Finance Stoch (2012) 16:17–43
DOI 10.1007/s00780-010-0148-2

# Cross hedging with stochastic correlation

**Stefan Ankirchner · Gregor Heyne**

Received: 30 March 2009 / Accepted: 21 January 2010 / Published online: 27 November 2010


**Abstract** This paper is concerned with the study of quadratic hedging of contingent claims with basis risk. We extend existing results by allowing the correlation between the hedging instrument and the underlying of the contingent claim to be random itself. We assume that the correlation process $\rho$ evolves according to a stochastic differential equation with values between the boundaries $-1$ and 1. We keep the correlation dynamics general and derive an integrability condition on the correlation process that allows to describe and compute the quadratic hedge by means of a simple hedging formula that can be directly implemented. Furthermore, we show that the conditions on $\rho$ are fulfilled by a large class of dynamics. The theory is exemplified by various explicitly given correlation dynamics.

**Keywords** Cross hedging · Incomplete markets · Correlation · Local risk minimisation · BSDE

**Mathematics Subject Classification (2000)** 91G20 · 60H10 · 60H07

**JEL Classification** C30 · G13

## 1 Introduction

If a hedging instrument is not perfectly correlated with the risk to be hedged, then a non-hedgeable risk, called *basis risk*, remains. A prominent example for financial

S. Ankirchner
Institute for Applied Mathematics, University of Bonn, Endenicher Allee 60, 53115 Bonn, Germany
e-mail: ankirchner@hcm.uni-bonn.de

G. Heyne (✉)
Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
e-mail: heyne@math.hu-berlin.de

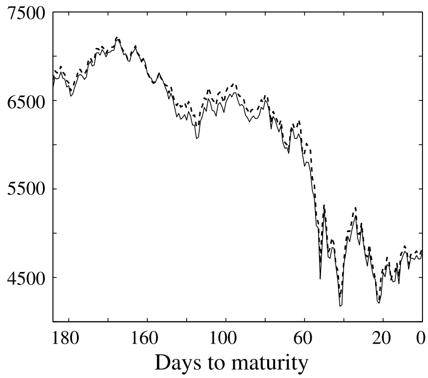


**Fig. 1** Daily average DAX values (*continuous line*) between April and December 2008 and daily EUREX average prices of the DAX futures (*dashed line*) issued in April 2008 and with maturity December 2008

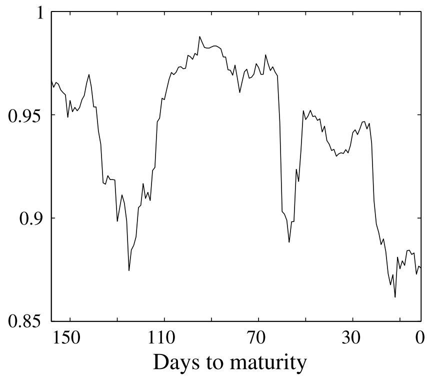


**Fig. 2** Correlation between the log-returns of the DAX and DAX futures of Fig. 1

derivatives that entail basis risk are basket options. As an example, think of options on stock market indices like the Dow Jones or the DAX. In practice, such options may be hedged by trading futures or forwards written on the index. A futures on a stock index is usually highly correlated with the index itself. Figure 1 shows the daily DAX values between April and December 2008, and the daily EUREX average prices of the DAX futures with maturity December 2008.

There are many papers dealing with optimal hedging with basis risk; see for instance [3] and [9] and the references therein. In the literature, two different optimality criteria for the hedge have been applied so far. The first is a utility based approach that aims at maximising the exponential utility of the terminal wealth minus the hedging costs (see, for instance [1, 3, 8]). The second is a quadratic approach that aims at minimising the quadratic hedging error. In all the hedging literature concerned, the correlation between the tradable and non-tradable assets is supposed to be constant. There is empirical evidence, however, that often the correlation is random itself and fluctuates over time. Figure 2 shows the correlation between the DAX and EUREX DAX futures.

In this paper, we extend results on quadratic hedging with basis risk by allowing the correlation to be random. As usual, we assume that the price of the tradable asset and the value of the non-tradable index evolve according to geometric Brownian motions. However, we assume that the correlation between the driving Brownian motions is not constant, but a random process with values between $-1$ and 1. More precisely, we assume that the correlation process is the solution of a stochastic differential equation (SDE) to guarantee that the correlation process possesses the Markov property.

We consider European options on the non-tradable index and derive the asset hedging strategy that locally minimises the quadratic hedging error, the so-called *locally risk minimising strategy* (see Sect. 2 for an introduction to local risk minimisation). Essentially, the optimal hedge can be described by the following factors: the *asset hedge ratio*, defined as

$$\rho_t \frac{\text{index vola}}{\text{asset vola}},$$

where $\rho_t$ is the correlation process, and the *correlation hedge ratio*, defined as

$$\gamma \frac{\text{correlation vola}}{\text{asset vola}},$$

where $\gamma$ is the correlation between the asset and $\rho_t$. The derivative with respect to the asset (resp. the correlation) of the expected value of the option under the so-called *minimal equivalent local martingale measure* will be called *asset delta* (resp. *correlation delta*). We show that the optimal hedge is the asset hedge ratio multiplied with the asset delta plus the correlation hedge ratio multiplied with the correlation delta, i.e.,

$$\begin{aligned}\text{optimal hedge} &= \text{asset hedge ratio} \times \text{asset delta}\\ &\quad + \text{correlation hedge ratio} \times \text{correlation delta}.\end{aligned}$$

In general, by assuming a stochastic correlation, there is no closed formula for the asset delta, but it is straightforward to show that it has a representation in terms of a simple expectation. The main effort, however, lies in showing that the *correlation delta* can be expressed as a simple expectation as well. Under a natural integrability condition on the correlation process, we show that one may differentiate under the expectation and hence obtains the desired representation. With this at hand, one may compute the hedging strategy by simple Monte Carlo simulations.

It will be the topic of future research to compare the performance of the derived hedging formula with naive hedging strategies assuming constant correlation. For models with constant correlation, the performance of quadratic hedging has been compared in [9] with exponential utility based hedging as described in [12].

We want to point out two papers that allow for stochastic correlation in *pricing* contingent claims. Van Emmerich [18] prices quanto options by assuming that the exchange rate is stochastically correlated with the underlying, and Frei and Schweizer [6, 7] deal with exponential utility indifference valuation of contingent claims based on risk sources that are stochastically correlated with assets traded on financial markets.

The paper is organised as follows. In Sect. 2, we give a short introduction to local risk minimisation. Section 3 introduces our model and gives an overview of our main results. The details we use to derive our hedge formula are provided in Sect. 4. We continue in Sect. 5 by analysing the boundary behaviour and integrability properties of correlation processes. We conclude in Sect. 6 by giving some explicit examples of correlation processes for which our main results hold.

## 2 A brief review of local risk minimisation

In this section, we give a short introduction to the theory of local risk minimisation in a quadratic sense. The presented material is a streamlined version of [17].

We start with a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $T > 0$ is a finite time horizon and the filtration $(\mathcal{F}_t)_{0\le t\le T}$ satisfies the usual conditions, i.e., $(\mathcal{F}_t)_{0\le t\le T}$ is right-continuous and completed by the $\mathbb{P}$-nullsets. We consider a financial market with one risky asset $S$ and one non-risky asset, say a money market account with dynamics $B$. We suppose that the discounted asset price $X = S/B$ is an $\mathbb{R}$-valued continuous semimartingale, and we assume that $X$ satisfies the so-called *structure condition (SC)*. This means that $X$ is a special semi-martingale with canonical decomposition

$$X = X_0 + M + A = X_0 + M + \int \lambda \, d\langle M\rangle,$$

where $M$ is a locally square-integrable martingale with $M_0 = 0$ and $\lambda$ is an $\mathbb{R}$-valued predictable process such that the so-called *mean-variance tradeoff* process $K = \int_0^\cdot \lambda \, dA = \int_0^\cdot \lambda^2 \, d\langle M\rangle$ satisfies $K_T < \infty$ $\mathbb{P}$ a.s. It is well known that (SC) is related to an absence-of-arbitrage condition; see [17] for a reference.

**Definition 2.1** ([17], Definition 1.1) The space $\Theta_S$ consists of all $\mathbb{R}$-valued predictable processes $\xi$ such that the stochastic integral process $\int \xi \, dX$ is well defined and

$$\mathbb{E}\left[\int_0^T \xi_s^2 \, d\langle M\rangle_s + \left(\int_0^T |\xi_s \, dA_s|\right)^2\right] < \infty.$$

An $L^2$-*strategy* is a pair $\varphi = (\xi, \eta)$, where $\xi \in \Theta_S$ and $\eta$ is a real-valued adapted process such that the so-called *value process* $V(\varphi) = \xi X + \eta$ is right-continuous and square-integrable. $\varphi$ is called *0-achieving* if $V_T(\varphi) = 0$ $\mathbb{P}$-a.s.

As usual, a strategy $\varphi = (\xi, \eta)$ describes the investment decisions of an agent trading in the financial market. An investor following the strategy $\varphi$ holds $\xi_t$ shares of the discounted asset $X$ at time $t$, and keeps $\eta_t$ units in the money market account. In this section, we use the money market account as numeraire so that we need not bother about the interest rate.

We next consider a payment stream $H = (H_t)_{0\le t\le T}$ kept fixed throughout this introduction. Mathematically, $H$ is right-continuous, adapted, real-valued and square-integrable; the interpretation is that $H_t \in L^2(\mathbb{P})$ represents the total payments on $[0, t]$ arising due to some financial contract. A European contingent claim with maturity $T$ would have $H_t = 0$ for all $t < T$ and just an $\mathcal{F}_T$-measurable payoff $H_T \in L^2(\mathbb{P})$ due at time $T$; in general, the process $H$ involves both cash inflows and outlays, and can (but need not) be of finite variation.

**Definition 2.2** ([17], Definition 1.2) Fix a payment stream $H$. The *(cumulative) cost process* of an $L^2$-strategy $\varphi = (\xi, \eta)$ is

$$C_t^H(\varphi) = H_t + V_t(\varphi) - \int_0^t \xi_s \, dX_s, \quad 0 \le t \le T.$$

$\varphi$ is called *self-financing* (for $H$) if $C^H(\varphi)$ is constant, and *mean-self-financing* if $C^H(\varphi)$ is a martingale (which is then square-integrable). Under the assumption that $X$ fulfils the structure condition (SC) and that the mean-variance tradeoff process $K$ is continuous, we say that an $L^2$-strategy $\varphi$ is *locally risk minimising* if $\varphi$ is 0-achieving and mean-self-financing, and the cost process $C^H(\varphi)$ is strongly orthogonal to $M$, i.e., $\langle M, C^H(\varphi)\rangle_t = 0$ for all $t \in [0, T]$.

Thus, $C_t(\varphi)$ comprises the hedger's accumulated costs during $[0, t]$ including the payments $H_t$, and $V_t(\varphi)$ should therefore be interpreted as the value of the portfolio $\varphi_t = (\xi_t, \eta_t)$ held at time $t$ after the payments $H_t$. In particular, $V_T(\varphi)$ is the value of the portfolio $\varphi_T$ upon settlement of all liabilities, and a natural condition is then to restrict to 0-achieving strategies as in Definition 2.1.

*Remark 2.3* ([17], Remark 1.3) Observe that if $\varphi_t = (\xi_t, \eta_t)$ is a 0-achieving and mean-self-financing $L^2$-strategy for $H$, then $\varphi$ is uniquely determined from $\xi$ (and of course $H$).

It is well known that the locally risk-minimising strategy can be obtained via the so-called *Föllmer–Schweizer (FS) decomposition* of the final payment $H_T$. This is the decomposition of $H_T$ into

$$H_T = H_T^{(0)} + \int_0^T \xi_s^{H_T} \, dX_s + L_T^{H_T} \quad \mathbb{P}\text{-a.s.}, \tag{2.1}$$

where $H_T^{(0)} \in L^2(\mathbb{P})$ is $\mathcal{F}_0$-measurable, $\xi^{H_T}$ is in $\Theta_S$, and the process $L^{H_T}$ is a (right-continuous) square-integrable martingale strongly orthogonal to $M$ and satisfying $L_0^{H_T} = 0$. Notice that such a decomposition can be shown to be unique. Once we have (2.1), the desired strategy $\varphi = (\xi, \eta)$ is then given by

$$\xi = \xi^{H_T}, \qquad \eta = V^{H_T} - \xi^{H_T} X,$$

with

$$V_t^{H_T} = H_T^{(0)} + \int_0^t \xi_s^{H_T} \, dX_s + L_t^{H_T} - H_t, \quad 0 \le t \le T$$

(see Proposition 5.2 of [17]). Furthermore, the associated cost process is given by

$$C_t^H(\varphi) = H_T^{(0)} + L_t^{H_T}, \quad 0 \le t \le T.$$

## 3 The model and the main results

Let $W=(W^1,W^2,W^3)$ be a three-dimensional Brownian motion on a probability space $(\Omega,\mathcal{F},P)$. Consider two processes with dynamics

$$dS_t = S_t\big(\mu_X\,dt+\sigma_X\,dW_t^1\big),$$
$$dU_t = U_t\Big(\mu_I\,dt+\sigma_I\Big(\rho_t\,dW_t^1+\sqrt{1-\rho_t^2}\,dW_t^2\Big)\Big).$$

To simplify the presentation, we assume throughout that all coefficients are constant; more precisely, $\mu_X,\mu_I\in\mathbb{R}$ and $\sigma_X,\sigma_I\in\mathbb{R}\setminus\{0\}$.

We assume that $S$ is the price process of a tradable asset, and $U$ the value process of a non-tradable index. The correlation $\rho$ is assumed to follow

$$d\rho_t = a(\rho_t)\,dt+g(\rho_t)\,d\hat{W}_t,\quad t\ge 0, \tag{3.1}$$

where $\hat{W}$ is given by $\hat{W}=\gamma W^1+\delta W^2+\sqrt{1-\gamma^2-\delta^2}W^3$, and $\gamma$ and $\delta$ are real numbers such that $\delta^2+\gamma^2\le 1$. For the moment, we assume that the coefficients $a$ and $g$ of the correlation dynamics belong to $\mathcal{C}^1(-1,1)$, and that there exists a unique solution $\rho$ of (3.1) with values in $[-1,1]$.

Throughout, we suppose that the interest $r>0$ is constant and let $B_t=e^{rt}$. The discounted processes $S$ and $U$ will be denoted by

$$X_t=e^{-rt}S_t,\qquad I_t=e^{-rt}U_t.$$

Notice that

$$\begin{aligned} dX_t &= X_t\big((\mu_X-r)\,dt+\sigma_X\,dW_t^1\big),\\ dI_t &= I_t\Big((\mu_I-r)\,dt+\sigma_I\Big(\rho_t\,dW_t^1+\sqrt{1-\rho_t^2}\,dW_t^2\Big)\Big). \end{aligned} \tag{3.2}$$

Consider a derivative $d(U_T)$ depending on the non-tradable index. Define the function $h$ by $h(x)=e^{-rT}d(e^{rT}x)$; then $d(U_T)=e^{rT}h(I_T)$. Our goal is to analyse how to hedge the liability $h(I_T)$ by trading the asset $X$. Since the market is incomplete, we need to choose a criterion according to which strategies are chosen and the prices of contingent claims are computed. We use the framework of local risk minimisation of Sect. 2.

Our first main result is an explicit hedge formula, which can be easily implemented, for example by simple Monte Carlo simulation. We state it right away in Theorem 3.1, after a brief collection of some notations and assumptions.

We need the conditional versions of the processes $I$ and $\rho$, which are given by

$$I_s^{t,y,v}=y+\int_t^s I_u^{t,y,v}\Big((\mu_I-r)\,du+\sigma_I\Big(\rho_u^{t,v}\,dW_u^1+\sqrt{1-\big(\rho_u^{t,v}\big)^2}\,dW_u^2\Big)\Big), \tag{3.3}$$

$$\rho_s^{t,v}=v+\int_t^s a\big(\rho_u^{t,v}\big)\,du+\int_t^s g\big(\rho_u^{t,v}\big)\,d\hat{W}_u, \tag{3.4}$$

for $t \in [0, T)$, $(y, v) \in \mathbb{R} \times (-1, 1)$. In order to find a nice representation of the quadratic hedge, we also need the dynamics of the derivatives of $I^{t,y,v}$ and $\rho^{t,v}$ with respect to the initial values $y$ and $v$. Note that the derivative with respect to $y$ of $I^{t,y,v}$ is given by $\frac{\partial}{\partial y} I^{t,y,v} = \frac{I^{t,y,v}}{y} = I^{t,1,v}$, and obviously $\frac{\partial}{\partial y}\rho^{t,v} = 0$. If the correlation process neither attains $-1$ nor 1 up to time $T$, then the derivatives of $I^{t,y,v}$ and $\rho^{t,v}$ with respect to $v$ are defined. Moreover, the processes $\frac{\partial}{\partial v} I^{t,y,v}$ and $\frac{\partial}{\partial v}\rho^{t,v}$, denoted by $\bar{I}^{t,y,v}$ and $\bar{\rho}^{t,v}$, respectively, solve the SDEs

$$\bar{I}_s^{t,y,v} = \int_t^s \bar{I}_u^{t,y,v}\Big((\mu_I - r)\,du + \sigma_I\Big(\rho_u^{t,v}\,dW_u^1 + \sqrt{1-\big(\rho_u^{t,v}\big)^2}\,dW_u^2\Big)\Big) + \int_t^s I_u^{t,y,v}\sigma_I\Big(\bar{\rho}_u^{t,v}\,dW_u^1 - \frac{\rho_u^{t,v}}{\sqrt{1-(\rho_u^{t,v})^2}}\bar{\rho}_u^{t,v}\,dW_u^2\Big), \tag{3.5}$$

$$\bar{\rho}_s^{t,v} = 1 + \int_t^s a'\big(\rho_u^{t,v}\big)\bar{\rho}_u^{t,v}\,du + \int_t^s g'\big(\rho_u^{t,v}\big)\bar{\rho}_u^{t,v}\,d\hat{W}_u, \tag{3.6}$$

for $s \in [t, T]$ (see Theorem V.38 in [14]). Notice that the correlation boundaries $-1$ and 1 are not attained if and only if the stopping times

$$\tau^v = \tau^{t,v} = \inf\big\{s \geq t : \rho_s^{t,v} \in \{-1, 1\}\big\}$$

satisfy $\tau^{t,v} > T$ $\mathbb{P}$-a.s. We formulate this as condition

(H1) $\tau^{t,v} > T$ $\mathbb{P}$-a.s., for any $v \in (-1, 1)$.

Notice that (H1) guarantees that $\int_t^T (g'(\rho_u^{t,v}))^2\,du < \infty$ $\mathbb{P}$-a.s. and

$$\int_t^T \frac{(\bar{\rho}_s^{t,v})^2}{1-(\rho_s^{t,v})^2}\,du < \infty \quad \mathbb{P}\text{-a.s.}, \tag{3.7}$$

and hence the stochastic integrals appearing in (3.5) and (3.6) are defined. Finally, for our aim of deriving an explicit representation of the quadratic hedge, we need to impose a stronger integrability condition on $\rho$ and $\bar{\rho}$ than (3.7). More precisely, we assume that the following condition is satisfied:

(H2) There exists a $p > 1$ such that for every $t \in [0, T]$ and $v_0 \in (-1, 1)$, there exists an open neighbourhood $U \subset (-1, 1)$ of $v_0$ such that

$$\sup_{v \in U} \mathbb{E}\bigg[\int_t^T \bigg|\frac{(\bar{\rho}_s^{t,v})^2}{1-(\rho_s^{t,v})^2}\bigg|^p ds\bigg] < \infty.$$

We are now ready to state our first main result which gives an explicit expression for the locally risk minimising strategy in terms of expectations with respect to the measure $\tilde{\mathbb{P}}$ with density $d\tilde{\mathbb{P}}/d\mathbb{P} = \mathcal{E}(-\int_0^{\cdot} \frac{\mu_X - r}{\sigma_X}\,dW^1)_T$, where $\mathcal{E}(\cdot)$ denotes the Doléans–Dade exponential. Note that $\tilde{\mathbb{P}}$ corresponds to the so-called *minimal martingale measure*; see [17].

**Theorem 3.1** *Suppose that the coefficients $a$ and $g$ in the dynamics of $\rho$ are continuously differentiable on* $(-1, 1)$. *Assume furthermore that both conditions* (H1) *and* (H2) *are satisfied. Let $h$ be Lipschitz and such that the weak derivative $h'$ is Lebesgue-almost everywhere continuous. Then there exists a locally risk minimising strategy $\varphi = (\xi, \eta)$ for the derivative $h(I_T)$, and $\xi$ is given by $\xi_t = \xi(t, I_t, X_t, \rho_t)$, where*

$$\xi(t, y, x, v) = \frac{y}{x}\left[v\frac{\sigma_I}{\sigma_X}\tilde{\mathbb{E}}\big[h'\big(I_T^{t,y,v}\big)I_T^{t,1,v}\big] + \frac{g(v)\gamma}{\sigma_X}\tilde{\mathbb{E}}\big[h'\big(I_T^{t,y,v}\big)\bar{I}_T^{t,1,v}\big]\right]. \tag{3.8}$$

The proof of this theorem is postponed to Sect. 4.

*Remark 3.2* In terms of the original processes $S$ and $U$, the hedge would be given by $\xi_t = \hat{\xi}(t, U_t, S_t, \rho_t)$, where

$$\hat{\xi}(t, y, x, v) = \frac{y}{x}\left[v\frac{\sigma_I}{\sigma_X}\tilde{\mathbb{E}}\big[d'\big(U_T^{t,y,v}\big)U_T^{t,1,v}\big] + \frac{g(v)\gamma}{\sigma_X}\tilde{\mathbb{E}}\big[d'\big(U_T^{t,y,v}\big)\bar{U}_T^{t,1,v}\big]\right]e^{-r(T-t)},$$

with $\bar{U}$ obtained in the same way as $\bar{I}$.

Before we state our second main contribution, let us apply the previous result to derive the locally risk minimising strategy for a European call option.

**Corollary 3.3** *Suppose the correlation is a deterministic function of time. For a strike $K > 0$, let $d(x) = \max\{x - K, 0\}$. Then there exists a locally risk minimising strategy $\varphi = (\xi, \eta)$ for the derivative $d(U_T)$, and $\xi$ is given by*

$$\xi_t = \rho_t \frac{U_t\sigma_I}{S_t\sigma_X}\Delta_{\mathrm{BS}}(t, U_t, \kappa_t, \sigma_I),$$

*where $\kappa_t = -(\mu_I - r)(T-t) + \sigma_I(\frac{\mu_X - r}{\sigma_X})\int_t^T \rho_s\, ds$ and, with $\Phi$ denoting the standard normal cumulative distribution function,*

$$\Delta_{\mathrm{BS}}(t, y, q, \sigma) = \exp(-q)\Phi\left(\frac{\ln(y/K) + (r + \sigma_I^2/2)(T-t) - q}{\sigma_I\sqrt{T-t}}\right)$$

*is the Black–Scholes delta for options on stocks with continuous dividend yield $q$.*

The content of the preceding corollary is only a slight extension of a result already mentioned in [9] for the case of constant correlation. The proof is a simple straightforward calculation.

*Remark 3.4* From the locally risk minimising strategy, we can easily deduce the so-called *mean-variance* optimal hedging strategy of the payoff $h(I_T)$. The mean-variance hedge is defined to be the self-financing strategy minimising the variance of the *global* hedging error, and usually differs from the locally risk minimising strategy (see [16] for an introduction to mean-variance hedging). In the model considered here, the mean-variance tradeoff process is deterministic, and hence, by an appeal to Theorems 4.6 and 4.7 of Schweizer [16], the mean-variance hedge has a representation allowing to derive it numerically by a simply recursion. Namely, letting

$w = \tilde{\mathbb{E}}[h(I_T)]$, the mean-variance optimal strategy $(\tilde{\xi}, \tilde{\eta})$ for $h(I_T)$ satisfies, with $\tilde{\xi} = \xi^{(w)}$,

$$\xi_t^{(w)} = \xi_t + \frac{\mu_X - r}{\sigma_X^2 X_t}\Big(\tilde{\mathbb{E}}\big[h(I_T) \mid \mathcal{F}_t\big] - w - \int_0^t \xi_s^{(w)}\, dX_s\Big)$$

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

$$\tilde{\eta}_t = w + \int_0^t \xi_s^{(w)}\, dX_s - \xi_t^{(w)} X_t$$

for all $t \in [0, T]$.

Our second main result concerns conditions on the coefficients $a$ and $g$ of $\rho$ such that conditions (H1) and (H2) are fulfilled.

**Theorem 3.5** *Let $a$ and $g$ be continuously differentiable with bounded derivatives. We assume that $g(-1) = g(1) = 0$, and that $g$ does not have any roots in $(-1, 1)$. If*

$$\limsup_{x\uparrow 1} \frac{2a(x)(1-x)}{g^2(x)} < 0 \quad \textit{and} \quad \liminf_{x\downarrow -1} \frac{2a(x)(1+x)}{g^2(x)} > 0, \tag{3.9}$$

*then both conditions* (H1) *and* (H2) *are satisfied, and hence the delta hedge is given as in Theorem* 3.1.

The preceding theorem can be generalised, which will enable us to give an example in Sect. 6 where the derivative of $g$ is unbounded. This, however, requires a little more notation, which for ease of exposition is avoided here. See Sect. 5 for a proof of Theorem 3.5 and the more general Proposition 5.3.

## 4 Derivation of the hedge formula

In this section, we derive and prove the hedge formula stated in Theorem 3.1. In Sect. 4.1, we use BSDEs to derive the Föllmer–Schweizer decomposition, which is the key to obtain the formula (3.8). In Sects. 4.2 and 4.3, we elaborate on details which we need along the way. It is in those details that we need conditions (H1) and (H2). Let us first recall the definition of a BSDE.

As in Sect. 3, let $W = (W^1, W^2, W^3)$ be a three-dimensional Brownian motion. The filtration generated by $W$ and completed by the $\mathbb{P}$-nullsets is denoted by $(\mathcal{F}_t)$. Let $T > 0$, $\xi$ an $\mathcal{F}_T$-measurable random variable and $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$ a measurable function such that for all $(y, z) \in \mathbb{R} \times \mathbb{R}^3$, the mapping $f(\cdot, \cdot, y, z)$ is predictable. A solution of the BSDE with *terminal condition* $\xi$ and *generator* $f$ is defined to be a pair of predictable processes $(Y, Z)$ such that almost surely we have $\int_0^T |Z_s|^2\, ds < \infty$, $\int_0^T |f(s, Y_s, Z_s)|\, ds < \infty$, and for all $t \in [0, T]$,

$$Y_t = \xi - \int_t^T Z_s\, dW_s + \int_t^T f(s, Y_s, Z_s)\, ds.$$

The solution processes $(Y, Z)$ are often shown to satisfy some integrability properties. To this end, one usually verifies whether they belong to the following function spaces. Let $p \geq 1$. We denote by $\mathbb{H}^p(\mathbb{R}^3)$ the set of all $\mathbb{R}^3$-valued predictable processes $\zeta$ such that $E[\int_0^T |\zeta_t|^p \, dt] < \infty$, and by $\mathbb{S}^p(\mathbb{R})$ the set of all $\mathbb{R}$-valued predictable processes $\delta$ satisfying $E[\sup_{s\in[0,T]} |\delta_s|^p] < \infty$.

## 4.1 Deriving the FS decomposition with BSDEs

As stated in Sect. 3, we use the framework of local risk minimisation. Accordingly, note first that $X$ satisfies the structure condition (SC), i.e., $X$ is a special semimartingale with canonical decomposition given by $M_t = \int_0^t \sigma_X X_s \, dW_s^1$, $\lambda_t = \frac{\mu_X - r}{\sigma_X^2 X_t}$ and hence $K_t = (\frac{\mu_X - r}{\sigma_X})^2 t$. In order to find the FS decomposition, we consider a BSDE with terminal condition $h(I_T)$ and driver $f$ to be specified later, namely

$$Y_t = h(I_T) - \int_t^T Z_s \, dW_s + \int_t^T f(s, Y_s, Z_s) \, ds. \tag{4.1}$$

Assume that $f$ can be chosen such that

$$\int_0^t \xi \, dX_s = \int_0^t Z_s^1 \, dW_s^1 - \int_0^t f(s, Y_s, Z_s) \, ds \tag{4.2}$$

for all $t \in [0, T]$. Also, by using (3.2), we have

$$\int_0^t \xi \, dX_s = \int_0^t \xi \sigma_X X_s \, dW_s^1 + \int_0^t \xi X_s(\mu_X - r) \, ds. \tag{4.3}$$

Uniqueness of semimartingale decompositions yields that the martingale parts of (4.2) and (4.3) coincide, and therefore it must hold $\xi_t = \frac{Z_t^1}{\sigma_X X_t}$ $\mathbb{P} \otimes \lambda$-a.e. Moreover, the driver $f$ has to satisfy

$$f(s, y, z) = -z^1 \frac{\mu_X - r}{\sigma_X}. \tag{4.4}$$

Indeed, one can show that the solution of the BSDE with generator (4.4) provides the FS decomposition. We summarise this in the next result, which is in fact a special case of Proposition 1.1 in [5].

**Lemma 4.1** *The FS decomposition of $h(I_T)$ is given by*

$$h(I_T) = Y_0 + \int_0^T \frac{Z_s^1}{\sigma_X X_s} \, dX_s + \int_0^T Z_s^2 \, dW_s^2 + \int_0^T Z_s^3 \, dW_s^3,$$

*where $(Y_s, Z_s)_{0\leq s\leq T}$ is the solution of the BSDE (4.1) with generator $f$ defined as in (4.4).*

In order to obtain a characterisation of the solution of

$$Y_t = h(I_T) - \int_t^T Z_s \, dW_s - \int_t^T Z_s^1 \frac{\mu_X - r}{\sigma_X} \, ds,$$

we consider the conditional forward SDEs given by (3.3) and (3.4) and the associated conditional BSDE

$$Y_s^{t,y,v} = h\big(I_T^{t,y,v}\big) - \int_s^T Z_u^{t,y,v}\, dW_u - \int_s^T Z_u^{1,t,y,v} \frac{\mu_X - r}{\sigma_X}\, du \tag{4.5}$$

for $s \in [t,T]$. Since the BSDE (4.5) is linear, we know by standard results (see for example [5]) that

$$Y_s^{t,y,v} = \mathbb{E}\big[h\big(I_T^{t,y,v}\big)\Gamma_T^s \mid \mathcal{F}_s\big] = \mathbb{E}\left[h\big(I_T^{t,y,v}\big)\frac{\Gamma_T^0}{\Gamma_s^0} \,\middle|\, \mathcal{F}_s\right],$$

where $\Gamma_s^t = \exp(-\frac{\mu_X - r}{\sigma_X}(W_s^1 - W_t^1) - \frac{(\mu_X - r)^2}{2\sigma_X^2}(s-t))$ is the solution of

$$d\Gamma_s^t = -\Gamma_s^t \frac{\mu_X - r}{\sigma_X}\, dW_s^1, \qquad \Gamma_t^t = 1.$$

Let $\tilde{\mathbb{P}}$ be the probability measure with density $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \Gamma_T^0$, and denote with $\psi$ the function defined by

$$\psi(t,y,v) = \tilde{\mathbb{E}}\big[h\big(I_T^{t,y,v}\big)\big]. \tag{4.6}$$

That the function $\psi$ is well defined and has first derivatives with respect to $y$ and $v$ follows from Sects. 4.2 and 4.3. The value process of the solution of the BSDE (4.5) satisfies $Y_s^{t,y,v} = \psi(s, I_s^{t,y,v}, \rho_s^{t,v})$. Our main goal is to derive the explicit hedge formula (3.8). With the help of Lemma 4.9, we get for $Z_s^{t,y,v}$ the representation

$$Z_s^{t,y,v} = \sigma\big(I_s^{t,y,v}, \rho_s^{t,v}\big)^* \begin{pmatrix} \partial_y \psi(s, I_s^{t,y,v}, \rho_s^{t,v}) \\ \partial_v \psi(s, I_s^{t,y,v}, \rho_s^{t,v}) \end{pmatrix},$$

where the volatility matrix $\sigma(y,v)$ is given by

$$\sigma(y,v) = \begin{pmatrix} y\sigma_I v & y\sigma_I\sqrt{1-v^2} & 0 \\ g(v)\gamma & g(v)\delta & g(v)\sqrt{1-\gamma^2-\delta^2} \end{pmatrix}, \qquad y \in \mathbb{R},\ v \in (-1,1).$$

Hence, we have

$$Z_s^{1,t,y,v} = I_s^{t,y,v}\sigma_I\rho_s^{t,v}\partial_y\psi\big(s, I_s^{t,y,v}, \rho_s^{t,v}\big) + g\big(\rho_s^{t,v}\big)\gamma\,\partial_v\psi\big(s, I_s^{t,y,v}, \rho_s^{t,v}\big),$$

i.e., the hedge formula is given by

$$\xi(t,y,x,v) = \frac{y\sigma_I v\partial_y\psi(s,y,v) + g(v)\gamma\,\partial_v\psi(s,y,v)}{\sigma_X x}. \tag{4.7}$$

Thus, by plugging in the explicit representations of $\partial_y\psi(s,y,v)$ and $\partial_v\psi(s,y,v)$ given in Sect. 4.3, we obtain (3.8), i.e., we have proved Theorem 3.1.

*Remark 4.2* Note that the approach we take by characterising the Föllmer–Schweizer decomposition via the solution of a linear BSDE is the same as in Example 1.3 in [5].

In our model, however, the inverse of the volatility matrix of the asset processes $X$ and $I$ is unbounded, and hence does not fall within the specifications of Hypothesis 1.1 in [5]. Moreover, the coefficients of the volatility matrix of the forward processes $I$ and $\rho$ associated with the BSDE do not satisfy the prerequisites of Proposition 5.9 in [5], i.e., we do not have uniformly bounded derivatives. In order to recover our hedge formula in spite of these extensions, we apply the results of our Sects. 4.2–4.4 and Sect. 5.

## 4.2 Differentiability with respect to the initial conditions

In this section, we want to make some remarks on the system of SDEs given by (3.3)–(3.6), concerning existence, uniqueness, continuity and differentiability with respect to the initial values $y$ and $v$. We also observe the following.

**Lemma 4.3** *Suppose that* (H1) *holds. Then the SDE for* $\bar{I}^{t,y,v}$ *in* (3.5) *has a unique solution which is given by*

$$\bar{I}^{t,y,v}_s = y\mathcal{E}\big(G^{t,v}\big)_s \int_t^s \mathcal{E}\big(G^{t,v}\big)^{-1}_u \, dH^{t,1,v}_u, \tag{4.8}$$

*where*

$$H^{t,y,v}_s = \int_t^s I^{t,y,v}_u \sigma_I \left( \bar{\rho}^{t,v}_u \, dW^1_u - \frac{\rho^{t,v}_u}{\sqrt{1-(\rho^{t,v}_u)^2}} \bar{\rho}^{t,v}_u \, dW^2_u \right)$$

*and*

$$G^{t,v}_s = \int_t^s (\mu_I - r)\, du + \int_t^s \sigma_I \left( \rho^{t,v}_u \, dW^1_u + \sqrt{1-\big(\rho^{t,v}_u\big)^2} \, dW^2_u \right).$$

*Proof* Due to the assumption (H1), we can define the semimartingales $(H^{t,y,v}_s)_{t\le s\le T}$ and $(G^{t,v}_s)_{t\le s\le T}$ as above. By looking at the dynamics (3.5), we see immediately that $\bar{I}^{t,y,v}_s$ is the solution of the linear stochastic equation

$$\bar{I}^{t,y,v}_s = H^{t,y,v}_s + \int_t^s \bar{I}^{t,y,v}_u \, dG^{t,v}_u. \tag{4.9}$$

The solution of (4.9) is given by

$$\bar{I}^{t,y,v}_s = \mathcal{E}\big(G^{t,v}\big)_s \left( H^{t,y,v}_t + \int_t^s \mathcal{E}\big(G^{t,v}\big)^{-1}_u \big( dH^{t,y,v}_u - d\langle H^{t,y,v}, G^{t,v}\rangle_u \big) \right)$$

(see Proposition IX.2.3 in [15]). Notice that

$$\begin{aligned} d\langle H^{t,y,v}, G^{t,v}\rangle_u &= I^{t,y,v}_u \sigma_I^2 \bar{\rho}^{t,v}_u \rho^{t,v}_u \, du - I^{t,y,v}_u \sigma_I^2 \sqrt{1-\big(\rho^{t,v}_u\big)^2} \frac{\rho^{t,v}_u}{\sqrt{1-(\rho^{t,v}_u)^2}} \bar{\rho}^{t,v}_u \, du \\ &= 0. \end{aligned}$$

Since $H^{t,y,v}_t = 0$ and $H^{t,y,v} = yH^{t,1,v}$, we obtain (4.8). □

*Remark 4.4* The process $I^{t,y,v}$ is given by

$$I_s^{t,y,v} = y \exp\biggl(\int_t^s \sigma_I \Bigl(\rho_u^{t,v}\, dW_u^1 + \sqrt{1-\bigl(\rho_u^{t,v}\bigr)^2}\, dW_u^2\Bigr) + \int_t^s \Bigl(-\frac{1}{2}\sigma_I^2 + \mu_I - r\Bigr)\, du\biggr).$$

Moreover suppose that $\int_t^s g'(\rho_u^{t,v}) d\hat{W}_u$ is well defined. Then the SDE in (3.6) has a unique solution $\bar{\rho}^{t,v}$ given by

$$\bar{\rho}_s^{t,v} = \mathcal{E}\biggl(\int_t^s g'\bigl(\rho_u^{t,v}\bigr)\, d\hat{W}_u + \int_t^s a'\bigl(\rho_u^{t,v}\bigr)\, du\biggr).$$

Before we end this section, we give an auxiliary result which will be used in the sequel.

**Lemma 4.5** *Consider two predictable processes $c^v$ and $d^v$, depending on a parameter $v \in (-1,1)$. Suppose that there exists a continuous function $D : (-1,1) \to \mathbb{R}_+$ such that $c^v + |d^v| \le D(v)$ for all $v \in (-1,1)$. Then the process*

$$b_s^{t,v} = \mathcal{E}\biggl(\int_t^s d_u^v\, d\hat{W}_u + \int_t^s c_u^v\, du\biggr)$$

*is well defined, and for all $p \ge 1$ and $v_0 \in (-1,1)$, there exists an open neighbourhood $U \subset (-1,1)$ of $v_0$ such that*

$$\sup_{v\in U} \mathbb{E}\Bigl[\sup_{t\le u\le T} \bigl|b_u^{t,v}\bigr|^p\Bigr] < \infty.$$

*Proof* Let $p \ge 1$. The Burkholder–Davis–Gundy inequality implies that for a constant $C_p$, depending only on $p$, we have

$$\begin{aligned}
\mathbb{E}\Bigl[\sup_{t\le u\le T} \bigl|b_u^{t,v}\bigr|^p\Bigr] &\le e^{p(T-t)D(v)} \mathbb{E}\biggl[\sup_{t\le u\le T} \biggl|\mathcal{E}\biggl(\int_t^u d_w^v\, d\hat{W}_w\biggr)\biggr|^p\biggr] \\
&\le C_p e^{p(T-t)D(v)} \mathbb{E}\biggl[\biggl|\int_t^T \biggl|\bigl(d_u^v\bigr)^2 \mathcal{E}\biggl(\int_t^u d_w^v\, d\hat{W}_w\biggr)\biggr|^2 du\biggr|^{p/2}\biggr] \\
&\le C_p e^{p(T-t)D(v)} D^p(v) \mathbb{E}\biggl[\biggl|\int_t^T \biggl|\mathcal{E}\biggl(\int_t^u d_w^v\, d\hat{W}_w\biggr)\biggr|^2 du\biggr|^{p/2}\biggr].
\end{aligned}$$

In the rest of the proof, we have to distinguish between $p \ge 2$ and $p \in [1,2)$. We first consider $p \ge 2$. By Jensen's inequality and Fubini's theorem, we get

$$\begin{aligned}
\mathbb{E}\biggl[\biggl|\int_t^T \biggl|\mathcal{E}\biggl(\int_t^u d_w^v\, d\hat{W}_w\biggr)\biggr|^2 du\biggr|^{p/2}\biggr] &\le \mathbb{E}\biggl[\int_t^T \biggl|\mathcal{E}\biggl(\int_t^u d_w^v\, d\hat{W}_w\biggr)\biggr|^p du\biggr] \\
&= \int_t^T \mathbb{E}\biggl[\biggl|\mathcal{E}\biggl(\int_t^u d_w^v\, d\hat{W}_w\biggr)\biggr|^p\biggr] du. \quad (4.10)
\end{aligned}$$

Notice that

$$\left|\mathcal{E}\left(\int_t^u d_w^v\, d\hat{W}_w\right)\right|^p = \exp\left(\int_t^u p d_w^v\, d\hat{W}_w + \int_t^u p^2 \left(d_w^v\right)^2 dw - \int_t^u p^2 \left(d_w^v\right)^2 dw - \frac{p}{2}\int_t^u \left(d_w^v\right)^2 dw\right),$$

and thus Hölder's inequality implies that the left-hand side of (4.10) can be further estimated against

$$\begin{aligned} &\mathbb{E}\left[\left|\int_t^T \left|\mathcal{E}\left(\int_t^u d_w^v\, d\hat{W}_w\right)\right|^2 du\right|^{p/2}\right] \\ &\leq \int_t^T \left(\mathbb{E}\left[\exp\left(\int_t^u 2p d_w^v\, d\hat{W}_w - \frac{1}{2}\int_t^u 4p^2 \left(d_w^v\right)^2 dw\right)\right]\right)^{1/2} \\ &\quad \times \left(\mathbb{E}\left[\exp\left(\int_t^u 2p^2 \left(d_w^v\right)^2 dw - p\int_t^u \left(d_w^v\right)^2 dw\right)\right]\right)^{1/2} du \\ &\leq \int_t^T e^{p^2(T-t)D(v)} \left(\mathbb{E}\left[\mathcal{E}\left(\int_t^u 2p d_w^v\, d\hat{W}_w\right)\right]\right)^{1/2} du \\ &= (T-t)e^{p^2(T-t)D(v)}, \end{aligned}$$

which yields

$$\mathbb{E}\Big[\sup_{t\leq u\leq T} \left|b_u^{t,v}\right|^p\Big] \leq C_p D^p(v)(T-t)e^{(p+p^2)(T-t)D(v)},$$

whence we deduce the result for $p \geq 2$. For $1 \leq p < 2$, we use Jensen's inequality to obtain

$$\mathbb{E}\Big[\sup_{t\leq u\leq T} \left|b_u^{t,v}\right|^p\Big] \leq C\mathbb{E}\left[\left|\int_t^T \left|\mathcal{E}\left(\int_t^u d_w^{t,v}\, d\hat{W}_w\right)\right|^2 du\right|\right]^{p/2},$$

and continue with the same arguments as for $p > 2$. Hence the result follows. □

*Remark 4.6* Since $I^{t,y,v}$ is lognormally distributed, independently of $v$, we have

$$\sup_{v\in(-1,1)} \mathbb{E}\left[\left|I_s^{t,y,v}\right|^p\right] < \infty$$

for all $p \geq 1$. Let $K$ be a compact subset of $(-1, 1)$ and suppose $\sup_{v\in K} g'(\rho_s^{t,v})$ is bounded and $\sup_{v\in K} a'(\rho_s^{t,v})$ is bounded above, uniformly for all $t \leq s \leq T$. Then we have $\sup_{v\in K} \mathbb{E}[\int_t^T |\bar{\rho}_s^{t,v}|^p\, ds] < \infty$ for all $p \geq 1$, by Lemma 4.5.

## 4.3 Differentiability of $\psi$

In order to derive the hedge formula (4.7), we need to ensure that $\psi$ defined in (4.6) is continuously differentiable with respect to $v$ and $y$. We only consider the differentiability in $v$, since for $y$ it is comparatively simpler. Since we want to use uniform integrability, this is where Conditions (H1) and (H2) come into play.

**Lemma 4.7** *Suppose conditions* (H1) *and* (H2) *hold. Then for all* $v_0 \in (-1,1)$, *there exists an open neighbourhood* $U \subset (-1,1)$ *such that*

$$\sup_{v\in U} \mathbb{E}\big[\big|\bar{I}_T^{t,y,v}\big|^{p'}\big] < C \tag{4.11}$$

*for all* $p' \in [1,p)$, *with p from assumption* (H2). *Moreover, C is a constant that depends only on p from condition* (H2), *the model parameters and* $U$.

*Proof* Let $p' \geq 1$ be such that $p > p'$ with $p$ from assumption (H2). Let $G^{t,v}$ and $H^{t,y,v}$ be defined as in Lemma 4.3. Notice that $G^{t,v}$ is lognormally distributed. Since the distribution does not depend on the correlation, there exists a constant $C \in \mathbb{R}_+$ such that we have $\mathbb{E}[|\mathcal{E}(G^{t,v})|^{2p'}] \leq C$ for all $v \in (-1,1)$. The Cauchy–Schwarz inequality, the Burkholder–Davis–Gundy inequality and Jensen's inequality imply

$$\begin{aligned}
&\mathbb{E}\big[\big|\bar{I}_T^{t,y,v}\big|^{p'}\big]\\
&\quad= \mathbb{E}\bigg[\bigg|\mathcal{E}\big(G^{t,v}\big)_T \int_t^T \mathcal{E}\big(G^{t,v}\big)_u^{-1}\, dH_u^{t,y,v}\bigg|^{p'}\bigg]\\
&\quad\leq \sqrt{C}\bigg(\mathbb{E}\bigg[\sup_{t\leq s\leq T}\bigg|\int_t^s \mathcal{E}\big(G^{t,v}\big)_u^{-1}\, dH_u^{t,y,v}\bigg|^{2p'}\bigg]\bigg)^{1/2}\\
&\quad\leq \sqrt{C}C_{p'}\bigg(\mathbb{E}\bigg[\bigg|\int_t^T \mathcal{E}\big(G^{t,v}\big)_u^{-2}\big(I_u^{t,y,v}\big)^2\sigma_I^2\frac{(\bar{\rho}_u^{t,v})^2}{1-(\rho_u^{t,v})^2}\,du\bigg|^{p'}\bigg]\bigg)^{1/2}\\
&\quad\leq \sqrt{C}C_{p'}\bigg(\mathbb{E}\bigg[\bigg|\int_t^T \mathcal{E}\big(G^{t,v}\big)_u^{-2p'}\big(I_u^{t,y,v}\big)^{2p'}\sigma_I^{2p'}\bigg(\frac{(\bar{\rho}_u^{t,v})^2}{1-(\rho_u^{t,v})^2}\bigg)^{p'}du\bigg|\bigg]\bigg)^{1/2}.
\end{aligned}$$

Now choose $\hat{p} > 1$ such that $\hat{p}p' < p$. An application of the Hölder inequality yields, with $\hat{q} = \frac{\hat{p}}{\hat{p}-1}$,

$$\begin{aligned}
\big(\mathbb{E}\big[\big|\bar{I}_T^{t,y,v}\big|^{p'}\big]\big)^2 \leq CC_{p'}^2&\bigg(\mathbb{E}\bigg[\bigg|\int_t^T \mathcal{E}\big(G^{t,v}\big)_u^{-2p'\hat{q}}\big(I_u^{t,y,v}\big)^{2p'\hat{q}}\,du\bigg|\bigg]\bigg)^{1/\hat{q}}\\
&\times\bigg(\mathbb{E}\bigg[\bigg|\int_t^T\bigg(\frac{(\bar{\rho}_u^{t,v})^2}{1-(\rho_u^{t,v})^2}\bigg)^{p'\hat{p}}du\bigg|\bigg]\bigg)^{1/\hat{p}}.
\end{aligned}$$

For any $U \subset (-1,1)$, the term $\sup_{v\in U}\mathbb{E}[|\int_t^T \mathcal{E}(G^{t,v})_u^{-2p'\hat{q}}(I_u^{t,y,v})^{2p'\hat{q}}\,du|]$ is finite due to the lognormal distribution of $I^{t,y,v}$ and the normal distribution of $G^{t,v}$, distributions which do not depend on the correlation process $\rho$. Therefore, with $U$ from assumption (H2), we get $\sup_{v\in U}\mathbb{E}[|\bar{I}_T^{t,y,v}|^{p'}] < C$. □

The following lemma states conditions under which $\psi$ is differentiable with respect to $v$.

**Lemma 4.8** *Let $h$ be Lipschitz and such that the weak derivative $h'$ is Lebesgue-almost everywhere continuous. Under the conditions* (H1) *and* (H2), $\psi(t,y,v)$ *is continuously differentiable with respect to $v$ and the partial derivative $\partial_v\psi(t,y,v)$ is given by*

$$\partial_v\psi(t,y,v)=\tilde{\mathbb{E}}\big[h'\big(I_T^{t,y,v}\big)\bar{I}_T^{t,y,v}\big].$$

*Proof* Let $v_0$ be an element of $(-1,1)$ and $p>1$ as in condition (H2). According to Lemma 4.7, we can choose a real $\delta>0$ with $(v_0-\delta,v_0+\delta)\subset(-1,1)$ such that for all $p'\in[1,p)$, we have $\sup_{v\in(v_0-\delta,v_0+\delta)}\mathbb{E}[|\bar{I}_T^{t,y,v}|^{p'}]<\infty$. For all $v\in(v_0-\delta, v_0+\delta)$, we show that

1. $\psi(t,y,v)$ is well defined
2. $h(I_T^{t,y,v})$ is absolutely continuous in $v$
3. $\tilde{\mathbb{E}}[h'(I_T^{t,y,v})\bar{I}_T^{t,y,v}]$ is continuous at $v=v_0$, and
4. $\tilde{\mathbb{E}}[\int_{-\delta}^{\delta}|h'(I_T^{t,y,v_0+u})\bar{I}_T^{t,y,v_0+u}|du]<\infty$.

By standard arguments, these four statements imply the result (see for instance [4, Theorem A9.1]).

The properties of $h$ imply that there is a constant $C>0$ with $|h(x)|\le C(1+|x|)$, and hence with Remark 4.6 we have, with $q=\frac{p}{p-1}$,

$$\begin{aligned}\tilde{\mathbb{E}}\big[\big|h\big(I_T^{t,y,v}\big)\big|\big]&\le\mathbb{E}\big[\big|\Gamma_T^0\big|^q\big]^{1/q}\mathbb{E}\big[\big|h\big(I_T^{t,y,v}\big)\big|^p\big]^{1/p}\\&\le\mathbb{E}\big[\big|\Gamma_T^0\big|^q\big]^{1/q}\,2C\big(1+\mathbb{E}\big[I_T^{t,y,v}\mid{}^p\big]\big)^{1/p}\\&<\infty,\end{aligned}$$

and therefore $\psi$ is well defined.

Since $h$ is Lipschitz, it is absolutely continuous. Besides, $I_T^{t,y,v}$ is differentiable and continuous in $v$ (see Sect. 3), and consequently, the composition $h(I_T^{t,y,v})$ is absolutely continuous in $v$. With the Hölder inequality, we have for $p'\in[1,p)$ and $\hat{p}>1$ such that $p>p'\hat{p}>1$ that

$$\tilde{\mathbb{E}}\big[\big|h'\big(I_T^{t,y,v}\big)\bar{I}_T^{t,y,v}\big|^{p'}\big]\le C\big(\mathbb{E}\big[\big|\bar{I}_T^{t,y,v}\big|^{p'\hat{p}}\big]\big)^{1/\hat{p}}.$$

Thus, by Lemma 4.7, we get

$$\sup_{v\in(v_0-\delta,v_0+\delta)}\tilde{\mathbb{E}}\big[\big|h'\big(I_T^{t,y,v}\big)\bar{I}_T^{t,y,v}\big|^{p'}\big]\le C<\infty.$$

Hence, the family of random variables $(h'(I_T^{t,y,v})\bar{I}_T^{t,y,v})_{v\in(v_0-\delta,v_0+\delta)}$ is uniformly integrable with respect to $\tilde{\mathbb{P}}$. Now let $(v_n)_{n\in\mathbb{N}}$ be any sequence in $(v_0-\delta,v_0+\delta)$ with limit $v_0$. Then by continuity of $h'(I_T^{t,y,v})\bar{I}_T^{t,y,v}$ and uniform integrability, we get

$$\begin{aligned}&\lim_{n\to\infty}\big|\tilde{\mathbb{E}}\big[h'\big(I_T^{t,y,v_n}\big)\bar{I}_T^{t,y,v_n}\big]-\tilde{\mathbb{E}}\big[h'\big(I_T^{t,y,v_0}\big)\bar{I}_T^{t,y,v_0}\big]\big|\\&\quad\le\lim_{n\to\infty}\tilde{\mathbb{E}}\big[\big|h'\big(I_T^{t,y,v_n}\big)\bar{I}_T^{t,y,v_n}-h'\big(I_T^{t,y,v_0}\big)\bar{I}_T^{t,y,v_0}\big|\big]=0,\end{aligned}$$

i.e., the continuity of $\tilde{\mathbb{E}}[h'(I_T^{t,y,v})\bar{I}_T^{t,y,v}]$ at $v = v_0$. We use boundedness of $h'$ and Fubini's theorem to get

$$\tilde{\mathbb{E}}\left[\int_{-\delta}^{\delta}\left|h'\left(I_T^{t,y,v_0+u}\right)\bar{I}_T^{t,y,v_0+u}\right|du\right] \le C\int_{-\delta}^{\delta}\mathbb{E}\left[\left|\bar{I}_T^{t,y,v_0+u}\right|\right]du$$
$$\le C \sup_{v\in(v_0-\delta,v_0+\delta)}\mathbb{E}\left[\left|\bar{I}_T^{t,y,v}\right|\right],$$

which is finite by Lemma 4.7. Since we verified all four points above, the proof of Lemma 4.8 is complete. □

## 4.4 The hedge as variational derivative

The control process $Z^{t,y,v}$ of the linear BSDE (4.5) has a representation in terms of the gradient of $\psi$ and the matrix-valued function defined by

$$\sigma(y,v) = \begin{pmatrix} y\sigma_I v & y\sigma_I\sqrt{1-v^2} & 0 \\ g(v)\gamma & g(v)\delta & g(v)\sqrt{1-\gamma^2-\delta^2}\end{pmatrix}, \quad y\in\mathbb{R},\ v\in(-1,1).$$

**Lemma 4.9** *Assume that* (H1) *and* (H2) *hold, that a and g are continuously differentiable and let h be Lipschitz and such that the weak derivative $h'$ is Lebesgue-almost everywhere continuous. Then*

$$Z_s^{t,y,v} = \sigma\left(I_s^{t,y,v}, \rho_s^{t,v}\right)^* \begin{pmatrix}\partial_y\psi(s, I_s^{t,y,v}, \rho_s^{t,v}) \\ \partial_v\psi(s, I_s^{t,y,v}, \rho_s^{t,v})\end{pmatrix} \tag{4.12}$$

*Proof* By Lemma 4.8, $\psi(s,y,v)$ is continuously differentiable in $y$ and $v$. As is shown in the fundamental Theorem 5.2 in [10], considering also Remark 5.3.i of [10], this is sufficient for the relationship (4.12) to hold. □

Note that Imkeller et al. [10] establish relations as in (4.12) by using only elementary methods. However, up to now the standard method of deriving these relationships was to interpret $Z^{t,y,v}$ as the Malliavin derivative, or more precisely the Malliavin trace, of $Y^{t,y,v}$. Compared to the approach given in [10], this has the disadvantage that additional regularity assumptions are needed which originate in the use of the Malliavin calculus. Nevertheless we want to outline how Malliavin calculus can be used to derive (4.12), thus giving a proof of (4.12) in this paper (though not in full generality). Since this approach entails variational derivatives of the forward processes $I$ and $\rho$ (see (4.13)), we need the additional assumption that the coefficients $a$ and $g$ of the dynamics of $\rho$ have bounded derivatives.

*Malliavin-based proof of Lemma 4.9 under the additional assumptions that a and g have bounded derivatives:* Let $I^n$ be the solution of the SDE

$$dI_t^n = I_t^n(\mu_I - r)\,dt + \sigma_I I_t^n\left(\left(1-\frac{1}{n}\right)\rho_t\,dW_t^1 + \sqrt{1-\left(1-\frac{1}{n}\right)^2\rho^2}\,dW_t^2\right).$$

It is straightforward to show that $I_T^n$ converges to $I_T$ in $L^2$. By taking a subsequence, we may assume that $I_T^n$ converges to $I_T$ almost surely.

Next we approximate the payoff function $h$ by a sequence of everywhere differentiable and globally Lipschitz-continuous functions. More precisely, we denote by $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ the density of a standard normal distribution, and $\varphi^n(x) := n\varphi(nx)$ for all $n \geq 1$. We define $h^n$ as the convolution with $\varphi^n$, i.e., $h^n = h * \varphi^n$. Observe that $h^n$ is Lipschitz-continuous with the same Lipschitz constant as $h$. Note that Lipschitz-continuity of $h$ implies uniform convergence of $h^n$ to $h$; hence, $h^n(I_T^n)$ converges a.s. to $h(I_T)$. Moreover, $h^n$ is differentiable.

As before, we denote by $I^{n,t,y,v}$ the process $I^n$ conditioned on $I_t^n = y$ and $\rho_t = v$. We further define $\psi^n(t, y, v) = \tilde{\mathbb{E}}[h^n(I_T^{n,t,y,v})]$ for all $n \geq 1$, where $\tilde{\mathbb{E}}$ denotes the expectation with respect to the measure $\tilde{\mathbb{P}}$ defined in Sect. 3. Note that by the same methods as in Sect. 4.3, it can be shown that the $\psi^n$ are differentiable; indeed, due to the factor $(1 - \frac{1}{n})$, the integrability condition (4.11) is trivial. Moreover, the derivative of each $\psi^n$ is bounded, i.e., the $\psi^n$ are Lipschitz-continuous.

We proceed by showing that $\psi^n$ converges pointwise to $\psi$. Indeed, with $L \in \mathbb{R}_+$ being the Lipschitz constant of $h$, we have

$$
\begin{aligned}
&\lim_n \left|\psi^n(t, y, v) - \psi(t, y, v)\right| \\
&\quad \leq \lim_n \sqrt{\tilde{\mathbb{E}}\left[\left|h^n\left(I_T^{n,t,y,v}\right) - h\left(I_T^{t,y,v}\right)\right|^2\right]} \\
&\quad \leq \lim_n \sqrt{4\left(\tilde{\mathbb{E}}\left[\left|h^n\left(I_T^{n,t,y,v}\right) - h\left(I_T^{n,t,y,v}\right)\right|\right]^2 + \tilde{\mathbb{E}}\left[\left|h\left(I_T^{n,t,y,v}\right) - h\left(I_T^{t,y,v}\right)\right|^2\right]\right)} \\
&\quad \leq \lim_n \sqrt{4\left(\left\|h^n - h\right\|_\infty + L^2 \tilde{\mathbb{E}}\left[\left|I_T^n - I_T\right|^2\right]\right)} = 0.
\end{aligned}
$$

Let $(Y^{n,t,y,v}, Z^{n,t,y,v})$ be the solution of the BSDE

$$
Y_s^{n,t,y,v} = h^n\left(I_T^{n,t,y,v}\right) - \int_s^T Z_u^{n,t,y,v}\, dW_u - \int_s^T Z_u^{n,1,t,y,v} \frac{\mu_X - r}{\sigma_X}\, du
$$

for $s \in [t, T]$. Since $h^n(I_T^{n,t,y,v})$ converges to $h(I_T^{t,y,v})$ in $L^2(\mathbb{P})$, standard a priori estimates for Lipschitz BSDEs, or simply the Itô isometry under the measure $\tilde{\mathbb{P}}$, imply that $(Y^n, Z^n)$ converges to $(Y, Z)$ in $\mathbb{S}(\mathbb{R}) \otimes \mathbb{H}_T^2(\mathbb{R}^3)$.

Notice that due to the Markov property, we have

$$
Y_s^{n,t,y,v} = \tilde{\mathbb{E}}\left[h^n\left(I_T^{n,t,y,v}\right) \mid \mathcal{F}_s\right] = \psi^n\left(s, I_s^{n,t,y,v}, \rho_s^{t,v}\right).
$$

Since the approximations $\psi^n$ are Lipschitz-continuous, we may apply the chain rule, which yields

$$
D_u Y_s^{n,t,y,v} = \partial_y \psi^n\left(s, I_s^{n,t,y,v}, \rho_s^{t,v}\right) D_u I_s^{n,t,y,v} + \partial_v \psi^n\left(s, I_s^{n,t,y,v}, \rho_s^{t,v}\right) D_u \rho_s^{t,v}, \tag{4.13}
$$

$u \in [t, T]$, where $D_u$ denotes the Malliavin derivative of $Y^{n,t,y,v}$, $I^{n,t,y,v}$ and $\rho^{t,v}$ respectively. $D_u I^{n,t,y,v}$ and $D_u \rho^{t,v}$ are solutions of linear SDEs (see Theorem 2.2.1 in [13]). In particular, this guarantees right-continuity of $D_u Y_s^{n,t,y,v}$ in $s$.

By the Clark–Ocone formula, the process $Z^{n,t,y,v}$ is the predictable projection of $h^n(I_T^{n,t,y,v})$ under the measure $\tilde{\mathbb{P}}$. More precisely, for all $s \in [t,T]$, we have $Y_s^{n,t,y,v} = \int_t^s \tilde{\mathbb{E}}[D_u Y_s^{n,t,y,v} \mid \mathcal{F}_u]\, d\tilde{W}_u$, where $\tilde{W}_t = (W_t^1 + \frac{\mu_X - r}{\sigma_X} t, W_t^2, W_t^3)$ and where $\tilde{\mathbb{E}}[\cdot \mid \mathcal{F}_u]$ stands for the predictable projection operator with respect to $\tilde{\mathbb{P}}$. Due to the right-continuity of $D_u Y_s^{n,t,y,v}$ in $s$, we may interchange the Malliavin derivative and the predictable projection operator, which yields

$$
\begin{aligned}
Z_u^{n,t,y,v} &= \tilde{\mathbb{E}}\big[D_u Y_s^{n,t,y,v} \mid \mathcal{F}_u\big] = \lim_{s \downarrow u} D_u Y_s^{n,t,y,v} \\
&= \sigma^n\big(u, I_u^{n,t,y,v}, \rho_u^{t,v}\big)^* \begin{pmatrix} \partial_y \psi^n(u, I_u^{n,t,y,v}, \rho_u^{t,v}) \\ \partial_v \psi^n(u, I_u^{n,t,y,v}, \rho_u^{t,v}) \end{pmatrix},
\end{aligned}
\tag{4.14}
$$

where

$$
\sigma^n(y,v) = \begin{pmatrix} y\sigma_I(1-\frac{1}{n})v & y\sigma_I\sqrt{1-(1-\frac{1}{n})^2 v^2} & 0 \\ g(v)\gamma & g(v)\delta & g(v)\sqrt{1-\gamma^2-\delta^2} \end{pmatrix}.
$$

We next show that the partial derivatives $\partial_y \psi^n$ and $\partial_v \psi^n$ converge pointwise to $\partial_y \psi$ and $\partial_v \psi$, respectively. To this end, denote again the derivatives of $I_t^{n,t,y,v}$ with respect to $v$ by $\bar{I}_t^{n,t,y,v}$. Lemma 4.7 yields that $\sup_n \mathbb{E}[|\bar{I}_T^n|]^p < \infty$, which further implies that the sequence $|\bar{I}_T^n|$ is uniformly integrable. Moreover,

$$
\begin{aligned}
\big|\partial_v \psi^n - \partial_v \psi\big| &\le \tilde{\mathbb{E}}\big[\big|(h^n)'\big(I_T^{n,t,y,v}\big)\bar{I}_T^{n,t,y,v} - h'\big(I_T^{t,y,v}\big)\bar{I}_T^{t,y,v}\big|\big] \\
&\le \tilde{\mathbb{E}}\big[\big|(h^n)'\big(I_T^{n,t,y,v}\big)\big|\big|\bar{I}_T^{n,t,y,v} - \bar{I}_T^{t,y,v}\big|\big] \\
&\quad + \tilde{\mathbb{E}}\big[\big|\bar{I}_T^{t,y,v}\big|\big|(h^n)'\big(I_T^{n,t,y,v}\big) - h'\big(I_T^{t,y,v}\big)\big|\big].
\end{aligned}
\tag{4.15}
$$

We show separately that both summands of (4.15) converge to 0 as $n \to \infty$. Since the approximating functions $h^n$ have one common Lipschitz constant $L \in \mathbb{R}_+$, the derivatives satisfy $|(h^n)'| \le L$ for all $n \ge 1$. Consequently,

$$
\lim_n \tilde{\mathbb{E}}\big[\big|(h^n)'\big(I_T^{n,t,y,v}\big)\big|\big|\bar{I}_T^{n,t,y,v} - \bar{I}_T^{t,y,v}\big|\big] \le L \lim_n \tilde{\mathbb{E}}\big[\big|\bar{I}_T^{n,t,y,v} - \bar{I}_T^{t,y,v}\big|\big].
$$

Next, let $\tau_k = T \wedge \inf\{t \ge 0 : |\rho_t| = (1 - \frac{1}{k})\}$ for all $k \ge 1$. Then the stopped processes $\bar{I}_{t\wedge\tau_k}^n$ converge to $\bar{I}_{t\wedge\tau_k}$ in $L^2$ as $n \to \infty$ (see [14, Theorem V.4.11]). Therefore, by dominated convergence, for every $k \ge 1$ we have

$$
\begin{aligned}
\lim_n \tilde{\mathbb{E}}\big[\big|\bar{I}_T^{n,t,y,v} - \bar{I}_T^{t,y,v}\big|\big] &\le \lim_n \big(\tilde{\mathbb{E}}\big[\big|1_{\{\tau_k < T\}}\big(\bar{I}_T^{n,t,y,v} - \bar{I}_T^{t,y,v}\big)\big|\big] \\
&\quad + \tilde{\mathbb{E}}\big[\big|1_{\{\tau_k \ge T\}}\big(\bar{I}_T^{n,t,y,v} - \bar{I}_T^{t,y,v}\big)\big|\big]\big) \\
&\le \tilde{\mathbb{E}}\big[1_{\{\tau_k < T\}}\big(\big|\bar{I}_T^{n,t,y,v}\big| + \tilde{\mathbb{E}}\big|\bar{I}_T^{t,y,v}\big|\big)\big].
\end{aligned}
$$

Recall that $\bar{I}_t^n$ is uniformly integrable and that $\lim_k \tilde{\mathbb{P}}[\tau_k = T] = 1$; hence, by letting $k \to \infty$, we get that $\lim_n \tilde{\mathbb{E}}[|\bar{I}_T^{n,t,y,v} - \bar{I}_T^{t,y,v}|] = 0$, and hence the first summand in (4.15) converges to 0.

In order to show that the second summand in (4.15) vanishes, we first show that for $x_n \to x$, we have $(h^n)'(x_n) \to h'(x)$. If $x$ is a point of continuity of $h'$, then we have, using $z := y - x_n + x$, the estimate

$$\begin{aligned}
\big|(h^n)'(x_n) - h'(x)\big| &= \left|\int \varphi_n(x_n - y)h'(y)\,dy - h'(x)\right| \\
&= \left|\int \varphi_n(x_n - x + x - y)h'(y)\,dy - h'(x)\right| \\
&= \left|\int \varphi_n(x - z)h'(z + x_n - x)\,dz - h'(x)\right| \\
&\le \left|\int \varphi_n(x - z)h'(z)\,dz - h'(x)\right| \\
&\quad + \left|\int \varphi_n(x - z)\big[h'(z + x_n - x) - h'(z)\big]\,dz\right|.
\end{aligned}$$

Applying the transformation $y := n(x - z)$ in each term on the right-hand side of the inequality, together with dominated convergence and the continuity of $h'$ in $x$, yields that $\lim_n h'(x_n) = h'(x)$. Since $I_T$ has a density, $(h^n)'(I_T^{n,t,y,v})$ converges to $h'(I_T^{t,y,v})$ almost everywhere. So, dominated convergence gives

$$\lim_n \tilde{\mathbb{E}}\big[\big|\bar{I}_T^{t,y,v}\big|\big|(h^n)'\big(I_T^{n,t,y,v}\big) - h'\big(I_T^{t,y,v}\big)\big|\big] = 0.$$

Thus we have shown that $\lim_n \partial_v \psi^n(t, y, v) = \partial_v \psi(t, y, v)$ for all $t \in [0, T]$, $y \in \mathbb{R}$ and $v \in (-1, 1)$.

Notice that $I_s^{n,t,y,v} = y I_s^{n,t,1,v}$ and $\bar{I}_s^{n,t,y,v} = y \bar{I}_s^{n,t,1,v}$, which implies that for all $t \ge 0$ and $v \in (-1, 1)$, the sequence $\partial_v \psi^n(t, \cdot, v)$ converges to $\partial_v \psi(t, \cdot, v)$ uniformly in $y$ on all compact sets of $\mathbb{R}$. Similarly, one can show locally uniform convergence in $y$ of the partial derivatives $\partial_y \psi^n$ to $\partial_y \psi$. This finally yields that $\partial_v \psi^n(s, I_s^{n,t,y,v}, \rho_s^{t,v})$ converges to $\partial_v \psi(s, I_s^{t,y,v}, \rho_s^{t,v})$, and $\partial_y \psi^n(s, I_s^{n,t,y,v}, \rho_s^{t,v})$ to $\partial_y \psi(s, I_s^{t,y,v}, \rho_s^{t,v})$, almost surely. Moreover, by combining this with (4.14), we get that $Z^n$ converges almost surely to $\sigma(s, I_s^{t,y,v}, \rho_s^{t,v})^* \binom{\partial_y \psi(s, I_s^{t,y,v}, \rho_s^{t,v})}{\partial_v \psi(s, I_s^{t,y,v}, \rho_s^{t,v})}$. Since $Z^n$ converges also to $Z$ in $\mathbb{H}^2$, we obtain the result.

## 5 A class of correlation dynamics which fulfil the main assumptions

In this part of the work, we characterise a class of dynamics which fulfil conditions (H1) and (H2). The result has already been mentioned as Theorem 3.5 in Sect. 3 above. For ease of reference we state it again below. Moreover, we give an extension of Theorem 3.5 in Proposition 5.3. We shall see in the example section that this extension enables us to show that the so-called Jacobi processes also fit into our framework. In contrast to the dynamics given in Theorem 3.5, the diffusion coefficient of a Jacobi process has unbounded derivatives in $-1$ and $1$.

We first collect some notation and facts on attainability of boundaries for diffusions. The material is taken from [11]. Suppose we are given a general diffusion

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t, \quad \ell \le X_0 \le r,$$

where $\ell$ (resp. $r$) denote the left (resp. right) boundary. In the following, we only consider the analysis for the left boundary. We define, for $x \in (\ell, r)$,

$$\begin{aligned} s(v) &= \exp\left(-\int_{v_0}^{v} \frac{2\mu(w)}{\sigma^2(w)}\,dw\right), \quad v_0 \in (\ell, x), \\ S(x) &= \int_{x_0}^{x} s(v)\,dv, \quad x_0 \in (\ell, x), \\ S[c,d] &= S(d) - S(c), \quad (c,d) \in (\ell, r), \\ S(\ell, x] &= \lim_{c\to\ell} S[c,x]. \end{aligned}$$

We already indicate that $x_0$ and $v_0$ will be of no relevance in the following. $S$ is called the *scale measure* whereas $M$ is the *speed measure*, defined via

$$\begin{aligned} m(x) &= \frac{1}{\sigma^2(x)s(x)}, \\ M[c,d] &= \int_c^d m(x)\,dx. \end{aligned}$$

We also need

$$\Sigma(\ell) = \lim_{c\to\ell} \int_c^x M[v,x]\,dS(v).$$

According to [11], the boundary $\ell$ is attracting if $S(\ell, x] < \infty$, and this criterion applies independently of $x \in (\ell, r)$. Moreover the boundary $\ell$ is said to be

1. attainable if $\Sigma(\ell) < \infty$
2. unattainable if $\Sigma(\ell) = \infty$.

For a proof of the following lemma, see Chap. 15.6 in [11].

**Lemma 5.1** $S(\ell, x] = \infty$ *implies* $\Sigma(\ell) = \infty$.

With this at hand, we can find sufficient conditions on the coefficients of the correlation dynamics so that condition (H1) is satisfied. Let again $\rho^{t,v}$ and $\bar{\rho}^{t,v}$ be defined as in the SDEs (3.4) and (3.6). To simplify notation, we suppress from now on the dependence on $t$ and $v$ and only write $\rho$ resp. $\bar{\rho}$.

**Lemma 5.2** *Let $a$ and $g$ be continuously differentiable. We assume that $g$ does not have any roots in* $(-1, 1)$. *If*

$$\limsup_{x\uparrow 1} \frac{2a(x)(1-x)}{g^2(x)} < 0 \quad \textit{and} \quad \liminf_{x\downarrow -1} \frac{2a(x)(1+x)}{g^2(x)} > 0, \tag{5.1}$$

*then condition* (H1) *is satisfied.*

*Proof* We show that $\rho$ does not reach $-1$. By condition (5.1) there exist $\varepsilon > 0$, $\delta > 0$ and $v_0 \in (-1, 1)$ such that $\frac{2a(w)}{g^2(w)} \geq \frac{\varepsilon}{1+w}$, for all $-1 < w < v_0 < -1 + \delta$. Hence,

$$s(v) = \exp\left(\int_v^{v_0} \frac{2a(w)}{g^2(w)} dw\right) \geq \exp\left(\int_v^{v_0} \frac{\varepsilon}{1+w} dw\right)$$
$$= C \exp\big(-\log(1+v)\big) = C \frac{1}{1+v}.$$

Hence,

$$S[c, x] \geq C \int_c^x \frac{1}{1+v} dv \longrightarrow \infty$$

for $c \to -1$, i.e., by Lemma 5.1 we obtain that $\rho$ does not reach $-1$. We treat the boundary 1 similarly, and hence get assumption (H1). □

The next proposition provides conditions under which condition (H2) is satisfied. We need two auxiliary processes $\tilde{a}$ and $\tilde{g}$ defined by

$$\tilde{a}_u = \frac{2\rho_u}{(1-\rho_u^2)} g(\rho_u) + 2g'(\rho_u),$$
$$\tilde{g}_u = \frac{2\rho_u}{(1-\rho_u^2)} a(\rho_u) + 2a'(\rho_u) + \frac{g^2(\rho_u)}{1-\rho_u^2} + \frac{4\rho_u^2 g^2(\rho_u)}{(1-\rho_u^2)^2}$$
$$+ \big(g'(\rho_u)\big)^2 + \frac{4\rho_u}{1-\rho_u^2} g(\rho_u) g'(\rho_u).$$

**Proposition 5.3** *Assume the conditions of Lemma 5.2 are satisfied. Then assumption* (H1) *holds and, therefore, $\tilde{a}$ and $\tilde{g}$ are well defined. Suppose $\tilde{a}$ is bounded and $\tilde{g}$ is bounded from above. Then assumption* (H2) *is satisfied, and hence the delta hedge is given as in Theorem 3.1.*

*Proof* We start by applying Itô's formula to the process $\Phi_s = f(\rho_s, \bar{\rho}_s)$, where $f$ is given by $f(x, y) = \frac{y^2}{1-x^2}$. Note that $f_x(x, y) = \frac{2xy^2}{(1-x^2)^2}$, $f_{xx} = \frac{2y^2}{(1-x^2)^2} + \frac{8x^2y^2}{(1-x^2)^3}$, $f_y(x, y) = \frac{2y}{1-x^2}$, $f_{yy} = \frac{2}{1-x^2}$ and $f_{xy} = \frac{4xy}{(1-x^2)^2}$. We have

$$\Phi_s = \Phi_t + \int_t^s \frac{2\rho_u \bar{\rho}_u^2}{(1-\rho_u^2)^2} \big[a(\rho_u)\, du + g(\rho_u)\, d\hat{W}_u\big]$$
$$+ \int_t^s \frac{2\bar{\rho}_u}{1-\rho_u^2} \big[a'(\rho_u)\bar{\rho}_u\, du + g'(\rho_u)\bar{\rho}_u\, d\hat{W}_u\big]$$
$$+ \frac{1}{2}\int_t^s \left[\frac{2\bar{\rho}_u^2}{(1-\rho_u^2)^2} + \frac{8\rho_u^2\bar{\rho}_u^2}{(1-\rho^2)^3}\right] g^2(\rho_u)\, du + \frac{1}{2}\int_t^s \frac{2}{1-\rho_u^2}\big(g'(\rho_u)\big)^2 \bar{\rho}_u^2\, du$$
$$+ \int_t^s \frac{4\rho_u\bar{\rho}_u}{(1-\rho^2)^2} g(\rho_u) g'(\rho_u)\bar{\rho}_u\, du$$

$$= \Phi_t + \int_t^s \Phi_u \left[ \frac{2\rho_u}{(1-\rho_u^2)} g(\rho_u) + 2g'(\rho_u) \right] d\hat{W}_u$$
$$+ \int_t^s \Phi_u \left[ \frac{2\rho_u}{(1-\rho_u^2)} a(\rho_u) + 2a'(\rho_u) + \frac{g^2(\rho_u)}{1-\rho_u^2} + \frac{4\rho_u^2 g^2(\rho_u)}{(1-\rho_u^2)^2} \right.$$
$$\left. + \big(g'(\rho_u)\big)^2 + \frac{4\rho_u}{1-\rho_u^2} g(\rho_u) g'(\rho_u) \right] du.$$

Thus, $\Phi$ is the solution of a linear stochastic equation and given by

$$\Phi_s = \Phi_t \mathcal{E}\bigg( \int_t^s \tilde{a}_u \, d\hat{W}_u + \int_t^s \tilde{g}_u \, du \bigg).$$

Hence by our assumptions on $\tilde{a}$ and $\tilde{g}$, and by Lemma 4.5, all moments of $\sup_{t \le u \le T} \Phi_u$ are finite, which yields assumption (H2). □

We use the two preceding statements to prove Theorem 3.5, whose statement we recall here.

**Theorem 3.5** *Let $a$ and $g$ be continuously differentiable with bounded derivatives. We assume that $g(-1) = g(1) = 0$, and that $g$ does not have any roots in $(-1, 1)$. If*

$$\limsup_{x \uparrow 1} \frac{2a(x)(1-x)}{g^2(x)} < 0 \quad \text{and} \quad \liminf_{x \downarrow -1} \frac{2a(x)(1+x)}{g^2(x)} > 0,$$

*then conditions* (H1) *and* (H2) *are satisfied, and hence the delta hedge is given as in Theorem* 3.1.

*Proof* Condition (H1) follows from Lemma 5.2. Since 1 and $-1$ are roots of $g$, we can write $\frac{g(x)}{1-x^2} = \frac{1}{1+x} \frac{g(x)-g(1)}{1-x} = \frac{-1}{1+x} \frac{g(x)-g(1)}{x-1}$ and hence $\frac{g(x)}{1-x^2}$ is bounded for $x \nearrow 1$ by the derivative of $g$ at $x = 1$. We argue similarly for $x \searrow -1$, i.e., the fraction $\frac{g(x)}{1-x^2}$ is bounded on $[-1, 1]$. Moreover, Condition (5.1) implies that there exists an $\varepsilon \in (0, 1)$ such that all $x \in (-1, 1)$ with $|x| \ge 1 - \varepsilon$ satisfy $xa(x) < 0$. Hence, $\tilde{a}$ (resp. $\tilde{g}$) is bounded (resp. bounded from above), and therefore we obtain the result by Proposition 5.3. □

*Remark 5.4* (1) Note that the conditions on the coefficients $a$ and $g$ of the correlation dynamics in Theorem 3.5 are more restrictive than in Proposition 5.3. This is mainly for ease of exposition in Sect. 3. In Sect. 6.1, an example is given where the coefficient $g$ of the correlation dynamics does not have a bounded derivative.

(2) It is possible to prove Theorem 3.5 without considering the auxiliary processes $\tilde{a}$ and $\tilde{g}$ and using Proposition 5.3. In the following we give a rough sketch of a more intuitive proof of Theorem 3.5. That alternative proof consists in showing that all moments of the process $Y_t = \frac{1}{1-\rho_t^2}$ are finite, from which one easily deduces condition (H2) to be satisfied. From Itô's formula, we obtain

$$dY_t = \frac{2\rho_t}{(1-\rho_t^2)} g(\rho_t) Y_t \, d\hat{W}_t + 2\rho_t a(\rho_t) Y_t^2 dt + \big(1 + 3\rho_t^2\big) \frac{g^2(\rho_t)}{(1-\rho_t^2)^2} Y_t \, dt,$$

so that $Y$ is a linear SDE with an additional drift term growing quadratically in $Y$. Condition (3.9) implies that there is an $\varepsilon \in (0,1)$ such that all $x \in (-1,1)$ with $|x| \geq 1-\varepsilon$ satisfy $xa(x) < 0$. Moreover, $\{|\rho_s| \leq 1-\varepsilon\} = \{Y_s \leq \frac{1}{2\varepsilon-\varepsilon^2}\}$, and consequently, the quadratic drift term in the dynamics of $Y$ has a shrinking effect as soon as $Y$ exceeds $C_\varepsilon = \frac{1}{2\varepsilon-\varepsilon^2}$. In other words, $Y$ can be shown to be dominated by the SDE

$$d\check{Y}_t = \frac{2\rho_t}{(1-\rho_t^2)} g(\rho_t)\check{Y}_t \, d\hat{W}_t + \left(1+3\rho_t^2\right)\frac{g^2(\rho_t)}{(1-\rho_t^2)^2}\check{Y}_t \, dt + C_\varepsilon \, dt,$$

which by standard arguments can be shown to possess finite moments.

## 6 Examples

The aim of this final section is to give some explicit correlation dynamics which fall within the framework above. We start by modelling correlation processes directly as solutions of various SDEs with values in $[-1,1]$ in Sect. 6.1. Another approach is used in Sect. 6.2 where we use mappings of an Ornstein–Uhlenbeck process onto the open interval $(-1,1)$.

### 6.1 Modelling correlation directly

*Example 6.1* Of course, all processes that are bounded away from $-1$ and 1 also fulfil the conditions (H1) and (H2).

*Example 6.2* For $a(x) = \kappa(\vartheta - x)$ with $\kappa > 0$, $\vartheta \in (-1,1)$, and $g(x) = \alpha(1-x^2)$ in the dynamics of $\rho$, the prerequisites of Theorem 3.5 are fulfilled.

*Example 6.3* Let $a$ and $g$ be polynomials. Assume that $g(-1) = g(1) = 0$, and that $g$ does not have any roots in $(-1,1)$. If

$$\lim_{x\uparrow 1} \frac{a(x)}{g^2(x)} = -\infty \quad \text{and} \quad \lim_{x\downarrow -1} \frac{a(x)}{g^2(x)} = +\infty,$$

then the prerequisites of Theorem 3.5 are satisfied.

The common feature of the preceding two examples is that the coefficients in the dynamics of $\rho$ fulfil the prerequisites of Theorem 3.5, which include bounded derivatives. We now want to give an example where $g$ does not have bounded derivatives in $-1$ and 1. We consider so-called Jacobi processes, which are given by the solution of

$$\rho_s^{t,v} = v + \int_t^s \kappa\left(\vartheta - \rho_u^{t,v}\right) du + \int_t^s \alpha\sqrt{1-\left(\rho_s^{t,v}\right)^2}\, \hat{W}_u. \tag{6.1}$$

Jacobi processes might be of interest for modelling stochastic correlation, because their stationary and transitional densities are well known and can be obtained quite explicitly; see for example [2].

By exploiting the boundary theory at the beginning of Sect. 5 or by checking when condition (3.9) holds, one can easily show that for $\kappa, \alpha > 0$ and $\vartheta$ such that

$$\kappa \geq \frac{\alpha^2}{1 \pm \vartheta}, \tag{6.2}$$

the boundaries $-1$ and $1$ of the process in (6.1) are unattainable. Hence we have that assumption (H1) is fulfilled. We want to apply Proposition 5.3 and therefore have to check the boundedness of $\tilde{a}$ and upper boundedness of $\tilde{g}$. Note that $\tilde{g}$ turns into

$$\begin{aligned}\tilde{g}_u &= \frac{2\rho_u}{(1-\rho_u^2)} a(\rho_u) + 2a'(\rho_u) + \frac{\alpha^2}{1-\rho_u^2} \\ &= \frac{2\rho_u}{(1-\rho_u^2)} \big(\kappa(\vartheta - \rho_u)\big) - 2\kappa + \frac{\alpha^2}{1-\rho_u^2},\end{aligned}$$

and $\tilde{a}_u = 0$. In order to ensure upper boundedness of $\tilde{g}_u$ it is sufficient to show the existence of an $\varepsilon > 0$ such that $2\rho_u(\kappa(\vartheta - \rho_u)) + \alpha^2 < 0$ for all $|\rho_u| > 1 - \varepsilon$, $\mathbb{P}$-a.s. This is guaranteed by choosing $\kappa, \alpha > 0$ and $\vartheta$ such that the constants $\rho_{(1)}$ and $\rho_{(2)}$ defined by

$$\rho_{(1),(2)} = \frac{\vartheta}{2} \pm \sqrt{\frac{\vartheta^2}{4} + \frac{\alpha^2}{2\kappa}}$$

fulfil

$$-1 < \rho_{(1)} \leq \rho_{(2)} < 1. \tag{6.3}$$

Note for example that for $\alpha = 1$ and $\vartheta = 0.9$, condition (6.2) is satisfied by $\kappa = 10$ and that this choice of parameters also fulfils (6.3).

### 6.2 Modelling correlation with Ornstein–Uhlenbeck processes

In the previous section, we assumed that the stochastic correlation process is described in terms of the SDE (3.1). The correlation dynamics need not be modelled directly. Alternatively, one can use a continuous bijection $b : (-1, 1) \to \mathbb{R}$, and model first the transformed process $b(\rho)$ by an SDE. This has the advantage that $b(\rho)$ can be modelled as a diffusion on $\mathbb{R}$ with Lipschitz coefficients. The correlation may be modelled as a standard mean-reverting process, for example, an Ornstein–Uhlenbeck process, the dynamics of which can be calibrated via standard methods.

In this section, we discuss this alternative approach of modelling correlation. As a paradigm example, we choose as bijection $b(x) = \frac{x}{\sqrt{1-x^2}}$, and we assume that $U = b(\rho)$ is an Ornstein–Uhlenbeck process with dynamics

$$dU_t = a(\vartheta - U_t)\,dt + \sigma_U d\Big(\gamma\, dW_t^1 + \delta\, dW_t^2 + \sqrt{1-\gamma^2-\delta^2}\, dW_t^3\Big),$$

where $a > 0$, $\vartheta \in \mathbb{R}$, $\sigma_U > 0$ and $\gamma, \delta \in (-1, 1)$ are such that $\gamma^2 + \delta^2 \leq 1$. Notice that $\rho_t = \frac{U_t}{\sqrt{1+U_t^2}}$. We prove that the prerequisites of Theorem 3.1 are satisfied, and hence that the locally risk minimising strategy is defined as in (3.8).

**Lemma 6.4** *The correlation process* $\rho$ *satisfies conditions* (H1) *and* (H2), *and hence Theorem* 3.1 *holds in this setting*.

*Proof* The proof is a simple application of Itô's formula. The first and second derivatives of $b^{-1}: \mathbb{R} \to ]-1,1[, x \mapsto \frac{x}{\sqrt{1+x^2}}$ are given by $(b^{-1}(x))' = (1+x^2)^{-\frac{3}{2}}$ and $(b^{-1}(x))'' = -3x(1+x^2)^{-\frac{5}{2}}$. Again, we set $\hat{W}_t = \gamma W_t^1 + \delta W_t^2 + \sqrt{1-\gamma^2-\delta^2} W_t^3$. We obtain

$$d\rho_t = \left(1-\rho_t^2\right)^{\frac{3}{2}} \sigma_U \, d\hat{W}_t + \left(1-\rho_t^2\right)\left(a\vartheta\left(1-\rho_t^2\right)^{\frac{1}{2}} - a\rho_t - \frac{3}{2}\rho_t\left(1-\rho_t^2\right)\sigma_U^2\right) dt.$$

It is straightforward to show that the coefficients of this SDE satisfy the conditions of Theorem 3.5. Hence, the result follows. □

**Acknowledgements** We would like to thank an anonymous referee and an anonymous associate editor for their detailed and helpful feedback to the first submitted version. This work arose partially at Quantitative Products Laboratory Berlin and the Deutsche Bank chair of applied financial mathematics at Humboldt University. Financial support from Deutsche Bank is gratefully acknowledged.

## References

1. Ankirchner, S., Imkeller, P., Reis, G.: Pricing and hedging of derivatives based on non-tradeable underlyings. Math. Finance **20**, 289–312 (2010)
2. Boortz, C.: Modelling correlation risk. Master thesis. Technische Universität Berlin (2008). http://www.qpl.db.com/en/binaer_view.asp?binaernr=83
3. Davis, M.H.A.: Optimal hedging with basis risk. In: Kabanov, R.L.Y., Stoyanov, J. (eds.) From Stochastic Calculus to Mathematical Finance, pp. 169–187. Springer, Berlin (2006)
4. Durrett, R.: Probability: Theory and Examples, 2nd edn. Duxbury Press, Belmont (1996)
5. El Karoui, N., Peng, S., Quenez, M.C.: Backward stochastic differential equations in finance. Math. Finance **7**, 1–71 (1997)
6. Frei, C., Schweizer, M.: Exponential utility indifference valuation in two Brownian settings with stochastic correlation. Adv. Appl. Probab. **40**, 401–423 (2008)
7. Frei, C., Schweizer, M.: Exponential utility indifference valuation in a general semimartingale model. In: Delbaen, F., Rásonyi, M., Stricker, C. (eds.) Optimality and Risk—Modern Trends in Mathematical Finance. The Kabanov Festschrift, pp. 49–86. Springer, Berlin (2009)
8. Henderson, V.: Valuation of claims on nontraded assets using utility maximization. Math. Finance **12**, 351–373 (2002)
9. Hulley, H., McWalter, T.A.: Quadratic hedging of basis risk. Research Paper Series 225, Quantitative Finance Research Centre, University of Technology, Sydney, June 2008. http://www.business.uts.edu.au/qfrc/research/research_papers/rp225.pdf
10. Imkeller, P., Réveillac, A., Richter, A.: Differentiability of quadratic BSDEs generated by continuous martingales and hedging in incomplete markets. arXiv:0907.0941
11. Karlin, S., Taylor, H.M.: A Second Course in Stochastic Processes. Academic Press, New York (1981)
12. Monoyios, M.: Performance of utility-based strategies for hedging basis risk. Quant. Finance **4**, 245–255 (2004)
13. Nualart, D.: The Malliavin Calculus and Related Topics, 2nd edn. Probability and its Applications. Springer, Berlin (2006)
14. Protter, P.E.: Stochastic Integration and Differential Equations, 2nd edn. Applications of Mathematics, vol. 21. Springer, Berlin (2004). Stochastic Modelling and Applied Probability
15. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Grundlehren der Mathematischen Wissenschaften, vol. 293. Springer, Berlin (1999)

16. Schweizer, M.: A guided tour through quadratic hedging approaches. In: Jouini, E., Cvitanić, J., Musiela, M. (eds.) Option Pricing, Interest Rates and Risk Management, pp. 538–574. Cambridge Univ. Press, Cambridge (2001)
17. Schweizer, M.: Local risk-minimization for multidimensional assets and payment streams. Banach Cent. Publ. **83**, 213–229 (2008)
18. van Emmerich, C.: Modelling correlation as a stochastic process. Technical report, Bergische Universität Wuppertal, March 2006. http://www-num.math.uni-wuppertal.de//Files/amna_06_03.pdf