Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 122 (2012) 600–622

stochastic processes and their applications

www.elsevier.com/locate/spa

Hedging electricity swaptions using partial integro-differential equations

Peter Hepperger

Zentrum Mathematik, Technische Universität München, Chair of Mathematical Statistics, 85748 Garching, Germany

Received 20 November 2010; received in revised form 28 May 2011; accepted 9 September 2011
Available online 22 September 2011

Abstract

The basic contracts traded on energy exchanges are swaps involving the delivery of electricity for fixed-rate payments over a certain period of time. The main objective of this article is to solve the quadratic hedging problem for European options on these swaps, known as electricity swaptions. We consider a general class of Hilbert space valued exponential jump-diffusion models. Since the forward curve is an infinite-dimensional object, but only a finite set of traded contracts are available for hedging, the market is inherently incomplete. We derive the optimization problem for the quadratic hedging problem under the risk neutral measure and state a representation of its solution, which is the starting point for numerical algorithms.

MSC: 60H30; 91B25; 35R15

Keywords: Swaptions; Quadratic hedging; Hilbert space valued jump-diffusion; Infinite-dimensional stochastic analysis; Partial integro-differential equation

1. Introduction

During the last two decades, energy markets all over the world have been liberalized. Electricity is now traded liquidly on exchanges like the Scandinavian Nordpool and the German Energy Exchange (EEX). These relatively young markets are open to producers, consumers, and speculating investors. Traded products include spot, futures, forwards and options on these. In order to price these contracts and develop corresponding hedging strategies, mathematical

E-mail address: hepperger@ma.tum.de.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2011.09.005

models are called for. A compendium of methods for electricity markets can be found in [6]. There are substantial differences between stock and electricity markets. The electricity spot price exhibits several unique stylized features, including seasonality, large jumps (many times higher than the average price), and mean reversion. In addition, while stocks are sold at a single point in time, electricity contracts always imply delivery over a certain period of time. Therefore, electricity forwards and derivatives are written on a delivery period (a week, a month, or even a year). The following part describes the contracts and mathematical objects occurring in this context, which will be the basis for our model.

1.1. Electricity swaps and swaptions

The most liquidly traded products on energy exchanges like EEX or Nordpool are contracts of futures type. These are agreements traded at time $t \geq 0$ for a constant delivery of 1 MW of electricity over a certain future period of time $[T_1, T_2]$, while in return a fixed rate $F(t; T_1, T_2)$ is paid during this delivery period. Since a payment of a fixed rate is made in exchange for the (unknown) future spot price, these contracts are also known as *electricity swaps*. The relation of spot and forward prices is not clearly defined for electricity because of its non-storability [2,5]. This difficulty can be avoided by directly modeling the forward curve under a risk neutral (with respect to swap rates) measure [1,3,18]. For every maturity $u \in [T_1, T_2]$, let

$$f(t, u) := \lim_{v \to u} F(t; u, v)$$

be the corresponding value of the forward curve at time $t \leq u$. Due to no-arbitrage considerations, the following equality must hold for every $t \leq T_1$:

$$\int_{T_1}^{T_2} e^{-r(u-t)} F(t; T_1, T_2)\, \mathrm{d}u = \int_{T_1}^{T_2} e^{-r(u-t)} f(t, u)\, \mathrm{d}u,$$

where r is the constant risk free interest rate. Thus, the swap rate F can be written as the weighted integral

$$F(t; T_1, T_2) = \int_{T_1}^{T_2} \omega(u; T_1, T_2) f(t, u)\, \mathrm{d}u,$$

with the non-negative discounting factor

$$\omega(u; T_1, T_2) := \frac{e^{-ru}}{\int_{T_1}^{T_2} e^{-rv}\, \mathrm{d}v}. \tag{1}$$

Since no initial payment is needed to enter a swap contract, the swap rate $F(t; T_1, T_2)$ is a martingale under the risk neutral measure.

Fig. 1 illustrates the different prices and concepts from the energy market and their relation. One year's worth of daily spot prices are taken from actual EEX data. The seasonality function is a truncated Fourier series fitted to the spot. Each traded swap contract is represented by a single horizontal line; these are market data, too. The longest lines correspond to contracts with a delivery period of one year, whereas shorter lines represent quarterly and monthly products. Finally, the forward curve is obtained by smooth interpolation of the swap data, also taking the seasonality into account. For an overview of the fitting methods see, e.g., [4,19].

Consider now a European call option with maturity T and strike rate K, with the underlying being a swap. The value of such a swaption at time $t \leq T$ is given by

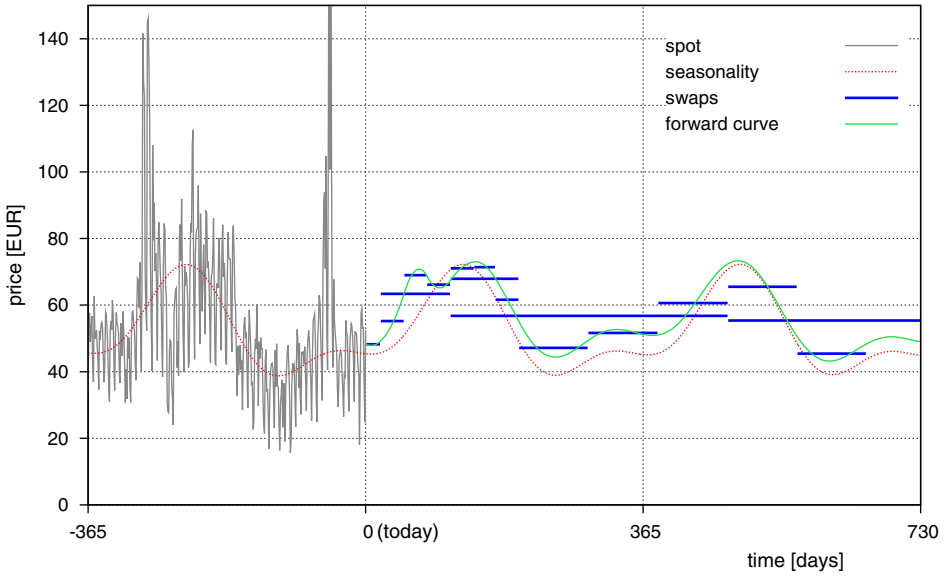

Fig. 1. EEX spot price data, fitted seasonality, traded swaps, and interpolated forward curve.

$$E\left[\left(\int_{T_1}^{T_2} e^{-r(u-t)} F(T;T_1,T_2)\,\mathrm{d}u - \int_{T_1}^{T_2} e^{-r(u-t)} K\,\mathrm{d}u\right)^+ \middle| \mathcal{F}_t\right]$$
$$= \kappa(t)\; E\left[\left(F(T;T_1,T_2) - K\right)^+ \middle| \mathcal{F}_t\right],$$

where

$$\kappa(t) := \kappa(t;T_1,T_2) := \int_{T_1}^{T_2} e^{-r(u-t)}\,\mathrm{d}u. \tag{2}$$

1.2. The objective and outline of the article

We use a Hilbert space valued, time-inhomogeneous exponential jump-diffusion process to model the forward curve. This enables us to reproduce a large variety of stylized features observed in electricity prices, e.g. the Samuelson effect of increasing volatilities close to maturity. The model is a generalization of, but not limited to, the models presented in [3,18]. In Section 2, we discuss the driving stochastic process in detail. In particular, we define the exponential of the jump-diffusion process and show that its values are elements of the Hilbert space themselves. We also examine how the drift has to be chosen in order to make the exponential a martingale.

It has been shown in [15] that, using this model, European swaptions can be priced with an efficient numerical algorithm based on partial integro-differential equations (PIDEs) and dimension reduction methods. The main goal of this article is to solve the corresponding hedging problem for European options. The challenge here is to hedge an option depending on an infinite-dimensional object (the forward curve) with a small set of traded contracts (swaps with various delivery periods). We may, e.g., want to hedge a monthly swaption with several weekly swaps and one monthly swap. It is inherent to the problem that no perfect hedge is possible, even in a

pure diffusion setting. There is a so called *basis risk*, which cannot be avoided or hedged with the given underlyings. Quadratic hedging therefore seems to be a reasonable approach. For an introduction to quadratic hedging in the Brownian case see, e.g., [12,21,28]. Hedging with more general driving processes is discussed in [7,25]. It is worth mentioning that, despite the fact that we are modeling forward curves, hedging methods for interest rate markets are not directly applicable here due to the special characteristics of electricity contracts.

In Section 3, we present our main results. We derive a representation of the (not necessarily unique) optimal hedging strategy as the solution of a linear equation system. This is in fact a generalization of the hedging formulas in one-dimensional jump-diffusion models. In order to improve readability, some of the more technical proofs needed for these results are postponed until Section 4. There, we discuss the properties of swap rates in detail. We show differentiability and calculate their stochastic dynamics. Moreover, the partial integro-differential equation (PIDE) satisfied by the swaption price is derived.

Similarly to a classical delta hedge, the optimal hedging strategy depends on partial derivatives of the option price. These derivatives can be approximated numerically by a dimension reduction approach, which is the topic of a separate, closely related paper [16]. To the best of our knowledge, the present article presents the first solution to the hedging problem for swaptions using traded swaps with various delivery periods.

2. The Hilbert space valued forward curve model

In this section we state the Hilbert space valued model which we will use throughout this article. We introduce the exponential additive process describing the forward curve and discuss moments and martingale conditions.

2.1. The Hilbert space valued exponential

Several authors propose exponential additive processes (also known as exponential time-inhomogeneous Lévy processes) of diffusion or jump-diffusion type to model the forward curve under the risk neutral measure [6,18]. Generalizing this approach, we now state the Hilbert space valued model used throughout this article. For a definition of stochastic processes and integration in Hilbert spaces with respect to Brownian motion see, e.g., [10,20]. An overview of Poisson random measures in Hilbert spaces can be found in [14]; the case of Lévy processes is treated in [24]. Infinite-dimensional stochastic analysis and its applications to interest rate theory are presented in [8].

We consider forward curves defined on the delivery period $D := [T_1, T_2]$ which are elements of a separable Hilbert space

$$H \subseteq L^2([T_1, T_2]; \mu_D),$$

with μ_D denoting the Lebesgue measure on D. For every $h \in H$ we denote the corresponding norm by

$$\|h\|_H := \sqrt{\int_{T_1}^{T_2} \big[h(u)\big]^2 \mu_D(\mathrm{d}u)}.$$

The basic driving stochastic process for our model is the H-valued additive process

$$X_t := \int_0^t \gamma_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W(s) + \int_0^t \int_E \eta_s(y)\, \widetilde{M}(\mathrm{d}y, \mathrm{d}s), \quad t \geq 0. \tag{3}$$

The diffusion part is driven by a U-valued Wiener process W, where $(U, \|\cdot\|_U)$ is a separable Hilbert space. The covariance of W is a symmetric non-negative definite trace class operator Q. The mark space $(E, \|\cdot\|_E)$ of the Poisson part of the process is a Banach space. The jumps are characterized by $\widetilde{M}$, the compensated random measure of an E-valued compound Poisson process

$$J_t = \sum_{i=1}^{N_t} Y_i, \quad t \geq 0,$$

which is independent of W. Here, N denotes a Poisson process with intensity λ and $Y_i \sim P^Y$ $(i = 1, 2, \ldots)$ are iid on E (and independent of N). The corresponding Lévy measure is denoted by $\nu = \lambda P^Y$. We denote by $L(U, H)$ and $L(E, H)$ the spaces of all bounded linear operators mapping U and E to H, respectively. We assume the drift $\gamma : [0, T] \to H$, the volatility $\sigma : [0, T] \to L(U, H)$ and the jump dampening factor $\eta : [0, T] \to L(E, H)$ to be deterministic functions. For an introduction to time dependent Bochner spaces, such as $L^2(0, T; H)$, see [13, Ch. 5.9]. The following hypothesis is assumed to hold.

Assumption 2.1. We assume that the second exponential moment of the jump distribution Y exists:

$$E[e^{2\|Y\|_E}] = \int_E e^{2\|y\|_E} P^Y(\mathrm{d}y) < \infty.$$

We assume further that $\|\eta_t\|_{L(E,H)} \leq 1$ for a.e. $t \in [0, T]$,

$$\gamma \in L^2(0, T; H), \quad \text{and} \quad \sigma \in L^2(0, T; L(U, H)).$$

By Assumption 2.1, $(X_t)_{t\geq 0}$ is an additive process with finite activity jump part and finite expectation.

In the spirit of a LIBOR model, we would now like to model the forward curve $f_t \in H$ as the exponential of the driving process X in some sense. To this end, we could take the pointwise exponential $f_t(u) = f_0(u) \exp\big(X_t(u)\big)$, $u \in D = [T_1, T_2]$. While this would be possible, several technical assumptions would then have to be made to ensure that f_t is again square integrable (and thus an element of the Hilbert space H). Since we are interested in swap rates, and not in pointwise evaluations of forward curves, a more natural way to define the exponential is the following: choose an orthonormal basis $\{e_k\}_{k\in\mathbb{N}}$ of H and set

$$f_t := \sum_{k\in\mathbb{N}} \langle f_0, e_k\rangle_H \, e^{\langle X_t, e_k\rangle_H} \, e_k \tag{4}$$

for $t \geq 0$, with $f_0 \in H$.

Note that the choice of the basis $\{e_k\}_{k\in\mathbb{N}}$ is part of the modeling process, like the choice of jump distributions and correlation structures. This allows us to solve the hedging problem for various modeling paradigms with the same, unified theoretical framework. One may use, e.g., eigenfunctions obtained from principal component analysis of the market data. In this case, (4) is nothing more than a factor model describing the dynamics of each component. Since electricity is usually traded on an hourly basis, another reasonable approach is to use piecewise constant indicator functions on hourly intervals. The Hilbert space actually considered then is a subspace of H of high but finite dimension. This is also the idea presented in [16]. Thus, (4) describes a

family of models corresponding to different ways of modeling the forward curves. For a finite-dimensional Hilbert space, definition (4) is equivalent to the pointwise exponential if $\{e_k\}$ are canonical unit vectors. This is, in particular, the setting for a multivariate stock market, where we have

$$f_i(t) = f_i(0)\, e^{X_i(t)}, \quad i = 1, \ldots, \dim H.$$

Remark 2.2. The forward curve in the general model (4) is not necessarily positive. Considering the fact that negative electricity prices have indeed been observed in the past, this may or may not be a disadvantage. Note that for the finite-dimensional space of piecewise hourly constant forward curves (choosing $\{e_k\}$ as normalized indicator functions of each hour) we *do* obtain positivity. Using different sets of basis vectors, the scores $\langle f_t, e_k\rangle_H$ remain positive, but the value of the forward curve may become negative. In a reasonably calibrated model, however, the probability of negative values should be negligible.

2.2. Properties of the model

The way that we have defined the exponential in (4) makes it easy to show the existence of moments and to derive sufficient conditions for f to be an H-valued martingale. We start with a proposition concerning the properties of the additive process X defined in (3).

Proposition 2.3. *The process $(X_t)_{t\geq 0}$ is square integrable and*

$$\sup_{0\leq t\leq T} E\|X_t\|_H^2 < \infty.$$

Let $\sigma_s^* \in L(H, U)$ be the adjoint operator of $\sigma_s = \sigma(s)$. The characteristic function of X_t is given by

$$E\left[e^{i\langle X_t, h\rangle_H}\right] = \exp\left[i\left\langle\int_0^t \gamma_s \,\mathrm{d}s, h\right\rangle_H - \frac{1}{2}\left\langle\left[\int_0^t \sigma_s Q\sigma_s^* \,\mathrm{d}s\right](h), h\right\rangle_H + \int_0^t\int_E \left(e^{i\langle\eta_s(y),h\rangle_H} - 1 - i\langle\eta_s(y), h\rangle_H\right)\nu(\mathrm{d}y)\mathrm{d}s\right]$$

for every $h \in H$.

Proof. For a proof, see [15, Ths. 2.2,2.3]. □

Due to the existence of second moments, the bounded linear covariance operator

$$\mathcal{C}_{X(T)}\colon \begin{cases} H \to H' \cong H \\ h \mapsto E\left[\langle X_T - E[X_T], h\rangle_H \langle X_T - E[X_T], \cdot\rangle_H\right] \end{cases}$$

is well-defined. By [15, Th. 2.4], it is a symmetric non-negative definite trace class operator (and thus compact).

Moreover, we can show the existence of certain Laplace transforms of X_t. This is similar to the properties of additive processes in the finite-dimensional case presented, e.g., in [27].

Proposition 2.4. *There are constants $C_1, C_2 > 0$ such that for every $h \in H$ with $\|h\|_H \leq 2$ and a.e. $t \in [0, T]$, we have*

$$E\left[e^{\langle X_t, h\rangle_H}\right] = \exp\left[\left\langle\int_0^t \gamma_s \,\mathrm{d}s, h\right\rangle_H + \frac{1}{2}\left\langle\left[\int_0^t \sigma_s Q\sigma_s^* \,\mathrm{d}s\right](h), h\right\rangle_H\right.$$

$$+\int_0^t\int_E\left(e^{\langle\eta_s(y),h\rangle_H}-1-\langle\eta_s(y),h\rangle_H\right)\nu(\mathrm{d}y)\mathrm{d}s\Bigg]$$
$$\leq C_1\,e^{C_2T}. \tag{5}$$

Proof. Using Assumption 2.1 and the Cauchy–Schwarz inequality, we obtain

$$\int_E e^{\langle\eta_t(y),h\rangle_H}\,\nu(\mathrm{d}y)\leq\int_E e^{\|y\|_E\|h\|_H}\,\nu(\mathrm{d}y)\leq\lambda\int_E e^{2\|y\|_E}\,P^Y(\mathrm{d}y)<\infty.$$

By [24, Th. 4.30], this is sufficient for the equality in (5). A theorem for interchanging linear operators and Bochner integrals [13, App. E, Th. 8] yields

$$\left\langle\int_0^t\gamma_s\,\mathrm{d}s,h\right\rangle_H=\int_0^t\langle\gamma_s,h\rangle_H\,\mathrm{d}s,$$
$$\left\langle\left[\int_0^t\sigma_sQ\sigma_s^*\mathrm{d}s\right](h),h\right\rangle_H=\int_0^t\left\langle\left[\sigma_sQ\sigma_s^*\right](h),h\right\rangle_H\,\mathrm{d}s.$$

Hence, we have the estimate

$$\begin{aligned}
&\exp\Bigg[\left\langle\int_0^t\gamma_s\,\mathrm{d}s,h\right\rangle_H+\frac{1}{2}\left\langle\left[\int_0^t\sigma_sQ\sigma_s^*\mathrm{d}s\right](h),h\right\rangle_H\\
&\quad+\int_0^t\int_E\left(e^{\langle\eta_s(y),h\rangle_H}-1-\langle\eta_s(y),h\rangle_H\right)\nu(\mathrm{d}y)\,\mathrm{d}s\Bigg]\\
&\leq\exp\Bigg[2\int_0^T\|\gamma_s\|_H\,\mathrm{d}s+\frac{4}{2}\int_0^T\|Q\|\,\|\sigma_s\|^2_{L(U,H)}\,\mathrm{d}s\\
&\quad+T\lambda\left(\int_E e^{2\|y\|_E}\,P^Y(\mathrm{d}y)+1+2\int_E\|y\|_E\,P^Y(\mathrm{d}y)\right)\Bigg]\\
&\leq\exp\Bigg[2C\|\gamma\|_{L^2(o,T;H)}+2\|Q\|\,\|\sigma\|^2_{L^2(o,T;l(U,H))}\\
&\quad+T\lambda\left(\int_E e^{2\|y\|_E}\,P^Y(\mathrm{d}y)+1+2\int_E\|y\|_E\,P^Y(\mathrm{d}y)\right)\Bigg].
\end{aligned}$$

This implies the statement of the proposition, again by Assumption 2.1. □

The next important step is to show that the forward curve $f_t=f(t)$ is indeed an element of the Hilbert space H.

Proposition 2.5. *The process $(f_t)_{0\leq t\leq T}$, which is defined as an exponential of X_t by (4), satisfies $\|f_t\|_H<\infty$ almost surely. Moreover, there are constants $C_1,C_2>0$ such that*

$$E\left[\|f_t\|_H^2\right]\leq C_1\,e^{C_2T}\,\|f_0\|_H^2\quad\text{for a.e. } t\in[0,T]. \tag{6}$$

Proof. It is enough to show (6), since this implies $\|f_t\|_H<\infty$ almost surely. To this end, we use monotone convergence and calculate

$$E\left[\langle f_tf_t\rangle_H\right]=E\left[\sum_{k\in\mathbb{N}}\langle f_0,e_k\rangle_H^2\,e^{2\langle X_t,e_k\rangle_H}\right]=\sum_{k\in\mathbb{N}}\langle f_0,e_k\rangle_H^2\,E\left[e^{\langle X_t,2e_k\rangle_H}\right].$$

Applying Proposition 2.4 with $h=2e_k$ yields (6). □

Finally, we can calculate the unique drift $\gamma \in L^2(0,T;H)$ which makes all swap rates martingales. We define Hilbert space martingales in the sense of Kunita [20], i.e., f is considered a Hilbert space valued martingale if and only if

$$\big(\langle f_t, h\rangle_H\big)_{t\geq 0}.$$

is a real-valued martingale for every $h \in H$.

Proposition 2.6. *The process $\big(f_t\big)_{0\leq t\leq T}$ is an H-martingale in the sense of Kunita, if and only if*

$$\gamma_t = \sum_{k\in\mathbb{N}}\left[-\frac{1}{2}\big\langle\big[\sigma_t Q\sigma_t^*\big](e_k), e_k\big\rangle_H - \int_E\Big(e^{\langle\eta_t(y),e_k\rangle_H} - 1 - \langle\eta_t(y), e_k\rangle_H\Big)\nu(\mathrm{d}y)\right]e_k \tag{7}$$

for a.e. $t \in [0,T]$.

Proof. By definition, f is an H-martingale if and only if

$$\big(\langle f_t, h\rangle_H\big)_{t\geq 0} = \left(\sum_{k\in\mathbb{N}}\langle f_0, e_k\rangle_H e^{\langle X_t, e_k\rangle_H}\langle h, e_k\rangle_H\right)_{t\geq 0}$$

is a martingale for every $h \in H$. By Proposition 2.4 and the Cauchy–Schwarz inequality, we obtain

$$E\left[\sum_{k\in\mathbb{N}}|\langle f_0, e_k\rangle_H|\, e^{\langle X_t,e_k\rangle_H}\,|\langle h, e_k\rangle_H|\right] \leq C_1\, e^{C_2 T}\|f_0\|_H\|h\|_H.$$

Hence, we may use dominated convergence to calculate

$$\begin{aligned}E\Big[\langle f_t, h\rangle_H\Big] &= \sum_{k\in\mathbb{N}}\langle f_0, e_k\rangle_H\langle h, e_k\rangle_H E\Big[e^{\langle X_t,e_k\rangle_H}\Big]\\ &= \sum_{k\in\mathbb{N}}\langle f_0, e_k\rangle_H\langle h, e_k\rangle_H \exp\left[\int_0^t\langle\gamma_s, e_k\rangle_H\,\mathrm{d}s + \frac{1}{2}\int_0^t\big\langle\big[\sigma_s Q\sigma_s^*\big](e_k), e_k\big\rangle_H\,\mathrm{d}s\right.\\ &\quad\left. + \int_0^t\int_E\Big(e^{\langle\eta_s(y),e_k\rangle_H - 1 - \langle\eta_s(y),e_k\rangle_H}\Big)\,\nu(\mathrm{d}y)\mathrm{d}s\right].\end{aligned}$$

Consequently, the drift γ given by (7) makes f an H-martingale, since

$$\sum_{k\in\mathbb{N}}\langle f_0, e_k\rangle_H\langle h, e_k\rangle_H = \langle f_0, h\rangle_H.$$

On the other hand, setting $h = e_k$ ($k \in \mathbb{N}$) in the calculation above shows that this is indeed the only possible choice for γ. □

Since swap rates are martingales under the risk neutral measure, we will subsequently assume that γ is defined by (7).

3. Hedging electricity swaptions

In this section, we present the main results of this article. We solve the quadratic hedging problem for European electricity swaptions. We do the hedging under the risk neutral measure. Since the interest rate is assumed to be constant, this coincides with the forward measure. For

a discussion of the advantages of hedging under the pricing measure compared to the physical measure, see, e.g., [9,22].

The basic challenge here is to hedge an infinite-dimensional object with a small, finite set of assets. The portfolio may only contain contracts which are available for trading, namely swaps with various delivery periods. Thus, it is inherent to the problem that we will not obtain a perfect hedge, even in a pure diffusion model. We first discuss the stochastic dynamics of the swaps in our portfolio and state the partial integro-differential equation (PIDE) satisfied by the swaption price. These results are then used to derive a representation of an optimal hedging strategy. Finally, we show that our solution can be interpreted as a generalization of the optimal hedge in a one-dimensional jump-diffusion model.

3.1. Stochastic dynamics of swap rates and swaption prices

We consider a portfolio of n swap contracts available for trading, whose delivery periods are given by $D_i := [T_1^i, T_2^i]$, $i = 1, \dots, n$. (We may for example want to hedge a quarterly swaption by trading in the quarterly swap itself as well as three monthly swaps.) The swap rates corresponding to the swaps in our portfolio are given by

$$F(t; T_1^i, T_2^i) = \int_{T_1^i}^{T_2^i} \omega_i(u) f(t,u)\,\mathrm{d}u,$$

where

$$\omega_i(u) := \omega(u; T_1^i, T_2^i) = \frac{e^{-ru}}{\int_{T_1^i}^{T_2^i} e^{-ru}\,\mathrm{d}u}$$

is the discounting factor defined in (1). We consider a European option with maturity T written on the swap with delivery period $D = [T_1, T_2]$. Since we cannot hedge with swaps whose delivery periods start before maturity of the option, we will assume $T \le T_1^i$ for every $i = 1, \dots, n$.

For the computation of an optimal hedging strategy, the stochastic dynamics of the swap rates $F(t; T_1^i, T_2^i)$, $i = 1, \dots, n$, play a central role. Each rate $F(t; T_1^i, T_2^i)$ is a real-valued, deterministic function of the forward curve f. More precisely,

$$\big(F(t; T_1^i, T_2^i)\big)_{0\le t\le T} = \big(\langle \omega_i, f_t\rangle_H\big)_{0\le t\le T}$$

is a real-valued martingale, since f is an H-martingale by Proposition 2.6. By (4), the forward curve is in turn a deterministic function of the driving jump-diffusion X defined in (3). We may thus introduce

$$F_i\colon \begin{cases} H \to \mathbb{R} \\ x \mapsto \left\langle \omega_i, \sum_{k\in\mathbb{N}} \langle f_0, e_k\rangle_H \exp\Big(\langle x, e_k\rangle_H\Big) e_k \right\rangle_H \end{cases} \tag{8}$$

and obtain

$$F_i(X_t) = F(t; T_1^i, T_2^i), \quad t \in [0, T].$$

We denote the Fréchet derivative of F_i at $x \in H$ by $D_x F_i(x) \in L(H, \mathbb{R})$. The stochastic dynamics of the swap rates is obtained by applying a Hilbert space valued version of Itô's

formula. We postpone the rather technical proof to Section 4, along with results concerning the differentiability of F_i.

Proposition 3.1. *The functions $F_i : H \to \mathbb{R}$, $i = 1, \ldots, n$, defined in (8), satisfy*

$$\mathrm{d}F_i(X_t) = D_x F_i(X_{t-})\,\sigma_t\,\mathrm{d}W_t + \int_E \left[F_i(X_{t-} + \eta_t(y)) - F_i(X_{t-})\right] \widetilde{M}(\mathrm{d}y, \mathrm{d}t).$$

We now consider the price process of a swaption. To this end, it is useful to introduce a centered version of X, which we denote by

$$Z_t := X_t - E[X_t] = X_t - \int_0^t \gamma_s\,\mathrm{d}s = \int_0^t \sigma_s\,\mathrm{d}W_s + \int_0^t \int_E \eta_s(y)\,\widetilde{M}(\mathrm{d}y, \mathrm{d}s).$$

Defining the swaption price in terms of Z instead of X simplifies the PIDE for the price, which we will derive below. It avoids the occurrence of an additional convection term in the PIDE. Moreover, it is also convenient for numerical approximations, which rely on projections of the centered process Z. We denote the price of the swaption at time $t \le T$ discounted to time 0 by

$$\widehat{V}(t, z) := e^{-rT} E\left[G(Z_T)\middle| Z_t = z\right], \tag{9}$$

where G is the payoff function of the option in terms of Z_T. We make the following assumption concerning the payoff.

Assumption 3.2. Suppose that the payoff function G is Lipschitz continuous on H with Lipschitz constant K_G.

Remark 3.3. Assumption 3.2 is not necessarily satisfied for payoffs depending on the exponential of Z_T, e.g. a plain call option depending on f_T. However, this can be easily remedied. In the specific case of a call, we can apply a put–call parity. More generally, every payoff can be truncated to a bounded domain (e.g. by multiplying with a smooth cutoff function). A payoff has finite expectation by definition; hence the error introduced by truncation is arbitrarily small. Since we have to localize the computational domain for any numerical calculation anyway (for details see [16]), Assumption 3.2 is no substantial restriction.

In addition, we will generalize two assumptions to the Hilbert space valued setting, which are usually made when pricing with PIDEs. The first one implies non-vanishing diffusion. Let $E_0(\mathcal{C}_{X(T)})$ be the eigenspace of the covariance operator $\mathcal{C}_{X(T)}$ corresponding to eigenvalue 0 (the kernel), a subspace which is with probability 1 never reached by X. Its orthogonal complement is $E_0(\mathcal{C}_{X(T)})^\perp$. As before, let Q be the covariance operator of W.

Assumption 3.4. Assume that for every $t \in [0, T]$, the restriction of the operator $\sigma_t Q \sigma_t^*$ to the subspace $E_0(\mathcal{C}_X)^\perp \subset H$ is positive definite, i.e.,

$$\langle \sigma_t Q \sigma_t^* h, h\rangle_H > 0 \quad \text{for every } h \in E_0(\mathcal{C}_{X(T)})^\perp \setminus \{0\}.$$

The second assumption deals with the regularity of the price process. It is common to assume that $\widehat{V}$ is twice continuously differentiable; see, e.g., [9,17,23]. For finite-dimensional spaces, this is indeed a direct consequence of Assumptions 3.2 and 3.4, as shown in [15, Th. 3.6]. We denote by $L_{HS}(H, H) \subset L(H, H)$ the space of Hilbert–Schmidt operators defined on H.

Assumption 3.5. Suppose that $\widehat{V} \in C^{1,2}((0,T) \times H, \mathbb{R}) \cap C([0,T] \times H, \mathbb{R})$, i.e. $\widehat{V}$ is continuously differentiable with respect to t and twice continuously Fréchet differentiable with respect to z. Moreover, assume that the second derivative satisfies $D_z^2 \widehat{V}(t,z) \in L_{HS}(H,H)$ for every $(t,z) \in [0,T] \times H$ and the mapping $D_z^2 \widehat{V}\colon (t,z) \to L_{HS}(H,H)$ is uniformly continuous on bounded subsets.

The stochastic dynamics for $\widehat{V}$ are very similar to those of F_i. The proof relies on Itô's formula on Hilbert spaces and is once again postponed until Section 4. We denote the trace of a nuclear operator A by $\mathrm{tr}(A)$.

Theorem 3.6. *For every $t \in [0,T]$, the discounted price $\widehat{V}$ defined in (9) satisfies*

$$\mathrm{d}\widehat{V}(t, Z_t) = D_z \widehat{V}(t, Z_{t-})\sigma_t \,\mathrm{d}W_t + \int_E \big[\widehat{V}(t, Z_{t-} + \eta_t(y)) - \widehat{V}(t, Z_{t-})\big]\, \widetilde{M}(\mathrm{d}y, \mathrm{d}t).$$

Moreover, it is a classical solution of the PIDE

$$\begin{aligned} -\partial_t \widehat{V}(t,z) = {} & \frac{1}{2}\mathrm{tr}\big[D_z^2 \widehat{V}(t,z)\sigma_t Q \sigma_t^*\big] \\ & + \int_E \Big\{\widehat{V}(t, z + \eta_t(y)) - \widehat{V}(t,z) - D_z \widehat{V}(t,z)\, \eta_t(y)\Big\} \nu(\mathrm{d}y), \end{aligned}$$

with terminal condition

$$\widehat{V}(T,z) = e^{-rT} G(z),$$

for every $t \in (0,T)$, $z \in E_0(\mathcal{C}_{X(T)})^{\perp}$.

It is possible to price the swaption with a PIDE, which relies on a Lévy model (with the same terminal distribution as the time-inhomogeneous model) and therefore uses constant coefficients. The PIDE needed for hedging, however, has time dependent coefficients. The stochastic dynamics of both the swap rate and the swaption price will be needed for the construction of an optimal hedging portfolio.

3.2. Optimal hedging strategies

In this section, we derive the optimal hedging strategy for quadratic hedging with a portfolio of swaps. Before we can compute the hedge, we need to discuss the set of admissible strategies and the corresponding value of the portfolio. A trading strategy is given by $\big(\theta_0(t), \theta(t)\big)$, $0 \le t \le T$, where $\theta_0 \in \mathbb{R}$ is the risk free investment and $\theta(t) = \big(\theta_1(t), \ldots, \theta_n(t)\big) \in \mathbb{R}^n$ describes the investment in each swap at time t. The value of the portfolio at time t is denoted by $V^{\theta}(t)$. The value S_0 of the risk free asset solves the differential equation

$$\mathrm{d}S_0(t) = r S_0(t) \mathrm{d}t.$$

Since a swap has no inherent value (you can enter the contract without paying anything), we have

$$V^{\theta}(t) = \theta_0(t) S_0(t).$$

Nevertheless, changes of the swap rates affect the wealth of the investor. In order to be self-financing, the discounted value of the portfolio must satisfy the following equation:

$$\mathrm{d}\widehat{V}^{\theta}(t) = \sum_{i=1}^{n} \theta_i(t) e^{-rt} \kappa_i(t)\, \mathrm{d}F_i(t),$$

where

$$\kappa_i(t) := \kappa(t; T_1^i, T_2^i) = \int_{T_1^i}^{T_2^i} e^{-r(u-t)}\,\mathrm{d}u.$$

The discounting factor κ has already been introduced in (2). A strategy $\big(\theta_0(t), \theta(t)\big)$ is admissible, if it is predictable, càglàd, and satisfies

$$E\left|\int_0^T \sum_{i=1}^n \theta_i(t) e^{-rt} \kappa_i(t)\,\mathrm{d}F_i(t)\right|^2 < \infty.$$

Quadratic hedging consists in minimizing the expected global quadratic hedging error

$$J(\theta) := E|\widehat{V}^\theta(T) - \widehat{V}(T)|^2. \tag{10}$$

In order to simplify and shorten notation, we define abbreviations for the jumps of swap rates and option price:

$$\begin{aligned}
&\delta F_i(t, y) := F_i\big(X_{t-} + \eta_t(y)\big) - F_i(X_{t-}), \quad i = 1, \ldots, n,\\
&\delta \widehat{V}(t, y) := \widehat{V}\big(t, Z_{t-} + \eta_t(y)\big) - \widehat{V}(t, Z_{t-}), \quad \text{for } y \in E.
\end{aligned}$$

Moreover, we will omit some of the more obvious function arguments and write e.g. $D_x F_i$ for $D_x F_i(X_{t-})$ and $D_z \widehat{V}$ for $D_z \widehat{V}(t, Z_{t-})$. The following matrix valued process M is essential for all our further computations. It describes the sensitivity of the traded swaps to changes of the driving stochastic processes.

$$\begin{aligned}
&m_{ij}(t) := e^{-2rt}\kappa_i(t)\kappa_j(t)\left(D_x F_i\, \sigma_t Q \sigma_t^*\, D_x F_j + \int_E \delta F_i \delta F_j\, \nu(\mathrm{d}y)\right),\\
&\quad i, j = 1, \ldots, n,\\
&M(t) := (m_{ij}(t))_{i,j=1}^n \in \mathbb{R}^{n\times n}.
\end{aligned} \tag{11}$$

Note that M is symmetric positive semi-definite by construction. Notice that we do not assume M to be strictly positive definite. Consequently, we allow for swaps in the portfolio, which are redundant or irrelevant to the hedging strategy. In particular, we cannot expect a unique optimal strategy under these weak assumptions. In practice, we could then introduce a second optimization criterion, e.g. minimizing the norm of θ.

The following proposition states a representation of the hedging error.

Theorem 3.7. *Let $M(t) \in \mathbb{R}^{n\times n}$ be the matrix valued process defined in (11). Define further*

$$b_i(t) := e^{-rt}\kappa_i(t)\left(D_x F_i\, \sigma_t Q \sigma_t^*\, D_z \widehat{V} + \int_E \delta F_i\, \delta \widehat{V}\, \nu(\mathrm{d}y)\right), \quad i = 1, \ldots, n,$$

and

$$c(t) := D_z \widehat{V}\, \sigma_t Q \sigma_t^*\, D_z \widehat{V} + \int_E (\delta \widehat{V})^2\, \nu(\mathrm{d}y).$$

Then the quadratic hedging error with strategy θ can be written as

$$J(\theta) = E\int_0^T \big[\theta(t)^T M(t)\theta(t) - 2b(t)^T\theta(t) + c(t)\big]\,\mathrm{d}t. \tag{12}$$

Proof. Inserting the dynamics of F_i and $\widehat{V}$, calculated in Proposition 3.1 and Theorem 3.6, into the definition (10) of J yields

$$\begin{aligned} J(\theta) = E\Bigg\{\Bigg[&\int_0^T \sum_{i=1}^n \theta_i(t) e^{-rt} \kappa_i(t)\, D_x F_i \sigma_t \,\mathrm{d}W_t \\ &+ \int_0^T \sum_{i=1}^n \theta_i(t) e^{-rt} \kappa_i(t) \int_E \delta F_i(t,y)\, \widetilde{M}(\mathrm{d}y, \mathrm{d}t) \\ &- \int_0^T D_z \widehat{V} \sigma_t \,\mathrm{d}W_t - \int_0^T \int_E \delta \widehat{V}(t,y)\, \widetilde{M}(\mathrm{d}y, \mathrm{d}t)\Bigg]^2\Bigg\}. \end{aligned}$$

By independence of $\widetilde{M}$ and W we hence have

$$\begin{aligned} J(\theta) = {} & E\left[\int_0^T \sum_{i=1}^n \theta_i(t) e^{-rt} \kappa_i(t)\, D_x F_i \sigma_t \,\mathrm{d}W_t - \int_0^T D_z \widehat{V} \sigma_t \,\mathrm{d}W_t\right]^2 \\ & + E\left[\int_0^T \sum_{i=1}^n \theta_i(t) e^{-rt} \kappa_i(t) \int_E \delta F_i(t,y)\, \widetilde{M}(\mathrm{d}y, \mathrm{d}t)\right. \\ & \left. - \int_0^T \int_E \delta \widehat{V}(t,y)\, \widetilde{M}(\mathrm{d}y, \mathrm{d}t)\right]^2 \\ =: {} & J_1 + J_2. \end{aligned} \tag{13}$$

We apply [10, Cor. 4.14] to the Brownian term J_1 and obtain

$$\begin{aligned} J_1 = E\int_0^T \mathrm{tr}\Bigg\{&\left[e^{-rt} \sum_{i=1}^n \theta_i(t) \kappa_i(t)\, D_x F_i - D_z \widehat{V}\right] \sigma_t Q \sigma_t^* \\ &\times \left[e^{-rt} \sum_{j=1}^n \theta_j(t) \kappa_j(t)\, D_x F_j - D_z \widehat{V}\right]^*\Bigg\} \,\mathrm{d}t. \end{aligned}$$

Note that the argument of the trace operator in this equation is a function mapping $\mathbb{R}$ to $\mathbb{R}$. Consequently, its "trace" is in fact the application of this function to 1. Moreover, the operators $D_x F_i$ and $D_z \widehat{V}$ are elements of $L(H, \mathbb{R})$, which we can identify with elements of H. Hence, we have

$$[D_x F]^*(1) = D_x F \quad \text{and} \quad [D_z \widehat{V}]^*(1) = D_z \widehat{V}.$$

Combining, we obtain

$$\begin{aligned} J_1 = E\int_0^T \Bigg[&e^{-2rt} \sum_{i=1}^n \sum_{j=1}^n \theta_i(t) \theta_j(t) \kappa_i(t) \kappa_j(t)\, D_x F_i\, \sigma_t Q \sigma_t^*\, D_x F_j \\ &- 2e^{-rt} \sum_{i=1}^n \theta_i(t) \kappa_i(t)\, D_x F_i\, \sigma_t Q \sigma_t^*\, D_z \widehat{V} + D_z \widehat{V}\, \sigma_t Q \sigma_t^*\, D_z \widehat{V}\Bigg] \,\mathrm{d}t. \end{aligned}$$

We use Theorem [24, Th. 23] to deal with the jump term J_2 in (13). This yields

$$J_2 = E\int_0^T \int_E \left\{ e^{-rt} \sum_{i=1}^n \theta_i(t)\kappa_i(t)\delta F_i(t,y) - \delta\widehat{V}(t,y) \right\}^2 \nu(\mathrm{d}y)\mathrm{d}t.$$

Adding the expressions for J_1 and J_2, we obtain (12) by definition of M, b and c. □

The expression (12) for the hedging error from the previous theorem involves a quadratic form with respect to θ. The following lemma states an important property of this quadratic form, which we will use to show the existence of an optimal hedging strategy.

Lemma 3.8. *For every $t \in [0,T]$, the vector $b(t)$ defined in Theorem 3.7 is an element of the range of $M(t)$.*

Proof. The vector $b(t)$ is an element of the range of $M(t)$ if and only if

$$\forall y \in \mathbb{R}^n : (y^T M(t) = 0 \Rightarrow y^T b(t) = 0)$$

holds. Let $y \in \mathbb{R}^n$ be such that $y^T M(t) = 0$. By definition of M, we have

$$\begin{aligned} 0 = y^T M(t) y = e^{-2rt} \Bigg[&\left(\sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \sigma_t Q \sigma_t^* \left(\sum_{j=1}^n y_j \kappa_j(t) D_x F_j \right) \\ &+ \int_E \left(\sum_{i=1}^n y_i \kappa_i(t) \delta F_i \right) \left(\sum_{j=1}^n y_j \kappa_j(t) \delta F_j \right) \nu(\mathrm{d}y) \Bigg]. \end{aligned}$$

Due to the positive semi-definiteness of Q, this yields

$$\left(\sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \sigma_t Q \sigma_t^* \left(\sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) = 0$$

and

$$\int_E \left(\sum_{i=1}^n y_i \kappa_i(t) \delta F_i \right)^2 \nu(\mathrm{d}y) = 0.$$

Using the Cauchy–Schwarz inequality, we obtain the estimate

$$\begin{aligned} |y^T b(t)| &= \left| e^{-rt} \sum_{i=1}^n y_i \kappa_i(t) \left(D_x F_i \, \sigma_t Q \sigma_t^* \, D_z \widehat{V} + \int_E \delta F_i \, \delta \widehat{V} \, \nu(\mathrm{d}y) \right) \right| \\ &\le e^{-rt} \left[\left(\sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \sigma_t Q \sigma_t^* \left(\sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \right]^{\frac{1}{2}} \left[D_z \widehat{V} \, \sigma_t Q \sigma_t^* \, D_z \widehat{V} \right]^{\frac{1}{2}} \\ &\quad + e^{-rt} \left[\int_E \left(\sum_{i=1}^n y_i \kappa_i(t) \delta F_i \right)^2 \nu(\mathrm{d}y) \right]^{\frac{1}{2}} \left[\int_E (\delta \widehat{V})^2 \nu(\mathrm{d}y) \right]^{\frac{1}{2}} \\ &= 0. \quad \square \end{aligned}$$

We are now able to derive the main result of this article, a representation of the optimal hedging strategy for portfolios containing an arbitrary number of swaps.

Theorem 3.9. *An investment strategy $\overline{\theta}$ minimizes the hedging error if and only if it solves*

$$M(t)\overline{\theta}(t) = b(t) \quad \text{for a.e. } t \in [0, T]. \tag{14}$$

There is at least one solution to this equation. It is unique if and only if $M(t)$ is strictly positive definite.

Proof. The minimal hedging error is achieved when the integrand

$$h\colon \begin{cases} \mathbb{R}^n & \to \mathbb{R} \\ \theta(t) & \mapsto \theta(t)^T M(t)\theta(t) - 2b(t)^T\theta(t) + c(t) \end{cases}$$

in (12) is minimized pointwise for a.e. $t \in [0, T]$. The matrix $M(t)$ is non-negative definite by construction. Consequently, h is a convex function (though not strictly convex). For convex functions, the necessary optimality condition of first order is already sufficient for a global minimum. Thus, every solution $\overline{\theta}(t)$ of (14) is an optimal hedging strategy. It is a direct consequence of Lemma 3.8 that $b(t)$ is an element of the range of $M(t)$. Hence, there is at least one such solution. The uniqueness property is then obvious. □

Comparison to one-dimensional hedging. The hedging portfolio computed in Theorem 3.9 is in fact a generalization of the optimal hedge in a one-dimensional jump-diffusion model. For the special case of a portfolio containing only a single swap ($n = 1$), with the same delivery period $D = [T_1, T_2]$ as the hedged swap itself, we obtain the following strategy.

Corollary 3.10. *The optimal investment for quadratic hedging with a single swap is given by*

$$\overline{\theta}(t) := \frac{D_x F\, \sigma_t Q \sigma_t^*\, D_z \widehat{V} + \int_E \delta F(t, y)\delta \widehat{V}(t, y)\, \nu(\mathrm{d}y)}{e^{-rt}\kappa(t)\Big[D_x F\, \sigma_t Q \sigma_t^*\, D_x F + \int_E \big(\delta F(t, y)\big)^2\, \nu(\mathrm{d}y)\Big]}. \tag{15}$$

We will now briefly show how this result relates to the hedging strategy for a stock market. To this end, we set all the Hilbert and Banach spaces in our model to $H = U = E = \mathbb{R}$. Since W is then a one-dimensional Brownian motion, we set $Q = \mathrm{Id}|_{\mathbb{R}}$, $\eta \equiv \mathrm{Id}|_{\mathbb{R}}$ and $\kappa \equiv 1$. Furthermore, in this case the stock price is modeled by

$$S_t = F(t, X_t) = S_0 \exp\left(\int_0^t \gamma(s)\mathrm{d}s + Z_t\right) \in \mathbb{R},$$

with an appropriate drift term γ. The option price can be written as a function of S_{t-}:

$$\widetilde{V}(t, S_{t-}) := e^{rt}\widehat{V}(t, Z_{t-}).$$

Hence, the following holds for the derivative of the price with respect to S:

$$D_z \widehat{V}(t, Z_{t-}) = e^{-rt} D_S \widetilde{V}(t, S_{t-})\, S_{t-}.$$

Finally, we calculate

$$D_x F(t, X_{t-}) = F(t, X_{t-}) = S_{t-}, \qquad \delta F(t, y) = (e^y - 1)S_t,$$

and

$$\delta \widehat{V}(t, y) = \widehat{V}(t, Z_{t-} + y) - \widehat{V}(t, Z_{t-}) = e^{-rt}\big[\widetilde{V}(t, S_{t-}e^y) - \widetilde{V}(t, S_{t-})\big].$$

Putting everything together, we can make a change of variable in (15) and obtain

$$\overline{\theta}(t) = \frac{\sigma_t^2 D_S \widetilde{V}(t, S_{t-}) + \frac{1}{S_{t-}} \int_{\mathbb{R}} (e^y - 1)\big[\widetilde{V}(t, S_{t-}e^y) - \widetilde{V}(t, S_{t-})\big]\, \nu(\mathrm{d}y)}{\sigma_t^2 + \int_{\mathbb{R}} (e^y - 1)^2\, \nu(\mathrm{d}y)}.$$

Note that this is exactly the formula for the optimal quadratic hedge in a stock market, calculated e.g. in [9, Rem. 10.3].

4. Deriving the swap dynamics

This section is concerned with the technical details of deriving the stochastic swap rate and swaption price dynamics. In particular, we will prove Proposition 3.1 and Theorem 3.6, using an Itô formula for Hilbert space valued processes. Before we can do so, however, we need to show certain differentiability properties for the swap rates F_i, $i = 1, \ldots, n$, defined in (8).

Swap rate derivatives. Since all subsequent results hold for every $i = 1, \ldots, n$, we choose i arbitrary but fixed, omit the index, and write F instead of F_i. Let us first recall the definition of derivatives on a Hilbert space (see, e.g., [11, Ch. VIII]). We denote the first and second Fréchet derivative of F at $x \in H$ by $D_x F(x) \in L(H, \mathbb{R})$ and $D_x^2 F(x) \in L(H, H)$ respectively. These are continuous linear operators such that

$$F(x + \xi) = F(x) + D_x F(x)\, \xi + \frac{1}{2} \langle D_x^2 F(x)\, \xi, \xi \rangle_H + o(\|\xi\|_H^2)$$

for every $\xi \in H$. It is often convenient to identify $D_x^2 F(x)$ with a bilinear form on $H \times H$, setting

$$D_x^2 F(x)\, (\xi_1, \xi_2) := \langle D_x^2 F(x)\, \xi_1, \xi_2 \rangle_H.$$

If F is Fréchet differentiable, then the Gâteaux derivatives

$$\partial_\xi F(x) := \frac{\partial}{\partial \xi} F(x) := \lim_{\varepsilon \to 0} \frac{F(x + \varepsilon\xi) - F(x)}{\varepsilon}$$

are also well-defined for every $\xi \in H$. They satisfy

$$\partial_\xi F(x) = D_x F(x)\, \xi.$$

If on the other hand F has linear and continuous Gâteaux derivatives, and the mapping $x \mapsto \partial_\cdot F(x) \in L(H, \mathbb{R})$ is continuous, then $F(x)$ is continuously Fréchet differentiable (i.e. F is of class C^1). The following theorem shows that the swap rate F is indeed twice differentiable.

Theorem 4.1. *The swap rate function F defined in (8) is of class C^2, i.e. it is twice continuously Fréchet differentiable. For every $x \in H$ and arbitrary $\xi, \xi_1, \xi_2 \in H$, the derivatives satisfy*

$$\begin{aligned} D_x F(x)\, \xi &= \sum_{k \in \mathbb{N}} \langle \omega, e_k \rangle_H \langle f_0, e_k \rangle_H e^{\langle x, e_k \rangle_H} \langle \xi, e_k \rangle_H \quad \text{and} \\ D_x^2 F(x)\, (\xi_1, \xi_2) &= \sum_{k \in \mathbb{N}} \langle \omega, e_k \rangle_H \langle f_0, e_k \rangle_H e^{\langle x, e_k \rangle_H} \langle \xi_1, e_k \rangle_H \langle \xi_2, e_k \rangle_H. \end{aligned} \tag{16}$$

Proof. We start by computing Gâteaux derivatives of F. By definition, we have

$$F(x + \varepsilon\xi) = \sum_{k \in \mathbb{N}} \langle \omega, e_k \rangle_H \langle f_0, e_k \rangle_H e^{\langle x + \varepsilon\xi, e_k \rangle_H}. \tag{17}$$

We define $c_k := \langle \omega, e_k\rangle_H \langle f_0, e_k\rangle_H$ and note that

$$\sum_{k\in\mathbb{N}} |c_k| \le \|\omega\|_H \|f_0\|_H < \infty.$$

Using the chain rule [11, Th. 8.2.1], we obtain

$$\frac{\partial}{\partial\varepsilon}\left(c_k e^{\langle x+\varepsilon\xi, e_k\rangle_H}\right) = c_k e^{\langle x+\varepsilon\xi, e_k\rangle_H}\langle \xi, e_k\rangle_H.$$

Moreover, the partial sums of these derivatives converge uniformly in ε for $|\varepsilon| < 1$, since

$$\left|\sum_{k=m}^{\infty} c_k e^{\langle x+\varepsilon\xi, e_k\rangle_H}\langle \xi, e_k\rangle_H\right| \le e^{\|x\|_H} e^{\|\xi\|_H}\|\xi\|_H \sum_{k=m}^{\infty}|c_k| \to 0 \quad \text{for } m\to\infty.$$

Thus, we may differentiate (17) term by term. This yields

$$\frac{\partial}{\partial\xi}F(x) = \frac{\partial}{\partial\varepsilon}F(x+\varepsilon\xi)\Big|_{\varepsilon=0} = \sum_{k\in\mathbb{N}} c_k e^{\langle x, e_k\rangle_H}\langle \xi, e_k\rangle_H.$$

These derivatives are obviously continuous in x, since $|e^{\langle x+\varepsilon\xi, e_k\rangle_H} - e^{\langle x, e_k\rangle_H}| \to 0$ for $\varepsilon \to 0$ uniformly in k. Since the Gâteaux derivatives of F are continuous, F is continuously Fréchet differentiable and

$$D_x F(x)\,\xi = \frac{\partial}{\partial\xi}F(x) = \sum_{k\in\mathbb{N}} c_k e^{\langle x, e_k\rangle_H}\langle \xi, e_k\rangle_H.$$

Due to the isometric isomorphism $L(H,\mathbb{R}) \cong H$, we may identify $D_x F(x)$ with an element in H and write

$$D_x F(x) = \sum_{k\in\mathbb{N}} c_k e^{\langle x, e_k\rangle_H} e_k.$$

By the very same arguments as for the first derivative, we obtain

$$\frac{\partial}{\partial\xi}D_x F(x) = \sum_{k\in\mathbb{N}} c_k e^{\langle x, e_k\rangle_H}\langle \xi, e_k\rangle_H\, e_k.$$

This implies the second equation in (16). □

In order to apply an Itô formula to F, one additional property for its derivatives is needed, which is stated in the following theorem.

Theorem 4.2. *The values of the second Fréchet derivative of the function $F: H \to \mathbb{R}$ defined in (8) are Hilbert–Schmidt operators. The mapping*

$$D_x^2 F: \begin{cases} H \to L_{LHS}(H,H) \\ x \mapsto D_x^2 F(x) \end{cases}$$

is uniformly continuous on bounded subsets.

Proof. The Hilbert–Schmidt norm of $D_x^2 F(x)$ is given by

$$\begin{aligned}\|D_x^2 F(x)\|^2_{L_{HS}(H,H)} &= \sum_{k\in\mathbb{N}} \langle D_x^2 F(x)\, e_k, D_x^2 F(x)\, e_k\rangle_H \\ &= \sum_{k\in\mathbb{N}} \langle f_0, e_k\rangle_H^2 \langle \omega, e_k\rangle_H^2 e^{2\langle x, e_k\rangle_H} \\ &\le \|f_0\|_H^2 \|\omega\|_H^2 e^{2\|x\|_H}.\end{aligned}$$

A similar calculation shows

$$\|D_x^2F(x_1) - D_x^2F(x_2)\|^2_{L_{HS}(H,H)} = \sum_{k\in\mathbb{N}} \langle f_0, e_k\rangle_H^2 \langle \omega, e_k\rangle_H^2 \left(e^{\langle x_1, e_k\rangle_H} - e^{\langle x_2, e_k\rangle_H}\right)^2$$
$$\leq \|f_0\|_H^2 \|\omega\|_H^2 e^{\max\{\|x_1\|_H, \|x_2\|_H\}} \|x_1 - x_2\|_H^2$$

for every $x_1, x_2 \in H$. This implies the uniform continuity on bounded subsets. □

The statement of the following lemma is a prerequisite for applying [24, Th. 8.23]. This will be useful for splitting the result of Itô 's formula into a martingale and a finite variation part.

Lemma 4.3. *The integrals*

$$\int_0^T \int_E E|F(X_{t-} + \eta_t(y)) - F(X_{t-})|^2 \, \nu(\mathrm{d}y)\mathrm{d}t$$

and

$$\int_0^T \int_E E|D_x F(X_{t-})\eta_t(y)|^2 \, \nu(\mathrm{d}y)\mathrm{d}t$$

are both finite.

Proof. For the first integral, we apply Young's inequality to obtain

$$|F(X_{t-} + \eta_t(y)) - F(X_{t-})|^2 \leq 2|F(X_{t-} + \eta_t(y))|^2 + 2|F(X_{t-})|^2.$$

We deal with the two terms separately. Using the definition of F we calculate

$$\int_0^T \int_E E|F(X_{t-})|^2 \, \nu(\mathrm{d}y)\mathrm{d}t = \lambda \int_0^T E|\langle \omega f_{t-}\rangle_H|^2 \, \mathrm{d}t$$
$$\leq \lambda \|\omega\|_H^2 \int_0^T E\|f_{t-}\|_H^2 \, \mathrm{d}t.$$

This expression is finite by Proposition 2.5. Similarly,

$$\int_0^T \int_E E|F(X_{t-} + \eta_t(y))|^2 \, \nu(\mathrm{d}y)\mathrm{d}t$$
$$= \int_0^T \int_E E\left|\sum_{k\in\mathbb{N}} \langle \omega, e_k\rangle_H \langle f_{t-}, e_k\rangle_H e^{\langle \eta_t(y), e_k\rangle_H}\right|^2 \nu(\mathrm{d}y)\mathrm{d}t$$
$$\leq \|\omega\|_H^2 \int_0^T E\|f_{t-}\|_H^2 \mathrm{d}t \int_E e^{2\|y\|_E} \, \nu(\mathrm{d}y).$$

This is finite by Proposition 2.5 and Assumption 2.1.

In order to show that the second integral in the statement of the lemma is finite, we plug in the derivative of F calculated in Theorem 4.1. This yields

$$\int_0^T \int_E E|D_x F(X_{t-})\eta_t(y)|^2 \, \nu(\mathrm{d}y)\mathrm{d}t$$
$$= \int_0^T \int_E E\left|\sum_{k\in\mathbb{N}} \langle \omega, e_k\rangle_H \langle f_{t-}, e_k\rangle_H \langle \eta(y), e_k\rangle_H\right|^2 \nu(\mathrm{d}y)\mathrm{d}t$$

$$\leq \|\omega\|_H^2 \int_0^T E\|f_{t-}\|_H^2 \mathrm{d}t \int_E \|y\|_E^2 \, \nu(\mathrm{d}y).$$

We proceed with Proposition 2.5 and Assumption 2.1 as above and the proof is finished. □

Applying Itô 's formula. We are now able to apply a Hilbert space valued version of Itô 's formula to derive the stochastic dynamics of F.

Lemma 4.4. *The function $F\colon H \to \mathbb{R}$ defined in (8) satisfies*

$$\begin{aligned} \mathrm{d}F(X_t) = {} & \frac{1}{2}\mathrm{tr}\left(D_x^2 F(X_{t-})\sigma_{t-} Q \sigma_{t-}^*\right)\mathrm{d}t \\ & + \int_E \left\{F(X_{t-} + \eta_t(y)) - F(X_{t-}) - D_x F(X_{t-})\,\eta_t(y)\right\}\nu(\mathrm{d}y)\,\mathrm{d}t \\ & + D_x F(X_{t-})\,\gamma_t\,\mathrm{d}t + D_x F(X_{t-})\,\sigma_t\,\mathrm{d}W_t \\ & + \int_E \left[F(X_{t-} + \eta_t(y)) - F(X_{t-})\right]\widetilde{M}(\mathrm{d}y, \mathrm{d}t). \end{aligned} \tag{18}$$

Proof. By Theorems 4.1 and 4.2, the assumptions of [24, Th. D.2] (Itô 's formula) are satisfied. The theorem yields

$$\begin{aligned} F(X_t) = {} & F(X_0) + \int_0^t D_x F(X_{s-})\,\mathrm{d}X_s + \frac{1}{2}\int_0^t D_x^2 F(X_{s-})\,\mathrm{d}[X, X]_s^c \\ & + \sum_{0\leq s\leq t}\left\{F(X_s) - F(X_{s-}) - D_x F(X_{s-})\,(X_s - X_{s-})\right\}, \end{aligned} \tag{19}$$

where $[X, X]^c$ denotes the continuous part of the predictable quadratic covariation as defined in [24]. By definition, we have

$$[X, X]_t^c = \sum_{i,j\in\mathbb{N}} e_i \otimes e_j\left([X_i, X_j]_t^c\right),$$

where $e_i \otimes e_j$ denotes the tensor product of the two basis elements and $X_i(t) := \langle X(t)e_i\rangle_H$ for $i \in \mathbb{N}$. Let $\mathcal{P}_i$ denote the projection onto the space spanned by the basis element e_i $(i \in \mathbb{N})$, and $\mathcal{P}_i^*$ be its adjoint operator:

$$\mathcal{P}_i\colon \begin{cases} H \to \mathbb{R} \\ h \mapsto \langle h e_i\rangle_H \end{cases} \quad \text{and} \quad \mathcal{P}_i^*\colon \begin{cases} \mathbb{R} \to H \\ a \mapsto a \cdot e_i. \end{cases}$$

By the properties of quadratic variations for real-valued processes and [10, Cor. 4.14], we obtain

$$\begin{aligned} [X_i, X_j]_t^c &= [X_i^c, X_j^c]_t = \left[\left\langle \int_0^\cdot \sigma_s\,\mathrm{d}W_s, e_i\right\rangle_H, \left\langle \int_0^\cdot \sigma_s\,\mathrm{d}W_s, e_j\right\rangle_H\right]_t \\ &= \int_0^t \mathcal{P}_i\,\sigma_s Q \sigma_s^*\,\mathcal{P}_j^*\,\mathrm{d}s = \int_0^t \langle \sigma_s Q \sigma_s^* e_j, e_i\rangle_H\,\mathrm{d}s. \end{aligned}$$

Thus, we have

$$\int_0^t D_x^2 F(X_{s-})\mathrm{d}[X, X]_s^c = \int_0^t \sum_{i,j\in\mathbb{N}} D_x^2 F(X_{s-})\,(e_i, e_j)\,\langle \sigma_s Q \sigma_s^* e_j, e_i\rangle_H\,\mathrm{d}s$$

$$= \int_0^t \sum_{j\in\mathbb{N}} D_x^2 F(X_{s-})\big(\sigma_s Q\sigma_s^* e_j, e_j\big)\,\mathrm{d}s$$

$$= \int_0^t \operatorname{tr}\left(D_x^2 F(X_{s-})\,\sigma_s Q\sigma_s^*\right)\mathrm{d}s. \tag{20}$$

It remains to reorganize the jump terms in (19). By Lemma 4.3 and the fact that $F(t)$ has finite expectation, the following holds:

$$\int_0^t D_x F(X_{s-}) \int_E \eta_s(y)\,\widetilde{M}(\mathrm{d}y,\mathrm{d}s)$$
$$= \sum_{0\le s\le t} D_x F(X_{s-})\,(X_s - X_{s-}) - \int_0^t \int_E D_x F(X_{s-})\,\eta_s(y)\,\nu(\mathrm{d}y)\mathrm{d}s.$$

Moreover, we have

$$\sum_{0\le s\le t}\Big[F(X_s) - F(X_{s-})\Big] = \int_0^t \int_E \Big[F(X_{s-}+\eta_s(y)) - F(X_{s-})\Big]\,\widetilde{M}(\mathrm{d}y,\mathrm{d}s)$$
$$+ \int_0^t \int_E \Big[F(X_{s-}+\eta_s(y)) - F(X_{s-})\Big]\,\nu(\mathrm{d}y)\,\mathrm{d}s.$$

Combined with (19) and (20), this implies (18). □

As a direct consequence of this lemma, we can state the proof of Proposition 3.1.

Proof of Proposition 3.1. By construction, $F(X)$ is a real-valued martingale. This is also true for the last two integrals in (18) by definition of the stochastic integral [24, Ths. 8.7,8.23]. Since continuous martingales of finite variation are a.s. constant [26, Th. 27], the sum of the remaining integral terms in (18) must equal 0. □

Swaption price dynamics. The remainder of the section is concerned with the proof of Theorem 3.6. We need a technical lemma similar to Lemma 4.3 in order to be able to rearrange the jump terms in the dynamics of $\widehat{V}$.

Lemma 4.5. *The integrals*

$$\int_0^T \int_E E|\widehat{V}(t, Z_{t-} + \eta_t(y)) - \widehat{V}(t, Z_{t-})|^2\,\nu(\mathrm{d}y)\,\mathrm{d}t$$

and

$$\int_0^T \int_E E|D_z\widehat{V}(t, Z_{t-})\,\eta_t(y)|^2\,\nu(\mathrm{d}y)\mathrm{d}t$$

are both finite.

Proof. By the definition of $\widehat{V}$ and the Lipschitz continuity of the payoff (Assumption 3.2) we obtain

$$|\widehat{V}(t, Z_{t-} + z) - \widehat{V}(t, Z_{t-})| \le e^{-rT} E\left[|G(Z_T + z) - G(Z_T)|\big|\mathcal{F}_t\right]$$
$$\le e^{-rT} K_G \|z\|_H \tag{21}$$

for every $z \in H$. Hence, the first integral satisfies

$$\int_0^T \int_E E|\widehat{V}(t, Z_{t-} + \eta_t(y)) - \widehat{V}(t, Z_{t-})|^2 \, \nu(\mathrm{d}y)\, \mathrm{d}t \le e^{-2rT} K_G^2 T \int_E \|y\|_E^2 \nu(\mathrm{d}y) < \infty.$$

For the second integral, note that by (21) we have

$$\|D_z \widehat{V}(t, Z_{t-})\|_{L(H,\mathbb{R})} \le e^{-rt} K_G.$$

This implies

$$\int_0^T \int_E E|D_z \widehat{V}(t, Z_{t-})\, \eta_t(y)|^2 \, \nu(\mathrm{d}y)\mathrm{d}t \le e^{-2rT} K_G^2 \int_0^T \int_E \|\eta_t(y)\|_H^2 \, \nu(\mathrm{d}y)\mathrm{d}t < \infty. \quad \square$$

The assumptions made for $\widehat{V}$ are almost identical to the results shown in Theorems 4.1 and 4.2 for the swap rate F. Hence, it is not surprising that we can derive very similar stochastic dynamics for $\widehat{V}$, using once again Itô's formula on Hilbert spaces. As before, we denote by $E_0(\mathcal{C}_{X(T)})$ the kernel of the covariance operator $\mathcal{C}_{X(T)}$ and by $E_0(\mathcal{C}_{X(T)})^\perp$ its orthogonal complement. Note that $\mathcal{C}_{X(T)}$ is also the covariance operator of the centered process Z.

Proof of Theorem 3.6. By Assumption 3.5, we may apply Itô 's formula [24, Th. D.2] to obtain

$$\begin{aligned}\widehat{V}(t, Z_t) = \widehat{V}(0, Z_0) &+ \int_0^t \partial_t \widehat{V}(s, Z_{s-})\, \mathrm{d}s + \int_0^t D_z \widehat{V}(s, Z_{s-})\, \mathrm{d}Z_s \\ &+ \frac{1}{2} \int_0^t D_z^2 \widehat{V}(s, Z_{s-})\, \mathrm{d}[Z, Z]_s^c \\ &+ \sum_{0 \le s \le t} \Big\{ \widehat{V}(s, Z_s) - \widehat{V}(s, Z_{s-}) - D_z \widehat{V}(s, Z_{s-})\, (Z_s - Z_{s-}) \Big\}.\end{aligned}$$

Since Z and X differ only with respect to a deterministic drift of finite variation, we have $[Z, Z]_s^c = [X, X]_s^c$. Consequently, the computations from the proof of Lemma 4.4 yield

$$\begin{aligned}\mathrm{d}\widehat{V}(t, Z_t) = \partial_t \widehat{V}(t, Z_{t-})\, \mathrm{d}t &+ \frac{1}{2} \mathrm{tr}\Big(D_z^2 \widehat{V}(t, Z_{t-})\, \sigma_{t-} Q \sigma_{t-}^* \Big)\, \mathrm{d}t \\ &+ \int_E \Big\{ \widehat{V}(t, Z_{t-} + \eta_t(y)) - \widehat{V}(t, Z_{t-}) - D_z \widehat{V}(t, Z_{t-})\, \eta_t(y) \Big\}\, \nu(\mathrm{d}y)\, \mathrm{d}t \\ &+ D_z \widehat{V}(t, Z_{t-})\, \sigma_t\, \mathrm{d}W_t + \int_E \big[\widehat{V}(t, Z_{t-} + \eta_t(y)) - \widehat{V}(t, Z_{t-}) \big]\, \widetilde{M}(\mathrm{d}y, \mathrm{d}t).\end{aligned}$$

Using the same arguments as in the proof of Proposition 3.1, this yields the stochastic dynamics (and the PIDE) for $\widehat{V}$ along almost every trajectory of Z. It remains to show that the PIDE does indeed hold for every $(t, z) \in (0, T) \times E_0(\mathcal{C}_{X(T)})^\perp$.

Fix an arbitrary $t \in (0, T)$ and $z \in E_0(\mathcal{C}_{X(T)})^\perp$. We denote by $B_\varepsilon(z)$ the ball with radius ε around z. Because of the non-vanishing diffusion (Assumption 3.4), the probability for $Z_{t-} \in B_\varepsilon(z)$ is non-zero for every $\varepsilon > 0$. Thus, for every $\varepsilon > 0$, we can find a $z_\varepsilon \in B_\varepsilon(z)$ such that the PIDE holds in (t, z_ε). Due to the regularity of $\widehat{V}$ (Assumption 3.5), we can conclude that the PIDE holds for every $(t, z) \in (0, T) \times E_0(\mathcal{C}_{X(T)})^\perp$. $\square$

5. Conclusion

In this article, quadratic hedging strategies for European electricity swaptions are discussed. The basic problem when hedging electricity is to hedge an infinite-dimensional object with a

finite set of traded assets (swaps with various delivery periods). We directly model the forward curve with an exponential time-inhomogeneous jump-diffusion process. We examine the moment and martingale properties of this model in detail. Stochastic dynamics and the corresponding PIDE for the swaption price are derived. We show that the optimal hedging strategy at each point in time is the solution of a linear equation system. Similarly to a classical delta hedge, the strategy depends on partial derivatives of the option price, which can be obtained from the PIDE.

The representation of the optimal hedging strategies given in this article is the starting point for efficient numerical approximation methods. The related paper [16] is concerned with the numerical computation of hedging strategies, using a dimension reduction technique.

Acknowledgments

I would like to thank Fred Espen Benth for initiating the project on hedging electricity and his greatly appreciated advice. Many thanks go also to Claudia Klüppelberg for helpful discussions and a critical reading of the manuscript.

The author was supported by the International Graduate School of Science and Engineering (IGSSE) of Technische Universität München.

References

[1] N. Audet, P. Heiskanen, J. Keppo, I. Vehviläinen, Modeling electricity forward curve dynamics in the Nordic market, in: D.W. Bunn (Ed.), Modelling Prices in Competitive Electricity Markets, in: Wiley Series in Financial Economics, Chichester, 2004 (Chapter 12).
[2] F.E. Benth, Á. Cartea, R. Kiesel, Pricing forward contracts in power markets by the certainty equivalence principle: explaining the sign of the market risk premium, Journal of Banking and Finance 32 (10) (2008) 2006–2021.
[3] F.E. Benth, S. Koekebakker, Stochastic modeling of financial electricity contracts, Energy Economics 30 (3) (2008) 1116–1157.
[4] F.E. Benth, S. Koekebakker, F. Ollmar, Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation, Journal of Derivatives 52 (15) (2007) 62–66.
[5] F.E. Benth, T. Meyer-Brandis, The information premium for non-storable commodities, Journal of Energy Markets 2 (3) (2009) 111–140.
[6] F.E. Benth, J. Šaltytė Benth, S. Koekebakker, Stochastic Modeling of Electricity and Related Markets, World Scientific, Singapore, 2008.
[7] O. Bobrovnytska, M. Schweizer, Mean–variance hedging and stochastic control: beyond the Brownian setting, IEEE Transactions on Automatic Control 49 (3) (2004).
[8] R. Carmona, M. Tehranchi, Interest Rate Models: An Infinite-Dimensional Stochastic Analysis Perspective, Springer, Berlin, 2006.
[9] R. Cont, P. Tankov, Financial Modelling with Jump Processes, Chapman & Hall, Boca Raton, 2004.
[10] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, Cambridge, 1992.
[11] J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York, 1960.
[12] D. Duffie, H.R. Richardson, Mean–variance hedging in continuous time, The Annals of Applied Probability 1 (1) (1991) 1–15.
[13] L.C. Evans, Partial Differential Equations, in: Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, 1998.
[14] E. Hausenblas, A note on the Itô formula of stochastic integrals in Banach spaces, Random Operators and Stochastic Equations 14 (1) (2006) 45–58.
[15] P. Hepperger, Option pricing in Hilbert space valued jump-diffusion models using partial integro-differential equations, SIAM Journal on Financial Mathematics 1 (2010) 454–489.
[16] P. Hepperger, Numerical hedging of electricity swaptions using dimension reduction (submitted for publication) preprint on http://www-m4.ma.tum.de, 2011.
[17] N. Hilber, N. Reich, C. Schwab, C. Winter, Numerical methods for Lévy processes, Finance and Stochastics 13 (2009) 471–500.

[18] R. Kiesel, G. Schindlmayr, R. Börger, A two-factor model for the electricity forward market, Quantitative Finance 9 (3) (2009) 279–287.
[19] T. Kluge, Pricing swing options and other electricity derivatives. Ph.D. Thesis, University of Oxford, 2006.
[20] H. Kunita, Stochastic integrals based on martingales taking values in Hilbert space, Nagoya Mathematical Journal 38 (1970) 41–52.
[21] J.P. Laurent, H. Pham, Dynamic programming and mean–variance hedging, Finance and Stochastics 3 (1) (1999) 83–110.
[22] D.B. Madan, Joint risk neutral laws and hedging, Robert H. Smith School Research Paper, No. RHS 06–140, 2010. http://ssrn.com/abstract=1603062.
[23] A.-M. Matache, T. von Petersdorff, C. Schwab, Fast deterministic pricing of options on Lévy-driven assets, ESAIM: Mathematical Modelling and Numerical Analysis 38 (1) (2004) 37–71.
[24] S. Peszat, J. Zabczyk, Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach, Cambridge University Press, Cambridge, 2007.
[25] H. Pham, On quadratic hedging in continuous time, Mathematical Methods of Operations Research 51 (2) (2000) 315–339.
[26] P.E. Protter, Stochastic Integration and Differential Equations, second ed., Springer, Berlin, 2005.
[27] K. Sato, Lévy Processes and Infinitely Divisible Distributions, Cambridge University Press, Cambridge, 1999.
[28] M. Schweizer, Mean–variance hedging for general claims, The Annals of Applied Probability 2 (1) (1992) 171–179.