times S_f^2 \rightarrow \mathbb{C}, k = 1, 2, 3 \) in these expressions are defined as

\[ J_1(t, z_1, z_2) = S_0^{z_1+z_2} e^{\xi_0} \left( \exp(\Psi_0(t, \xi_1, z_1 + z_2) + \Psi_1(t, \xi_1, z_1 + z_2)y_0) + \frac{\eta_2}{\delta_1} (D_2\Psi_0(t, \xi_1, z_1 + z_2) + \eta_1\delta_1 - \eta_2\delta_0) \right) \times \int_0^1 \left( \frac{\delta_1}{\delta_0} + \xi_1 s \right) e^{\frac{\delta_1}{\delta_0} \xi_1 s + \Psi_0 \left( t, z_1 + z_2, \delta_0 \left( \frac{\eta_1}{\delta_0} \log(s) + \xi_1 s + z_1 + z_2 \right) \right) + \Psi_1 \left( t, z_1 + z_2, \delta_0 \left( \frac{\eta_1}{\delta_0} \log(s) + \xi_1 s + z_1 + z_2 \right) \right) dt, \]

\[ J_2(t, z_1, z_2) = \frac{S_0^{z_1+z_2} e^{\xi_0}}{\delta_1} \exp(\Psi_0(t, \xi_1, z_1 + z_2) + \Psi_1(t, \xi_1, z_1 + z_2)y_0) \times \left( \eta_1 + (D_2\Psi_0(t, \xi_1, z_1 + z_2) + D_2\Psi_1(t, \xi_1, z_1 + z_2)y_0) \eta_2 \right), \]

\[ J_3(t, z_1, z_2) = \frac{S_0^{z_1+z_2} e^{\xi_0}}{\delta_0} \exp \left( (\psi_0(0, z_1) + \psi_0(0, z_2))(T - t) + \psi_0(0, z_1 + z_2)t \right). \]

The constants \( \delta_0, \delta_1 \) are defined as in Equation (4.4). The remaining variables are specified as follows:

\[ \alpha_j = \alpha_j(t, z_1, z_2) \]
\[ := \psi_j(\xi_1(t, z_1, z_2), z_1 + z_2) - \psi_j(\Psi_1(T - t, 0, z_1), z_1) - \psi_j(\Psi_1(T - t, 0, z_2), z_2), \]

\[ \eta_0 = \eta_0(t, z_1, z_2) := \delta_0\alpha_0(t, z_1, z_2) - \kappa_0(t, z_1)\kappa_0(t, z_2), \]

\[ \eta_1 = \eta_1(t, z_1, z_2) := \delta_0\alpha_1(t, z_1, z_2) + \delta_1\alpha_0(t, z_1, z_2) - \kappa_1(t, z_1)\kappa_0(t, z_2) - \kappa_1(t, z_2)\kappa_0(t, z_1), \]

\[ \eta_2 = \eta_2(t, z_1, z_2) := \delta_1\alpha_1(t, z_1, z_2) - \kappa_1(t, z_1)\kappa_1(t, z_2), \]

\[ \xi_j = \xi_j(t, z_1, z_2) := \Psi_j(T - t, 0, z_1) + \Psi_j(T - t, 0, z_2), \quad j = 0, 1, \]

with \( \kappa_0, \kappa_1 \) from (4.3). For ease of notation we dropped the arguments of some functions in the formulae above. The mappings \( \eta_0, \eta_1, \eta_2, \alpha_0, \alpha_1, \xi_0, \xi_1 \) are defined on \([0, T] \times S_f^2\).

Remark 4.3 1. The integrals in the previous theorem are to be understood in the sense that

\[ J_0 = \lim_{c \to \infty} \int_{S_f} \int_{S_f} \int_0^T J_k(t, z_1, z_2)dt \Pi(dz_1)\Pi(dz_2), \]
where
\[ S^c_f := \{ z \in \mathbb{C} : R' \leq \text{Re}(z) \leq R, |\text{Im}(z)| \leq c \}. \]

It is not obvious whether integrability holds on all of \( S_f \).

2. Condition \( \delta_1 = 0 \) actually implies that \( Z \) is a Lévy process. The integral of \( J_3 \) relative to \( t \) can easily be evaluated in closed form, which is done in [15, Theorem 3.2].

3. The minimal hedging error in (4.5) for fixed initial endowment \( \tilde{v} \) instead of the optimal \( v \) equals
\[ J_0 + (\tilde{v} - v)^2. \]

For concrete models and numerical results we refer the reader to the companion paper [20].

5 Proof of the main results

This section is devoted to the proofs of Theorems 4.1 and 4.2. We start by analyzing the martingales \( V(z) \) in Proposition 3.4.

**Lemma 5.1** \( V(z) \) in Proposition 3.4 is of the form (4.2).

**Proof.** This follows immediately from (3.1). □

The Galtchouk-Kunita-Watanabe (GKW) decomposition of \( V(z) \) is determined in the following lemma.

**Lemma 5.2** Fix \( z \in S_f \). If we set
\[ \vartheta(z)_t := \frac{V(z)_t - \kappa_0(t,z) + \kappa_1(t,z)y_t}{\delta_0 + \delta_1y_t}, \]
\[ L(z)_t := V(z)_t - \vartheta(z)_t \cdot S, \]

then \( \vartheta(z) \in \Theta \) (here liberally extended to complex-valued processes), \( L(z) \) is a square-integrable martingale and \( LS \) a local martingale. \( \kappa_j(t,z) \) and \( \delta_j, j = 0, 1 \) are defined as in (4.3) and (4.4), respectively.

**Proof.** The denominator in (5.1) is positive because the constants \( \delta_0, \delta_1 \in \mathbb{R}_+ \) do not both vanish (cf. Assumption 2.4). The differential characteristics \((b^S, c^S, F^S)\) of \( S \) can be obtained from (2.1) and (2.2). We have
\[ e_t^S = S_t^{2} \left( \gamma^{(2)}_{(0)} + \gamma^{(2,2)}_{(1)}y_t \right) \quad \text{and} \]
\[ F_t^S(G) = \int 1_G(S_t - (e^{x^2} - 1))\varphi_{(0)}(dx) + \int 1_G(S_t - (e^{x^2} - 1))\varphi_{(1)}(dx)y_t \]
for all $G \in \mathcal{B}$. By [17] II.2.29b this implies
\[
\langle S, S \rangle_t = \int_0^t \left( c_s^S + \int x^2 F_s^S (dx) \right) ds = \int_0^t S_{s-}^2 (\delta_0 + \delta_1 y_{s-}) ds,
\] (5.2)
where the second equality follows from
\[
\gamma_{(j)}^{2,2} + \int (e^{x^2} - 1)^2 \varphi_{(j)} (dx) = \psi_j (0, 2) - 2 \psi_j (0, 1) = \psi_j (0, 2) = \delta_j.
\]
Another application of Proposition A.3 allows to compute the differential characteristics
\[
(b_{(V(z), S)}, c_{(V(z), S)}, F_{(V(z), S)})
\]
\[
\text{of } (V(z), S) = g(I, y, Z), \text{ where } I_t = t \text{ denotes the identity process and}
\]
\[
g(t, x_1, x_2) := \left( S_0^\delta \exp \left( \Psi_0 (T - t, 0, z) + \Psi_1 (T - t, 0, z) x_1 + z x_2 \right) \right).
\]
Again using [17] II.2.29b this leads to
\[
\langle V(z), S \rangle_t = \int_0^t \left( (c_{(V(z), S)})^2 + \int x_1 x_2 F_{(V(z), S)} (dx) (x_1, x_2) \right) ds
\]
\[
= \int_0^t V(z)_{s-} S_{s-} (\kappa_0 (s, z) + \kappa_1 (s, z) y_{s-}) ds.
\] (5.3)
We conclude that
\[
\vartheta (z) \cdot \langle S, S \rangle = \langle V(z), S \rangle.
\] (5.4)
Hence $\langle L, S \rangle = 0$, which implies that $LS$ is a local martingale. (5.4) also yields
\[
\langle L(z), L(z) \rangle + |\vartheta(z)|^2 \cdot \langle S, S \rangle = \langle L(z), L(z) \rangle + (\vartheta(z) \overline{\vartheta(z)}) \cdot \langle S, S \rangle = \langle V(z), V(z) \rangle.
\]
Since $V(z)$ is a square-integrable martingale, we have $E(\langle V(z), V(z) \rangle_T) < \infty$, which implies $\vartheta(z) \in \Theta$ and $L(z) \in \mathcal{H}^2$.

For later use we need a technical result.

**Lemma 5.3** Let
\[
\tau_n := \inf \left\{ t > 0 : y_t \notin [1/n, n] \text{ or } S_t \notin [1/n, n] \right\} \wedge T.
\]
For any $n \in \mathbb{N}$ there exists a constant $c(n) < \infty$ such that
\[
\int_0^T E \left[ 1_{[0, \tau_n]} |V(z)|^2 (\alpha_0 (t, z, \overline{z}) + \alpha_1 (t, z, \overline{z}) y_t) \right] dt \leq c(n)
\] (5.5)
holds for all $z \in S_f$. The functions $\alpha_0$ and $\alpha_1$ are defined as in Theorem 4.2.
Using (2.5, 2.6) one obtains

\[ K_j(t, z, \overline{z}) = \psi_j(2 \text{Re}(\Psi_j(T-t, 0, z)), 2 \text{Re}(z)) - 2 \text{Re}(\psi_j(\Psi_j(T-t, 0, z), z)) \]

\[ \geq \left( \text{Re}(\Psi_j(T-t, 0, z)), \text{Re}(z) \right) \gamma_j \left( \text{Re}(\Psi_j(T-t, 0, z)), \text{Re}(z) \right)^T \]

\[ + \left( \text{Im}(\Psi_j(T-t, 0, z)), \text{Im}(z) \right) \gamma_j \left( \text{Im}(\Psi_j(T-t, 0, z)), \text{Im}(z) \right)^T \]

\[ + \int (e^{2 \text{Re}(\Psi_j(T-t, 0, z))x_1 + \text{Re}(z)x_2} - 1)^2 \varphi_j(dx) \geq 0 \]

for \( j = 0, 1 \) because \( \gamma_j \) is positive semi-definite. Hence

\[ K_1(z) := \int_0^T E \left[ 1_{[0,n]} e^{2 \text{Re}(\Psi_0(T-t, 0, z))} + 2 \text{Re}(\Psi_1(T-t, 0, z)) y_t S_t \right] \times \]

\[ \times (\alpha_0(t, z, \overline{z}) + \alpha_1(t, z, \overline{z}) y_t) dt \]

\[ \leq \int_0^T e^{2 \text{Re}(\Psi_0(T-t, 0, z)) + \frac{2}{n} \text{Re}(\Psi_1(T-t, 0, z))} c \left( \psi_0(2 \text{Re}(\Psi_j(T-t, 0, z)), 2 \text{Re}(z)) \right) \]

\[ - 2 \text{Re}(\psi_0(\Psi_1(T-t, 0, z), z)) + n \left( \psi_1(2 \text{Re}(\Psi_1(T-t, 0, z)), 2 \text{Re}(z)) \right) \]

\[ - 2 \text{Re}(\psi_1(\Psi_1(T-t, 0, z), z)) \) \right) \right) dt. \]

Using (2.5, 2.6) one obtains

\[ K_1(z) \leq c \int_0^T \left| e^{2 \text{Re}(\Psi_0(T-t, 0, z)) + \frac{2}{n} \text{Re}(\Psi_1(T-t, 0, z))} \psi_0(2 \text{Re}(\Psi_j(T-t, 0, z)), 2 \text{Re}(z)) \right| dt \]

\[ + c \int_0^T \left| e^{2 \text{Re}(\Psi_0(T-t, 0, z)) + \frac{2}{n} \text{Re}(\Psi_1(T-t, 0, z))} \psi_1(2 \text{Re}(\Psi_j(T-t, 0, z)), 2 \text{Re}(z)) \right| dt \]

\[ + c \left| \int_0^T \frac{\partial}{\partial t} \left( 2 \text{Re}(\Psi_0(T-t, 0, z)) + \frac{2}{n} \text{Re}(\Psi_1(T-t, 0, z)) \right) \right| \times \]

\[ e^{2 \text{Re}(\Psi_0(T-t, 0, z)) + \frac{2}{n} \text{Re}(\Psi_1(T-t, 0, z))} dt \]

\[ = c \left( K_2(z) + K_3(z) + |K_4(z)| \right). \]

And Jensen’s inequality yield \( 2 \text{Re}(\Psi_j(t, 0, z)) \leq \Psi_j(t, 0, 2 \text{Re}(z)) \) for all \( (t, z) \in [0,T] \times S_f \). By continuity of the mappings \( \Psi_0 \) and \( \Psi_1 \) (cf. Assumption 3.1) we have

\[ |K_4(z)| = \left| \exp \left( \frac{2}{n} \text{Re}(z) \right) - \exp \left( 2 \text{Re}(\Psi_0(T, 0, z)) + \frac{2}{n} \text{Re}(\Psi_1(T, 0, z)) \right) \right| \]

\[ \leq \exp \left( \frac{2}{n} \text{Re}(z) \right) + \exp \left( \Psi_0(T, 0, 2 \text{Re}(z)) + \frac{1}{n} \Psi_1(T, 0, 2 \text{Re}(z)) \right) \leq c. \]
Similarly we obtain estimates
\[
g(t, z) := 2\text{Re}(\Psi_1(T - t, 0, z)) \leq \Psi_1(T - t, 0, 2\text{Re}(z)) \leq c,
\]
\[
2\text{Re}(\Psi_0(T - t, 0, z)) \leq \Psi_0(T - t, 0, 2\text{Re}(z)) \leq c.
\]

For \( j = 0, 1 \) we have
\[
\left| \frac{1}{n} e^{\frac{j}{n}g(t, z)\psi_j(g(t, z), 2\text{Re}(z))} \right| \leq \left[ \frac{2}{n} \left|\beta^2_{(j)} + g|\gamma^2_{(j)}|\right| \phi + \left|\beta_{(j)}^1 + 2g|\gamma_{(j)}^1|\right| |g_n(t, z)| + \frac{n}{2} |\psi_{(j)}^1||g_n(t, z)|^2 \right] e^{g_n(t, z)}
\]
\[+ \frac{1}{n} e^{g_n(t, z)} \int \left| e^{g(t, z)x_1 + 2\text{Re}(z)x_2} - 1 - g(t, z)\tilde{h}(x_1) - 2\text{Re}(z)\tilde{h}(x_2) \right| \varphi_{(j)}(dx),
\]
where \( g_n(t, z) := \frac{1}{n} g(t, z) \). Boundedness of \( g \) implies that we have
\[
\exp(g_n(t, z))|g_n(t, z)|^k \leq c
\]
for \( k = 0, 1, 2 \) and some constant \( c \). Furthermore, we have
\[
\frac{1}{n} e^{g_n(t, z)} \int \left| e^{g(t, z)x_1 + 2\text{Re}(z)x_2} - 1 - g(t, z)\tilde{h}(x_1) - 2\text{Re}(z)\tilde{h}(x_2) \right| \varphi_{(j)}(dx)
\]
\[
\leq \frac{1}{n} e^{g_n(t, z)} \left( \int_{\{|x| > 1\}} e^{M_0 x_1 + 2R x_2} \varphi_{(j)}(dx) + \int_{\{|x| > 1\}} e^{M_0 x_1 + 2R x_2} \varphi_{(j)}(dx) + (1 + 2g + |g(t, z)|) \varphi_{(j)}(\{|x| > 1\}) \right)
\]
\[+ \frac{1}{n} e^{g_n(t, z)} \int_{\{|x| \leq 1\}} \left| e^{g(t, z)x_1 + 2\text{Re}(z)x_2} - 1 - g(t, z)x_1 - 2\text{Re}(z)x_2 \right| \varphi_{(j)}(dx).
\]
Since \( \varphi_{(j)} \) is a Lévy measure, \( \varphi_{(j)}(\{|x| > 1\}) \) is finite. The first two integrals on the right-hand side are bounded in view of Assumption 3.1(3). Note that
\[
\left| e^{g(t, z)x_1 + 2\text{Re}(z)x_2} - 1 - g(t, z)x_1 - 2\text{Re}(z)x_2 \right| \leq c \left( x_1^2 + x_2^2 \right)
\]
for all \( |x| \leq 1 \). Since \( \int_{\{|x| \leq 1\}} (x_1^2 + x_2^2) \varphi_{(j)}(dx) < \infty \), we obtain
\[
\frac{1}{n} e^{g_n(t, z)} \int \left| e^{g(t, z)x_1 + 2\text{Re}(z)x_2} - 1 - g(t, z)\tilde{h}(x_1) - 2\text{Re}(z)\tilde{h}(x_2) \right| \varphi_{(j)}(dx) \leq c.
\]
Altogether, we conclude that
\[
\left| \frac{1}{n} e^{\frac{j}{n}g(t, z)\psi_j(g(t, z), 2\text{Re}(z))} \right| \leq c < \infty
\]
uniformly in \( (t, z) \in [0, T] \times \mathcal{S}_f, \ j = 0, 1 \). This in turn yields \( K_2(z) \leq c, K_3(z) \leq c \) and hence
\[
0 \leq K_1(z) \leq c < \infty
\] (5.6)
for all \( z \in S_f \). Analogously one shows that

\[
K_5(z) := \int_0^T E \left[ 1_{[0,\tau_n]} \left\{ \Re(\Psi_1(T-t,0,z)) > 0 \right\} 2 \text{Re}(\Psi_0(T-t,0,z)) + 2 \text{Re}(\Psi_1(T-t,0,z)) y_t - S_t^2 \right] dt \leq c.
\]

(5.7)

The assertion (5.5) follows now from (5.6) and (5.7) because

\[
|V(z)_t|^2 = S_t^{2 \text{Re}(z)} \exp \left( 2 \text{Re}(\Psi_0(T-t,0,z)) + 2 \text{Re}(\Psi_1(T-t,0,z)) y_t \right)
\]

due to Lemma 5.1.

**Corollary 5.4** Using the notation of Lemma 5.3 we have

\[
E \left[ \left( \int |\vartheta(z)|^2 |\Pi|(dz) \right) \cdot (S, S)_{\tau_n} \right] < \infty,
\]

(5.8)

\[
\int E[|L(z)|^2] |\Pi|(dz) < \infty.
\]

(5.9)

**PROOF.** By (5.1) (5.2) we have

\[
K := E \left[ \left( \int |\vartheta(z)|^2 |\Pi|(dz) \right) \cdot (S, S)_{\tau_n} \right] = E \left[ \int_0^{\tau_n} \int |V(z)_t|^2 \frac{|\kappa_0(t,z) + \kappa_1(t,z)y_t|^2}{\delta_0 + \delta_1 y_t} |\Pi|(dz) dt \right].
\]

Similar to the derivation of Equation (5.3) we derive

\[
\langle V(z_1), V(z_2) \rangle_t = \int_0^t V(z_1)_s V(z_2)_s \left( \alpha_0(s,z_1,z_2) + \alpha_1(s,z_1,z_2)y_s \right) ds
\]

for \( z_1, z_2 \in S_f \). Equations (5.2) (5.3) imply

\[
\vartheta(z_2) \cdot \langle V(z_1), S \rangle_t = \vartheta(z_1) \cdot \langle S, V(z_2) \rangle_t = \left( \vartheta(z_1) \vartheta(z_2) \right) \cdot \langle S, S \rangle_t
\]

\[
= \int_0^t V(z_1)_s V(z_2)_s \frac{(\kappa_0(s,z_1) + \kappa_1(s,z_1)y_s)(\kappa_0(s,z_2) + \kappa_1(s,z_2)y_s)}{\delta_0 + \delta_1 y_s} ds.
\]

This leads to

\[
\langle L(z_1), L(z_2) \rangle_t = \int_0^t V(z_1)_s V(z_2)_s \left( \alpha_0(s,z_1,z_2) + \alpha_1(s,z_1,z_2)y_s \right)
\]

\[
- \frac{(\kappa_0(s,z_1) + \kappa_1(s,z_1)y_s)(\kappa_0(s,z_2) + \kappa_1(s,z_2)y_s)}{\delta_0 + \delta_1 y_s} ds
\]

(5.10)

for \( z_1, z_2 \in S_f \). The angle bracket \( \langle L(z), \overline{L(z)} \rangle = \langle L(z), L(z) \rangle \) is a nonnegative increasing process. Hence

\[
0 \leq \frac{|\kappa_0(t,z) + \kappa_1(t,z)y_t|^2}{\delta_0 + \delta_1 y_t} \leq \alpha_0(t,z,\overline{z}) + \alpha_1(t,z,\overline{z})y_t
\]

(5.11)
for all \( t \in [0, T] \) because \( \kappa_j(t, z) = \kappa_j(t, z) \) for \( j = 0, 1 \). With the help of Fubini’s theorem we deduce

\[
K \leq \int_0^T E \left[ 1_{[0, \tau_n]} |V(z)_t|^2 (\alpha_0(t, z, z) + \alpha_1(t, z, z)y_t) \right] dt |\Pi|(dz).
\]

Since \( \Pi \) is a finite measure on \( S_f \), the estimate (5.8) follows from Lemma 5.3. In view of (5.10, 5.11) we have

\[
\langle L(z), L(z) \rangle_{\tau_n} \leq \int_0^{\tau_n} |V(z)_t|^2 (\alpha_0(t, z, z) + \alpha_1(t, z, z)y_t) dt \tag{5.12}
\]

Hence (5.9) follows from Lemma 5.3 as well. \( \square \)

We can now determine the GKW decomposition of \( V \).

**Lemma 5.5**

\[ \vartheta^* := \int \vartheta(z) \Pi(dz) \]

defines a real-valued admissible trading strategy. Moreover,

\[ L := \int L(z) \Pi(dz) \]

is a real-valued square-integrable martingale orthogonal to \( S \) (i.e. such that \( LS \) is a local martingale). Finally, we have

\[ V = V_0 + \vartheta^* \cdot S + L. \]

**Proof.** Using Hölder’s inequality and (5.8) we obtain

\[
E \left[ \int_0^{\tau_n} |\vartheta^*_t|^2 d\langle S, S \rangle_t \right] \\
\leq E \left[ \int_0^{\tau_n} \left( \int |\vartheta(z)_t| |\Pi|(dz) \right)^2 d\langle S, S \rangle_t \right] \\
\leq |\Pi|(S_f) E \left[ \left( \int |\vartheta(z)|^2 |\Pi|(dz) \right) \cdot \langle S, S \rangle_{\tau_n} \right] < \infty.
\]

for the stopping times \( \tau_n \) from Lemma 5.3. Hence \( \vartheta^* \in L^2_{\text{loc}}(S) \). Lemma 5.2 implies that \( SL(z) \in \mathcal{M}_{\text{loc}} \) for all \( z \in S_f \). Let \( (\sigma_n)_{n \geq 1} \) denote a localizing sequence for \( S_{\sigma_n} \in \mathcal{H}^2_{\text{loc}} \). Since \( L(z) \in \mathcal{H}^2 \), it follows that \( (SL(z))_{\sigma_n} \in \mathcal{M} \), cf. [17] I.4.2. Now we set \( \tilde{\tau}_n := \tau_n \land \sigma_n \).

Since \( \langle L(z), L(z) \rangle \) is an increasing process, Jensen’s inequality yields that

\[
(E[|L(z)_{\tilde{\tau}}|^2])^2 \leq E[|L(z)_{\tilde{\tau}}|^2] \leq E[\langle L(z), L(z) \rangle_{\tilde{\tau}}] \tag{5.13}
\]

for any stopping time \( \tau \). In Lemma 5.3 we have shown that

\[
\int_0^T E \left[ 1_{[0, \tau_n]} |V(z)_t|^2 (\alpha_0(t, z, z) + \alpha_1(t, z, z)y_t) \right] dt
\]
is uniformly bounded in \( z \in S_f \). Due to \( S_{t^n}^z \in \mathbb{H}^2 \) we have
\[
K := \sup \left\{ E \left[ \left( S_t^z \right)^2 \right] : t \in [0, T] \right\} < \infty.
\]
In view of the Cauchy-Schwarz inequality and (5.13), (5.12) one obtains
\[
\int E[\left| S_t^z L(\tau)^{\tau_n} \right|] \Pi \, (dz)
\leq \sqrt{K} \int \left( \int_0^T E \left[ 1 \leq \int_0^T [V(z)]_t^2 (\alpha_0(t, z, \bar{z}) + \alpha_1(t, z, \bar{z}) y_t)] \, dt \right)^{\frac{1}{2}} \Pi \, (dz) < \infty
\]
and similarly
\[
E \left( \int |L(z)|_{\tau_n} \, \Pi \, (dz) \right) = \int E[|L(z)|_{\tau_n}] \Pi \, (dz) < \infty
\]
for \( t \in [0, T] \). In particular, the integral in the definition of \( L \) is finite. An application of Fubini’s theorem yields
\[
E \left[ \left( S_t^z L_t^{\tau_n} - S_s^z L_s^{\tau_n} \right) 1_A \right] = \int E \left[ \left( S_t^z L(z)^{\tau_n} - S_s^z L(z)^{\tau_n} \right) 1_A \right] \Pi \, (dz) = 0.
\]
for \( A \in \mathcal{F}_s \) and \( s \leq t \). Hence \( SL \in \mathcal{M}_{\text{loc}} \) and therefore \( \langle S, L \rangle = 0 \). A similar Fubini-type argument yields that \( L_{\tau_n} \in \mathcal{M} \) because \( L(z) \in \mathcal{H}^2 \) for all \( z \in S_f \) and \( \mathcal{H}^2 \) is stable under stopping. Furthermore, (5.9) leads to
\[
\sup_{t \in [0, T]} E[|L_t^{\tau_n}|^2] = E[|L_{\tau_n}|^2]
\leq \int \int E[|L(z_1)_{\tau_n} L(z_2)_{\tau_n}|] \Pi \, (dz_1) \Pi \, (dz_2)
\leq |\Pi| \langle S_f \rangle \int E[|L(z)_{\tau_n}|^2] \Pi \, (dz)
\leq |\Pi| \langle S_f \rangle \int E[|L(z), L(z)|_{\tau_n}] \Pi \, (dz) < \infty
\]
and hence \( L \in \mathcal{H}_{\text{loc}}^2 \). Obviously, \( L \) starts in zero. From Proposition 3.4 and Lemma 5.2 we know that
\[
V = V_0 + \int (\vartheta(z) \cdot S) \Pi \, (dz) + L.
\]
In view of (5.8), Fubini’s theorem for stochastic integrals [30, Theorem IV.65] yields
\[
\int (\vartheta(z) \cdot S) \Pi \, (dz) = \left( \int \vartheta(z) \Pi \, (dz) \right) \cdot S = \vartheta^* \cdot S.
\]
From
\[
(\vartheta^* - \overline{\vartheta}) \cdot \langle S, S \rangle = \langle V, S \rangle - \langle \overline{V}, S \rangle = 0
\]
it follows that \( \vartheta^* \) and hence also \( L \) is real-valued. The admissibility of \( \vartheta^* \) and the square-integrability of \( L \) follow as in Lemma 5.2 from \( V \in \mathcal{H}^2 \). \( \square \)
The GKW decomposition of $V$ leads to the variance-optimal hedge.

\textbf{Proof of Theorem 4.1.} The assertion follows from Lemma 5.5 combined with Theorem 3 and Corollary 10. \hfill \Box

Finally, we turn to the proof of the formula for the hedging error. Observe that the superscript $c$ in the following does \textit{not} refer to the continuous martingale part of processes.

\textbf{Proof of Theorem 4.2.} Let us consider the truncated payoff function

$$f^c(s) := \int s^2 \Pi^c(dz),$$

where $c \in \mathbb{R}_+$, $\Pi^c(A) := \Pi(S^c_f \cap A)$ for all $A \in \mathcal{B}(S_f)$, and

$$S^c_f := \{z \in \mathbb{C} : R' \leq \text{Re}(z) \leq R, |\text{Im}(z)| \leq c\}.$$

Observe that $f^c$ is real-valued because $\Pi$ is symmetric. $H^c := f^c(S_T)$ is the corresponding contingent claim. Its price process $V^c$ is defined by $V^c_t := E[H^c|\mathcal{F}_t]$ as in Section 3. By Proposition 3.4 we have

$$V^c_t = \int V(z)_t \Pi^c(dz) = \int_{S^c_f} V(z)_t \Pi(dz).$$

Hölder’s inequality yields

$$|V^c_T - V_T|^2 \leq \left( \int_{S^c_f} |V(z)_T| \Pi(dz) \right)^2 \leq |\Pi|(S_f) \int |V(z)_T|^2 \Pi(dz) \leq (|\Pi|(S_f))^2 \left(e^{2R\log S_T} + e^{2R'\log S_T}\right).$$

Since the expression on the right-hand side is square-integrable, we have

$$H^c = V^c_T \xrightarrow{L^2} V_T = H$$

by dominated convergence. If we set

$$(\vartheta^c)^* := \int \vartheta(z) \Pi^c(dz) \quad \text{and} \quad L^c := \int L(z) \Pi^c(dz),$$

then

$$V^c = V^c(0) + \vartheta^* \cdot S + L$$

is the Galtchouk-Kunita-Watanabe decomposition of $V^c$ in the sense of Lemma 5.5. Theorem 3.8] shows that $L^c_T$ converges to $L_T$ in $L^2$. Hence we have

$$J_0 = E[L_T^2] = \lim_{c \to \infty} E[(L^c_T)^2] = \lim_{c \to \infty} E[\langle L^c, L^c \rangle_T]. \quad (5.14)$$

In order to finish the proof we need the following Fubini-type theorem for the predictable covariation.
Lemma 5.6 We have
\[ \langle L^c, L^c \rangle = \int \int \langle L(z_1), L(z_2) \rangle \Pi^c(dz_1) \Pi^c(dz_2), \]
where \( \langle L(z_1), L(z_2) \rangle \) is given by (5.10).

PROOF. This is shown similarly as in the proof of [15 Theorem 3.2], cf. [28 Lemma 3.26] for details. \( \square \)

We can now complete the proof of Theorem 4.2. From (5.14) it follows with Lemma 5.6 that
\[ E[L^c_T] = \lim_{c \to \infty} E \left[ \int_{S^c_0} \int_{S^c_0} \langle L(z_1), L(z_2) \rangle \Pi(dz_1) \Pi(dz_2) \right]. \] (5.15)

We focus on the case \( \delta_0, \delta_1 \neq 0 \). The others follow along the same lines. Note that \( \langle X, \overline{X} \rangle \) is an increasing and integrable process for all \( \mathbb{C} \)-valued square-integrable martingales \( X \). From (5.10) it follows that
\[ E[\langle L(z_1), L(z_2) \rangle_T] = \int_0^T \left( \frac{\eta_2}{\delta_0} E[V(z_1), V(z_2)] + \frac{\eta_1 \delta_1 - \eta_2 \delta_0}{\delta_0^2} E[V(z_1), V(z_2)] +(5.16)\]
where \( \eta_0, \eta_1, \) and \( \eta_2 \) defined as in the assertion depend on \( t, z_1, \) and \( z_2 \). In view of Assumption 3.1(1) and Remark 3.2(2), we obtain
\[ E[V(z_1), V(z_2)] = e^{\xi_0} S_0^{z_1+z_2} \exp \left( \Psi_0(t, \xi_1, z_1 + z_2) + \Psi_1(t, \xi_1, z_1 + z_2) y_0 \right), \] (5.17)
where \( \xi_0 \) and \( \xi_1 \) defined as in the assertion depend on \( t, z_1, \) and \( z_2 \). For \( (x_1, x_2) \in S \) we set \( g_0(x_1, x_2) := \exp(x_1 y_1 + x_2 Z_1) \) and \( g_1(x_1, x_2) := E[g_0(x_1, x_2)]. \) Moreover, let \( B(\xi_1, \tilde{\varepsilon}) := \{ z \in \mathbb{C} : |z - \xi_1| < \tilde{\varepsilon} \} \) denote the ball around \( \xi_1 \) with radius \( \tilde{\varepsilon} \). Choose \( \tilde{\varepsilon} := \varepsilon/2 \) for \( \varepsilon \) as in Assumption 3.1(1). We have
\[ \tilde{\varepsilon} \sup \{ |D_1 g_0(\xi, z_1 + z_2)| : \xi \in B(\xi_1, \tilde{\varepsilon}) \} \leq e^{(\lambda_0 + 2\tilde{\varepsilon}) t} \left( e^{2RZ_1} + e^{2RZ_1} \right). \]
The right-hand side of this inequality has finite expectation by Assumption 3.1(1) and Remark 3.2(2). Since
\[ \left| \frac{g_0(\xi_1 + \eta, z_1 + z_2) - g_0(\xi_1, z_1 + z_2)}{\eta} \right| \leq \varepsilon + \sup \{ |D_1 g_0(\xi, z_1 + z_2)| : \xi \in B(\xi_1, \tilde{\varepsilon}) \} \]
for sufficiently small \( |\eta| \), dominated convergence yields
\[ E[V(z_1), V(z_2)] = e^{\xi_0} S_0^{z_1+z_2} E[D_1 g_0(\xi_1, z_1 + z_2)] = e^{\xi_0} S_0^{z_1+z_2} D_1 g_1(\xi_1, z_1 + z_2) \]
\[ = E[V(z_1), V(z_2)] \left( D_2 \Psi_0(t, \xi_1, z_1 + z_2) + D_2 \Psi_1(t, \xi_1, z_1 + z_2) y_0 \right). \] (5.18)
Moreover,

\[
e^{\xi_1(y_t + \frac{\delta_0}{\delta_1})}_{y_t + \frac{\delta_0}{\delta_1}} = \int_\Gamma e^{z(y_t + \frac{\delta_0}{\delta_1})} dz = \int_0^1 \left( \frac{\delta_1}{\delta_0} + \xi_1 s \right) \exp \left( \left( \frac{\delta_1}{\delta_0} \log(s) + \xi_1 s \right) y_t + \frac{\delta_0}{\delta_1} \xi_1 s \right) ds
\]

with \( \Gamma : [0, 1] \to \mathbb{C}, s \mapsto \frac{\delta_1}{\delta_0} \log(s) + \xi_1 s \). It follows that

\[
E \left[ \frac{V(z_1)_t V(z_2)_t}{\delta_0 + \delta_1 y_t} \right] = e^{\xi_1 S_0^{z_1} + z_2} e^{\frac{\delta_0}{\delta_1} \xi_1 s + (z_1 + z_2) Z_t} \int_0^1 \left( \frac{\delta_1}{\delta_0} + \xi_1 s \right) \exp \left( \left( \frac{\delta_1}{\delta_0} \log(s) + \xi_1 s \right) y_t + \frac{\delta_0}{\delta_1} \xi_1 s + (z_1 + z_2) Z_t \right) ds
\]

By \( \delta_0, \delta_1 > 0 \) we have

\[
E \left[ \frac{V(z_1)_t V(z_2)_t}{\delta_0 + \delta_1 y_t} \right] = e^{\xi_1 S_0^{z_1} + z_2} e^{\frac{\delta_0}{\delta_1} \xi_1 s + (z_1 + z_2) Z_t} \int_0^1 \left( \frac{\delta_1}{\delta_0} + \xi_1 s \right) \exp \left( \left( \frac{\delta_1}{\delta_0} \log(s) + \xi_1 s \right) y_t + \frac{\delta_0}{\delta_1} \xi_1 s + (z_1 + z_2) Z_t \right) ds.
\]

(5.19)

From (5.17–5.19) and from the continuity of \( \Psi_0, \Psi_1, D_2 \Psi_0, \) and \( D_2 \Psi_1 \) on \([0, T] \times S \), we obtain

\[
E[|V(z)_t|^2] = E[V(z)_tV(z)_t] \leq k_1(c),
\]

\[
E[|V(z)_t|^2 y_t] = E[V(z)_tV(z)_ty_t] \leq k_2(c),
\]

\[
E \left[ \frac{|V(z)_t|^2}{\delta_0 + \delta_1 y_t} \right] = E \left[ \frac{|V(z)_tV(z)_t|}{\delta_0 + \delta_1 y_t} \right] \leq k_3(c)
\]

for all \((t, z) \in [0, T] \times S_0^c\) and some finite numbers \( k_1(c), k_2(c), \) and \( k_3(c) \) which may depend on \( c \) but neither on \( t \) nor on \( z \). Consequently, (5.16) yields

\[
E[\langle L(z), L(z) \rangle_T] = E[\langle L(z), L(z) \rangle_T] \leq \int_0^T \left( |\eta_2(t, z, \bar{z})| \frac{k_2(c)}{\delta_1} + (\delta_1 |\eta_1(t, z, \bar{z})| + \delta_0 |\eta_2(t, z, \bar{z})|) \frac{k_1(c)}{\delta_1^2} \right) dt
\]

\[
+ \left( \delta_1^2 |\eta_0(t, z, \bar{z})| + \delta_0 |\eta_1(t, z, \bar{z})| + \delta_0^2 |\eta_2(t, z, \bar{z})| \right) \frac{k_3(c)}{\delta_1^3}
\]

\[
\leq k_4(c)
\]
for all \( z \in S_j^c \) and a positive constant \( k_c(c) \) only depending on \( c \) because the mappings

\[
(t, z) \mapsto \eta_0(t, z, z), \eta_1(t, z, z), \eta_2(t, z, z)
\]

are continuous on \([0, T] \times S_j^c\) by Assumption 3.1(2). Hence, Fubini’s theorem yields

\[
E[L_T^2] = \lim_{c \to \infty} \int_{S_j^c} \int_{S_j^c} E[(L(z_1), L(z_2))_T] \Pi(dz_1) \Pi(dz_2)
\]

in view of (5.15). The assertion follows now from Equations (5.16–5.18, 5.20). \( \square \)

### A Appendix

The dynamic behaviour of a semimartingale can be described in terms of predictable characteristics \((B, C, \nu)\) (cf. [17] for a definition and further background). For most continuous-time processes in applications the characteristics factorize as follows.

**Definition A.1** Let \( X \) be an \( \mathbb{R}^d \)-valued semimartingale with characteristics \((B, C, \nu)\) relative to some truncation function \( h : \mathbb{R}^d \to \mathbb{R}^d \). If there are some \( \mathbb{R}^d \)-valued predictable process \( b \), some predictable \( \mathbb{R}^{d \times d} \)-valued process \( c \) whose values are non-negative, symmetric matrices, and some transition kernel \( F \) from \((\Omega \times \mathbb{R}_+, \mathcal{F})\) into \((\mathbb{R}^d, \mathcal{B}_d)\) such that

\[
B_t = \int_0^t b_s ds, \quad C_t = \int_0^t c_s ds, \quad \nu([0, t] \times G) = \int_0^t F_s(G) ds \quad \text{for } t \in [0, T], G \in \mathcal{B}_d,
\]

we call \((b, c, F)\) differential characteristics of \( X \).

One can interpret \( b_t \) or rather \( b_t + \int (x - h(x))F_t(dx) \) as a drift rate, \( c_t \) as a diffusion coefficient, and \( F_t \) as a local jump measure. For a Lévy process the triplet \((b, c, F)\) is deterministic and does not depend on \( t \). In this case it coincides with the Lévy-Khintchine triplet appearing in the Lévy-Khintchine formula. In a sense the differential characteristics generalize this triplet for more general semimartingales. They are typically derived from other “local” representations of the process e.g. in terms of a stochastic differential equation.

**Proposition A.2** Let \( X \) be a \( \mathbb{R}^d \)-valued semimartingale with differential characteristics \((b, c, F)\) and \( \xi \) an \( \mathbb{R}^{n \times d} \)-valued predictable process whose components \( \xi_j^*, j = 1, \ldots, n \) are integrable with respect to \( X \). Then the \( \mathbb{R}^n \)-valued integral process \( \xi \cdot X := (\xi_j^* \cdot X)_{j=1,\ldots,n} \) has differential characteristics \((\tilde{b}, \tilde{c}, \tilde{F})\) of the form

\[
\tilde{b}_t = \xi_t b_t + \int \left( \tilde{h}(\xi_t x) - \xi_t h(x) \right) F_t(dx),
\]

\[
\tilde{c}_t = \xi_t c_t \xi_t^*,
\]

\[
\tilde{F}_t(G) = \int 1_G(\xi_t x)F_t(dx) \quad \forall G \in \mathcal{B}_n \quad \text{with} \quad 0 \notin G.
\]

Here, \( h \) and \( \tilde{h} \) denote truncation functions on \( \mathbb{R}^d \) and \( \mathbb{R}^n \), respectively.
Proposition A.3 Let $X$ denote an $\mathbb{R}^d$-valued semimartingale with differential characteristics $(b, c, F)$. Suppose that $g : U \rightarrow \mathbb{R}^n$ is twice continuously differentiable on some open subset $U \subset \mathbb{R}^d$ such that $X, X_-$ are $U$-valued. Then the $\mathbb{R}^n$-valued semimartingale $g(X)$ has differential characteristics $(\tilde{b}, \tilde{c}, \tilde{F})$ if the form

$$
\tilde{b}_t^i = \sum_{k=1}^d D_k g^i(X_{t-}) b_t^k + \frac{1}{2} \sum_{k,l=1}^d D^2_{k,l} g^i(X_{t-}) c_t^{k,l}
+ \int \left( \hat{h}^i (g(X_{t-} + x) - g(X_{t-})) - \sum_{k=1}^d D_k g^i(X_{t-}) h^k(x) \right) F_t(dx),
$$

$$
\tilde{c}_t^{i,j} = \sum_{k,l=1}^d D_k g^i(X_{t-}) c_t^{k,l} D_l g^j(X_{t-}),
$$

$$
\tilde{F}_t(G) = \int 1_G (g(X_{t-} + x) - g(X_{t-})) F_t(dx) \quad \forall G \in \mathcal{B}^n \quad \text{with} \quad 0 \notin \mathcal{G},
$$

$i, j = 1, \ldots, n$. Here, $D_k$ etc. denote partial derivatives and $h$ and $\hat{h}$ truncation functions on $\mathbb{R}^d$ and $\mathbb{R}^n$, respectively.

Proof. Cf. [12, Corollary A.6].

Acknowledgements

The first author gratefully acknowledges partial support through Sachbeihilfe KA 1682/2-1 of the Deutsche Forschungsgemeinschaft. Thanks are also due to Richard Vierthauer for valuable discussions.

References


