Hedging Electricity Swaptions Using Partial Integro-Differential Equations

Peter Hepperger *

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 is available for hedging, the market is inherently incomplete. We derive the optimization problem for the quadratic hedging problem and state a representation of its solution, which is the starting point for numerical algorithms.

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 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 the 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 paragraph describes the contracts and mathematical objects occurring in this context, which will be the basis for our model.

*Zentrum Mathematik, Technische Universität München, 85748 Garching bei München, Germany (hepperger@ma.tum.de). Supported by the International Graduate School of Science and Engineering (IGSSE) of Technische Universität München.
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 $t \leq T_1$.

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

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

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

with the non-negative discounting factor

$$\omega(u; T_1, T_2) := \frac{e^{-ru}}{\int_{T_1}^{T_2} e^{-rv} \, dv}.$$

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.

Figure 1 illustrates the different prices and concepts from the energy market and their relation. One year 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, shorter lines represent quarterly and monthly products. Finally, the forward curve is obtained by smooth interpolation of the swap data, taking also the seasonality into account. For an overview of the fitting methods confer, 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

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

where

\[
(2) \quad \kappa(t) := \kappa(t; T_1, T_2) := \int_{T_1}^{T_2} e^{-r(u-t)} \, du.
\]

1.2 Objective and Outline of the Article

We use a Hilbert space valued, time-inhomogeneous exponential jump-diffusion process to model the forward curve. It 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 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, 27]. Hedging with more general driving processes is discussed in [7, 24]. 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.

Similar 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 matter 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 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 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 [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 [23]. Infinite-dimensional stochastic analysis and its applications to interest-rate theory and Heath–Jarrow–Morton models are presented in [8].

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

\[
H := L^2([T_1, T_2]; \mu_D),
\]

with \( \mu_D \) denoting the Lebesgue measure on \( D \). For every \( h \in H \) we denote the corresponding norm by

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

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

\[
X_t := \int_0^t \gamma_s \, ds + \int_0^t \sigma_s \, dW(s) + \int_0^t \int_E \eta_s(y) \, \tilde{M}(dy, ds), \quad t \geq 0.
\]

The diffusion part is driven by an \( U \)-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 \( \tilde{M} \), 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 \( \lambda \) 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:

\[
E[e^{2\|Y\|_E}] = \int_E e^{2\|y\|_E} P_Y(dy) < \infty.
\]

We assume further \( \| \eta \|_{L(E,H)} \leq 1 \) 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.

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 point-wise exponential
\[
f_t(u) = f_0(u) \exp(X_t(u)), \quad u \in D = [T_1, T_2].
\]
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
\[
(4) \quad f_t := \sum_{k \in \mathbb{N}} \langle f_0, e_k \rangle_H e^{(X_t,e_k)_H} e_k
\]
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 modelling process, similar to 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 actually considered Hilbert space then is a high-but finite-dimensional subspace of \(H\). This is also the idea presented in [16]. Thus, \((4)\) describes a family of models corresponding to different ways of modelling the forward curves. For a finite-dimensional Hilbert space, definition \((4)\) is equivalent to the point-wise exponential if \(\{e_k\}\) are canonical unit vectors. This is, in particular, the setting for a multivariate stock market, where we have
\[
f_i(t) = f_i(0) e^{X_i(t)}, \quad i = 1, \ldots, \dim H.
\]

2.2 Properties of the Model

The way 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 2.2.** The process \((X_t)_{t \geq 0}\) is square-integrable and
\[
\sup_{0 \leq t \leq T} E[\|X_t\|_H^2] < \infty.
\]
Let \(\sigma^*_s \in L(H,\mathcal{U})\) be the adjoint operator of \(\sigma_s = \sigma(s)\). The characteristic function of \(X_t\) is
given by

\[
E \left[ e^{i \langle X_t, h \rangle_H} \right] = \exp \left[ i \int_0^t \gamma_s ds, h \right]_H - \frac{1}{2} \left[ \int_0^t \sigma_s Q \sigma_s^* ds \right] (h, h)_H + \int_0^t \int_E \left( e^{i \langle \eta_s(y), h \rangle_H} - 1 - i \langle \eta_s(y), h \rangle_H \right) v(dy) ds
\]

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

\[
C_{X(T)} : \begin{cases} 
H \to H' \cong H \\
h \mapsto E \left[ \langle X_T - E[X_T], h \rangle_H \langle X_T - E[X_T], \cdot \rangle_H \right]
\end{cases}
\]

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 [26].

**Proposition 2.3.** 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

\[
E \left[ e^{i \langle X_t, h \rangle_H} \right] = \exp \left[ \int_0^t \gamma_s ds, h \right]_H + \frac{1}{2} \left[ \int_0^t \sigma_s Q \sigma_s^* ds \right] (h, h)_H + \int_0^t \int_E \left( e^{i \langle \eta_s(y), h \rangle_H} - 1 - i \langle \eta_s(y), h \rangle_H \right) v(dy) ds \leq C_1 e^{C_2 T}.
\]

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

\[
\int_E e^{i \langle \eta_s(y), h \rangle_H} v(dy) \leq \int_E e^{\|y\|_E \|h\|_H} v(dy) \leq \lambda \int_E e^{2 \|y\|_E} P^Y(dy) < \infty.
\]

By [23, 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

\[
\left\langle \int_0^t \gamma_s ds, h \right\rangle_H = \int_0^t \langle \gamma_s, h \rangle_H ds,
\]

\[
\left\langle \left[ \int_0^t \sigma_s Q \sigma_s^* ds \right] (h), h \right\rangle_H = \int_0^t \langle \sigma_s Q \sigma_s^* \rangle (h, h) ds.
\]
Hence, we have the estimate
\[
\exp \left[ \left\langle \int_0^t \gamma_s \, ds, h \right\rangle_H \right] + \frac{1}{2} \left[ \int_0^t \sigma_t Q \sigma_t^* \, ds \right] \left\langle h, h \right\rangle_H \\
+ \int_0^t \int_E \left( e^{(\eta_t(y), h)_H} - 1 - \langle \eta_t(y), h \rangle_H \right) v(dy) \, ds
\]
\[
\leq \exp \left[ 2 \int_0^T \| \gamma_s \|_H \, ds + \frac{4}{2} \int_0^T \| Q \| \| \sigma_s \|_{L^2(U;H)} \, ds \right. \\
+ T \lambda \left( \int_E e^{2\|y\|=P} (dy) + 1 + 2 \int_E \| y \|=P \cdot P(dy) \right)
\]
\[
\leq \exp \left[ 2C \| \gamma \|_{L^2(0,T;H)} + 2 \| Q \| \| \sigma \|_{L^2(0,T;L^2(U,H))}^2 \\
+ T \lambda \left( \int_E e^{2\|y\|=P} (dy) + 1 + 2 \int_E \| y \|=P \cdot P(dy) \right) \right].
\]

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 2.4.** 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
\[
E[\| f_t \|^2_H] \leq C_1 e^{C_2 t} \| f_0 \|^2_H \quad \text{for a.e. } t \in [0, T].
\]

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

Applying Proposition 2.3 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
\[
\left( \left\langle f_t, h \right\rangle_H \right)_{t \geq 0}
\]
is a real-valued martingale for every \( h \in H \).

**Proposition 2.5.** The process \( (f_t)_{0 \leq t \leq T} \) is an \( H \)-martingale in the sense of Kunita, if and only if
\[
\gamma_t = \sum_{k \in \mathbb{N}} \left[ -\frac{1}{2} \left\langle \sigma_t Q \sigma_t^* \right\rangle (e_k), e_k \right]_H - \int_E \left( e^{(\eta_t(y), e_k)_H} - 1 - \langle \eta_t(y), e_k \rangle_H \right) v(dy) \, e_k
\]
for a.e. \( t \in [0, T] \).
Proof. By definition, $f$ is an $H$-martingale if and only if

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

is a martingale for every $h \in H$. By Proposition 2.3 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

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

$$= \sum_{k \in \mathbb{N}} \langle f_0, e_k \rangle_H \langle h, e_k \rangle_H \exp \left[ \int_0^t \langle \gamma_s, e_k \rangle_H ds + \frac{1}{2} \int_0^t \left( \langle \sigma_s Q \sigma_s^* \rangle (e_k), e_k \right)_H ds \right.$$ 

$$+ \int_0^t \int_E \left( e^{\langle \eta_s(y), e_k \rangle_H} - 1 - \langle \eta_s(y), e_k \rangle_H \right) v(dy) ds \right].$$

Consequently, the drift $\gamma$ 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 $\gamma$. 

Since swap rates are martingales under the risk neutral measure, we will subsequently assume that $\gamma$ 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. 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 $D_i := [T_{i1}, T_{i2}]$, $i = 1, \ldots, n$. (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

$$F(t; T_{i1}, T_{i2}) = \int_{T_{i1}}^{T_{i2}} \omega_i(u) f(t, u) \, du,$$

where

$$\omega_i(u) := \omega(u; T_{i1}, T_{i2}) = \frac{e^{-ru}}{\int_{T_{i1}}^{T_{i2}} e^{-ru} \, du}$$

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 $T \leq T_{i1}$ for every $i = 1, \ldots, n$.

For the computation of an optimal hedging strategy, the stochastic dynamics of the swap rates $F(t; T_{i1}, T_{i2})$, $i = 1, \ldots, n$, play a central role. Each rate $F(t; T_{i1}, T_{i2})$ is a real-valued, deterministic function of the forward curve $f$. More precisely,

$$\left( F(t; T_{i1}, T_{i2}) \right)_{0 \leq t \leq T} = \left( \langle \omega_i, f_i \rangle_H \right)_{0 \leq t \leq T}$$

is a real-valued martingale, since $f$ is an $H$-martingale by Proposition 2.5. By (4), the forward curve is in turn a deterministic function of the driving jump-diffusion $X$ defined in (3). We may thus introduce

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

and obtain

$$F_i(X_t) = F(t; T_{i1}, T_{i2}), \quad t \in [0, T].$$

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 until 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

$$dF_i(X_t) = D_x F_i(X_t-) \sigma_t \, dW_t + \int_E \left[ F_i(X_{t-} + \eta_t(y)) - F_i(X_{t-}) \right] \tilde{M}(dy, dt).$$
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

$$Z_t := X_t - E[X_t] = X_t - \int_0^t \gamma(s) \, ds = \int_0^t \sigma(t) \, dW_t + \int_0^t \int_E \eta(s)(y) \, \tilde{M}(dy, ds).$$

We denote the price of the swaption at time $t \leq T$ discounted to time 0 by

$$\hat{V}(t, z) := e^{-rT} E\left[G(Z_T) \mid Z_t = z\right],$$

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 confer \cite{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 $E_0(C_{X(T)})$ be the eigenspace of the covariance operator $C_{X(T)}$ corresponding to eigenvalue 0 (the kernel), a subspace which is with probability 1 never reached by $X$. Its orthogonal complement is $E_0(C_{X(T)})^\perp$. 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 $E_0(C_{X(T)})^\perp \subset H$ is positive definite, i.e.,

$$\langle \sigma(t)Q\sigma^*_t h, h \rangle_H > 0 \quad \text{for every } h \in E_0(C_{X(T)})^\perp \setminus \{0\}.$$

The second assumption deals with the regularity of the price process. It is common to assume that $\hat{V}$ is twice continuously differentiable, see, e.g., \cite{9, 17, 22}. For finite-dimensional spaces, this is indeed a direct consequence of Assumptions 3.2 and 3.4, as shown in \cite[Th. 3.6]{15}. 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 $\hat{V} \in C^{1,2}(\{0, T\} \times H, \mathbb{R}) \cap C(\{0, T\} \times H, \mathbb{R})$, i.e., $\hat{V}$ is continuously differentiable with respect to $t$ and twice continuously Fréchet differentiable with respect to $z$. Moreover, assume that the second derivative satisfies $D^2_t\hat{V}(t, z) \in L_{HS}(H, H)$ for every $(t, z) \in [0, T] \times H$ and the mapping $D^2_t\hat{V} : (t, z) \to L_{HS}(H, H)$ is uniformly continuous on bounded subsets.
The stochastic dynamics for $\hat{V}$ 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 $\text{tr}(A)$.

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

$$d\hat{V}(t, Z_t) = D_z\hat{V}(t, Z_{t-}) \sigma_t dW_t + \int_E \left[ \hat{V}(t, Z_{t-} + \eta_i(y)) - \hat{V}(t, Z_{t-}) \right] \tilde{M}(dy, dt).$$

Moreover, it is a classical solution of the PIDE

$$-\partial_t \hat{V}(t, z) = \frac{1}{2} \text{tr} \left[ D_z^2 \hat{V}(t, z) \sigma_t Q \sigma_t^* \right] + \int_E \left\{ \hat{V}(t, z + \eta_i(y)) - \hat{V}(t, z) - D_z \hat{V}(t, z) \eta_i(y) \right\} \nu(dy),$$

with terminal condition

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

for every $t \in (0, T), z \in E_0(C_X(T))$. 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 $(\theta_0(t), \theta(t))$, $0 \leq t \leq T$, where $\theta_0 \in \mathbb{R}$ is the risk free investment and $\theta(t) = (\theta_1(t), \ldots, \theta_n(t)) \in \mathbb{R}^n$ 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

$$dS_0(t) = rS_0(t)dt.$$

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:

$$d\tilde{V}^\theta(t) = \sum_{i=1}^n \theta_i(t) e^{-rt_i} \kappa_i(t) dF_i(t),$$
where
\[ \kappa_i(t) := \kappa(t; T^i_1, T^i_2) = \int_{T^i_1}^{T^i_2} e^{-r(u-T)} \, du. \]

The discounting factor \( \kappa(t) \) has already been introduced in (2). A strategy \((\theta_0(t), \theta(t))\) is admissible, if it is predictable, càglàd, and satisfies
\[ E \left| \int_0^T \sum_{i=1}^n \theta_i(t) e^{-rt} \kappa_i(t) \, dF_i(t) \right|^2 < \infty. \]

Quadratic hedging consists in minimizing the expected global quadratic hedging error
\[ J(\theta) := E \left| \hat{V}^\theta(T) - \hat{V}(T) \right|^2. \]

In order to simplify and shorten notation, we define abbreviations for the jumps of swap rates and option price:
\[ \delta F_i(t,y) := F_i(X_{t-} + \eta_i(y)) - F_i(X_{t-}), \quad i = 1, \ldots, n, \]
\[ \delta \hat{V}(t,y) := \hat{V}(t,Z_{t-} + \eta_i(y)) - \hat{V}(t,Z_{t-}), \quad \text{for } y \in E. \]

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 \( D_x \hat{V} \) for \( D_x \hat{V}(t,Z_{t-}) \). 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.
\[ m_{ij}(t) := e^{-2rt} \kappa_i(t) \kappa_j(t) \left( D_x F_i \sigma_i \sigma^*_j D_x F_j + \int_E \delta F_i \delta F_j \nu(dy) \right), \quad i, j = 1, \ldots, n, \]
\[ M(t) := (m_{ij}(t))_{i,j=1}^n \in \mathbb{R}^{n \times n}. \]

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 \( \theta \).

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
\[ b_i(t) := e^{-rt} \kappa_i(t) \left( D_x F_i \sigma_i Q \sigma^*_i D_x F_i + \int_E \delta F_i \delta \hat{V} \nu(dy) \right), \quad i = 1, \ldots, n, \]
and
\[ c(t) := D_x \hat{V} \sigma_i Q \sigma^*_i D_x \hat{V} + \int_E (\delta \hat{V})^2 \nu(dy). \]

Then the quadratic hedging error with strategy \( \theta \) can be written as
\[ J(\theta) = E \left[ \int_0^T [\theta(t)^T M(t) \theta(t) - 2b(t)^T \theta(t) + c(t)] \, dt \right]. \]
Proof. Inserting the dynamics of $F_t$ and $\hat{V}$, calculated in Proposition 3.1 and Theorem 3.6, into the definition (10) of $J$ yields

$$J(\theta) = E \left\{ \left[ \int_0^T \sum_{i=1}^n \theta_i(t)e^{-rt} \kappa_i(t) D_x F_i \sigma_i dW_t + \int_0^T \sum_{i=1}^n \theta_i(t)e^{-rt} \kappa_i(t) \int_E \delta F_i(t,y) \tilde{M}(dy,dt) - \int_0^T D_2 \tilde{V} \sigma_i dW_t - \int_0^T \int_E \delta \hat{V}(t,y) \tilde{M}(dy,dt) \right]^2 \right\}.$$ 

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

$$J(\theta) = E \left[ \int_0^T \sum_{i=1}^n \theta_i(t)e^{-rt} \kappa_i(t) D_x F_i \sigma_i dW_t - \int_0^T D_2 \tilde{V} \sigma_i dW_t \right]^2$$

(13)

$$=: J_1 + J_2.$$

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

$$J_1 = E \int_0^T \text{tr} \left\{ \left[ e^{-rt} \sum_{i=1}^n \theta_i(t) \kappa_i(t) D_x F_i - D_2 \tilde{V} \right] \sigma_i \sigma_i^* \times \left[ e^{-rt} \sum_{j=1}^n \theta_j(t) \kappa_j(t) D_x F_j - D_2 \tilde{V} \right]^* \right\} dt.$$

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 $D_2 \tilde{V}$ are elements of $L(H, \mathbb{R})$, which we can identify with elements of $H$. Hence, we have

$$[D_x F]^*(1) = D_x F \quad \text{and} \quad [D_2 \tilde{V}]^*(1) = D_2 \tilde{V}.$$

Combined, we obtain

$$J_1 = E \int_0^T \left[ e^{-2rt} \sum_{i=1}^n \sum_{j=1}^n \theta_i(t) \theta_j(t) \kappa_i(t) \kappa_j(t) D_x F_i \sigma_i \sigma_i^* D_x F_j - 2e^{-rt} \sum_{i=1}^n \theta_i(t) \kappa_i(t) D_x F_i \sigma_i \sigma_i^* D_2 \tilde{V} + D_2 \tilde{V} \sigma_i \sigma_i^* D_2 \tilde{V} \right] dt.$$

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

$$J_2 = E \int_0^T \int_E \left\{ e^{-rt} \sum_{i=1}^n \theta_i(t) \kappa_i(t) \delta F_i(t,y) - \delta \hat{V}(t,y) \right\}^2 v(dy) dt.$$
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 $\theta$. The following lemma states an important property of this quadratic form, which we will use to show existence of an optimal hedging strategy.

Lemma 3.8. The vector $b(t)$ defined in Theorem 3.7 satisfies
\[ \forall y \in \mathbb{R}^n : (y^T M(t) = 0 \Rightarrow y^T b(t) = 0) \]
for every $t \in [0, T]$.

Proof. Let $y^T M(t) = 0$. By definition of $M$, we have
\[
0 = y^T M(t)y = e^{-2rt} \left[ \left( \sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \sigma_t Q \sigma_t^* \left( \sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \right. \\
+ \left. \int_{E} \left( \sum_{i=1}^n y_i \kappa_i(t) \delta F_i \right) \left( \sum_{j=1}^n y_j \kappa_j(t) \delta F_j \right) v(dy) \right]
\]
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
\[
\int_{E} \left( \sum_{i=1}^n y_i \kappa_i(t) \delta F_i \right)^2 v(dy) = 0.
\]
Using the Cauchy–Schwarz inequality, we obtain the estimate
\[
\left| y^T b(t) \right| = \left| e^{-rt} \left[ \left( \sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \sigma_t Q \sigma_t^* D_x \hat{V} + \int_{E} \delta F_i \delta \hat{V} v(dy) \right] \right|
\leq e^{-rt} \left[ \left( \sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \sigma_t Q \sigma_t^* \left( \sum_{i=1}^n y_i \kappa_i(t) D_x F_i \right) \right]^{\frac{1}{2}} \left[ D_x \hat{V} \sigma_t Q \sigma_t^* D_x \hat{V} \right]^{\frac{1}{2}}
\leq e^{-rt} \left[ \int_{E} \left( \sum_{i=1}^n y_i \kappa_i(t) \delta F_i \right)^2 v(dy) \right]^{\frac{1}{2}} \left[ \int_{E} (\delta \hat{V})^2 v(dy) \right]^{\frac{1}{2}}
= 0.
\]

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 $\theta$ minimizes the hedging error if and only if it solves

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

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

$$h : \mathbb{R}^n \to \mathbb{R} \\
\theta(t) \mapsto \theta(t)^T M(t) \theta(t) - 2b(t)^T \theta(t) + c(t)$$

in (12) is minimized point-wise 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 $\theta(t)$ 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. $\square$

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

$$\bar{\theta}(t) := \frac{D_x F \sigma_t Q \sigma_t^* D_z \hat{V} + \int_E \delta F(t, y) \delta \hat{V}(t, y) \nu(dy)}{e^{-rt} \kappa(t)}.$$  

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 $Q = \text{Id}_{|\mathbb{R}}$, $\eta \equiv \text{Id}_{|\mathbb{R}}$ and $\kappa \equiv 1$. Furthermore, in this case the stock price is modeled by

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

with an appropriate drift term $\gamma$. The option price can be written as a function of $S_t$:

$$\tilde{V}(t, S_t) := e^{rt} \hat{V}(t, Z_t).$$

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

$$D_z \tilde{V}(t, Z_t) = e^{-rt} D_S \tilde{V}(t, S_t) S_t.$$
Finally, we calculate
\[ D_x F(t, X_t-) = F(t, X_t-) = S_{t-}, \quad \delta F(t, y) = (e^y - 1)S_t, \]
and
\[ \delta \tilde{V}(t, y) = \tilde{V}(t, S_{t-} + y) - \tilde{V}(t, Z_{t-}) = e^{-rt} \left[ \tilde{V}(t, S_{t-}e^y) - \tilde{V}(t, S_{t-}) \right]. \]
Putting everything together, we can do a change of variable in (15) and obtain
\[ \theta(t) = \sqrt{\bar{\sigma}^2 + \int_\mathbb{R} (e^y - 1)^2 v(dy)}. \]

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
\[ F(x + \xi) = F(x) + D_x F(x) \xi + \frac{1}{2} \langle D_x^2 F(x) \xi, \xi \rangle_H + o(\| \xi \|^2_H) \]
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 $x \mapsto \partial F(x) \in L(H, \mathbb{R})$ 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 $\xi_1, \xi_2 \in H$, the derivatives satisfy

$$
D_x F(x) \xi = \sum_{k \in \mathbb{N}} \langle \omega, e_k \rangle_H \langle f_0, e_k \rangle_H e^{(x, e_k)H} \langle \xi, e_k \rangle_H \quad \text{and}
$$

$$
D_x^2 F(x) (\xi_1, \xi_2) = \sum_{k \in \mathbb{N}} \langle \omega, e_k \rangle_H \langle f_0, e_k \rangle_H e^{(x, e_k)H} \langle \xi_1, e_k \rangle_H \langle \xi_2, e_k \rangle_H.
$$

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

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

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

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

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

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

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

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

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

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

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

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

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

$$
D_x F(x) = \sum_{k \in \mathbb{N}} c_k e^{(x, e_k)H} e_k.
$$
By the very same arguments as for the first derivative, we obtain
\[
\frac{\partial}{\partial \xi} D_x F(x) = \sum_{k \in \mathbb{N}} c_k e^{(x, e_k)_H} \langle e_k, e_k \rangle_H e_k.
\]
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
\[
D_x^2 F : \begin{cases}
H \to \mathcal{L}_{HS}(H, H) \\
x \mapsto D_x^2 F(x)
\end{cases}
\]
is uniformly continuous on bounded subsets.

**Proof.** The Hilbert-Schmidt norm of \( D_x^2 F(x) \) is given by
\[
\| D_x^2 F(x) \|^2_{\mathcal{L}_{HS}(H, H)} = \sum_{k \in \mathbb{N}} \langle D_x^2 F(x) e_k, D_x^2 F(x) e_k \rangle_H = \sum_{k \in \mathbb{N}} \langle f_0, e_k \rangle_H^2 \langle \omega, e_k \rangle_H^2 e^{2(x, e_k)_H} \leq \| f_0 \|^2_H \| \omega \|^2_H e^{2\|x\|_H}.
\]
A similar calculation shows
\[
\| D_x^2 F(x_1) - D_x^2 F(x_2) \|^2_{\mathcal{L}_{HS}(H, H)} = \sum_{k \in \mathbb{N}} \langle f_0, e_k \rangle_H^2 \langle \omega, e_k \rangle_H^2 \left( e^{(x_1, e_k)_H} - e^{(x_2, e_k)_H} \right)^2 \leq \| f_0 \|^2_H \| \omega \|^2_H e^{2\max\{\|x_1\|_H, \|x_2\|_H\}} \| x_1 - x_2 \|^2_H
\]
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 [23, Th. 8.23]. This will be useful for splitting the result of Itô’s formula in a martingale and a finite variation part.

**Lemma 4.3.** The integrals
\[
\int_0^T \int_E E |F(X_{t-} + \eta_t(y)) - F(X_{t-})|^2 \nu(dy)dt
\]
and
\[
\int_0^T \int_E E |D_x F(X_{t-}) \eta_t(y)|^2 \nu(dy)dt
\]
are both finite.
We proceed with Proposition 2. We plug in the derivative of Itô’s formula to derive the stochastic dynamics of $X_t$. This expression is finite by Proposition 2.4. Similarly,

$$
\int_0^T \int_E \left| F(X_t + \eta_t(y)) - F(X_t) \right|^2 \nu(dy)dt = \lambda \int_0^T \left| \langle \omega, f_{t-} \rangle_H \right|^2 dt 
\leq \lambda \| \omega \|^2_H \int_0^T E \| f_{t-} \|^2_H dt.
$$

This is finite by Proposition 2.4 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

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

We proceed with Proposition 2.4 and Assumption 2.1 as above and the proof is finished. \qed

**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 $F : H \rightarrow \mathbb{R}$ defined in (8) satisfies

$$
dF(X_t) = \frac{1}{2} \text{tr} \left( D^2_x F(X_t) \sigma_t - Q \sigma_t \right) dt 
+ \int_E \left\{ F(X_t + \eta_t(y)) - F(X_t) - D_x F(X_t) \eta_t(y) \right\} \nu(dy) dt 
+ D_x F(X_t) \gamma_t dt + D_x F(X_t) \sigma_t d\mathcal{W}_t 
+ \int_E \left[ F(X_t + \eta_t(y)) - F(X_t) \right] \tilde{M}(dy, dt).
$$

(18)
Proof. By Theorems 4.1 and 4.2, the assumptions of [23, Th. D.2] (Itô’s formula) are satisfied. The theorem yields

\begin{equation}
F(X_t) = F(X_0) + \int_0^t D_x F(X_{s-}) \, dX_s + \frac{1}{2} \int_0^t D_x^2 F(X_{s-}) \, d[X, X]^c_s
\end{equation}

(19)

where \([X, X]^c\) denotes the continuous part of the predictable quadratic covariation as defined in [23]. By definition, we have

\[ [X, X]^c = \sum_{i,j \in \mathbb{N}} e_i \otimes e_j \left( [X, X]^c_{ij} \right), \]

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 \(P_i\) denote the operator represented by the basis element \(e_i\) \((i \in \mathbb{N})\), and \(P_i^*\) be its adjoint operator:

\[ P_i : \mathbb{R} \rightarrow \mathbb{R} \quad \text{and} \quad P_i^* : \mathbb{R} \rightarrow \mathbb{R} \quad a \rightarrow a \cdot e_i. \]

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

\[ [X_{ij}, X_{ij}]^c_t = \left[ \left\langle \int_0^t \sigma_s \, dW_s, e_i \right\rangle_H, \left\langle \int_0^t \sigma_s \, dW_s, e_j \right\rangle_H \right]_t \]

\[ = \int_0^t P_i^* \sigma_s Q \sigma_s^* P_j s \, ds = \int_0^t \left\langle \sigma_s Q \sigma_s^* e_j, e_i \right\rangle \, ds. \]

Thus, we have

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

\[ = \int_0^t \sum_{j \in \mathbb{N}} D_x^2 F(X_{s-}) \left( \sigma_s Q \sigma_s^* e_j, e_j \right) \, ds \]

\[ = \int_0^t \text{tr} \left( D_x^2 F(X_{s-}) \sigma_s Q \sigma_s^* \right) \, ds \]

(20)

It remains to reorganize the jump terms in (19). Due to Lemma 4.3, the following holds:

\[ \int_0^t D_x F(X_{s-}) \int_E \eta_s(y) \tilde{M}(dy, ds) \]

\[ = \sum_{0 \leq s \leq t} D_x F(X_{s-}) (X_s - X_{s-}) - \int_0^t \int_E D_x F(X_{s-}) \eta_s(y) \nu(dy, ds). \]
Moreover, we have
\[
\sum_{0 \leq s \leq t} \left[ F(X_s) - F(X_{s-}) \right] = \int_0^t \int_E \left[ F(X_{s-} + \eta_s(y)) - F(X_{s-}) \right] \tilde{M}(dy, ds)
\]
\[+ \int_0^t \int_E \left[ F(X_{s-} + \eta_s(y)) - F(X_{s-}) \right] \nu(dy) ds.
\]
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 [23, Ths. 8.7, 8.23]. Since continuous martingales of finite variation are a.s. constant [25, 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 3.6. Similar to Lemma 4.3, we need a technical lemma in order to be able to rearrange the jump terms in the dynamics of \( \tilde{V} \).

**Lemma 4.5.** The integrals
\[
\int_0^T \int_E E \left| \tilde{V}(t, Z_{t-} + \eta_t(y)) - \tilde{V}(t, Z_{t-}) \right|^2 \nu(dy) dt
\]
and
\[
\int_0^T \int_E E \left| D_z \tilde{V}(t, Z_{t-}) \eta_t(y) \right|^2 \nu(dy) dt
\]
are both finite.

**Proof.** By definition of \( \tilde{V} \) and the Lipschitz continuity of the payoff (Assumption 3.2) we obtain
\[
\left| \tilde{V}(t, Z_{t-} + z) - \tilde{V}(t, Z_{t-}) \right| \leq e^{-rT} E |G(Z_T + z) - G(Z_T)| \big| F_t \big|
\]
\[\leq e^{-rT} K_G \| z \|_H
\]
for every \( z \in H \). Hence, the first integral satisfies
\[
\int_0^T \int_E E \left| \tilde{V}(t, Z_{t-} + \eta_t(y)) - \tilde{V}(t, Z_{t-}) \right|^2 \nu(dy) dt \leq e^{-2rT} K_G^2 T \int_E \| y \|^2_H \nu(dy) < \infty.
\]
For the second integral, note that by (21) we have
\[
\| D_z \tilde{V}(t, Z_{t-}) \|_{L(H, R)} \leq e^{-rT} K_G.
\]
This implies
\[
\int_0^T \int_E E \left| D_z \tilde{V}(t, Z_{t-}) \eta_t(y) \right|^2 \nu(dy) dt \leq e^{-2rT} K_G^2 \int_0^T \| \eta_t(y) \|^2_H \nu(dy) dt < \infty.
\]
\[\square\]
The assumptions made for $\hat{V}$ 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 $\hat{V}$, using once again Itô’s formula on Hilbert spaces. As before, we denote by $E_0(C_{X(T)})$ the kernel of the covariance operator $C_{X(T)}$ and its orthogonal complement by $E_0(C_{X(T)})^\perp$. Note that $C_{X(T)}$ 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 [23, Th. D.2] to obtain

$$
\hat{V}(t, Z_t) = \hat{V}(0, Z_0) + \int_0^t \partial_t \hat{V}(s, Z_{s-}) \, ds + \int_0^t D_z \hat{V}(s, Z_{s-}) \, dZ_s \\
+ \frac{1}{2} \int_0^t D_z^2 \hat{V}(s, Z_{s-}) \, d[Z, Z]_s \\
+ \sum_{0 \leq s \leq t} \left\{ \hat{V}(s, Z_s) - \hat{V}(s, Z_{s-}) - D_z \hat{V}(s, Z_{s-}) (Z_s - Z_{s-}) \right\}.
$$

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

$$
d\hat{V}(t, Z_t) = \\
\partial_t \hat{V}(t, Z_{t-}) \, dt + \frac{1}{2} \, \text{tr} \left( D_z^2 \hat{V}(t, Z_{t-}) \, \sigma_{t-} Q \sigma_{t-}^* \right) \, dt \\
+ \int_E \left\{ \hat{V}(t, Z_{t-} + \eta(y)) - \hat{V}(t, Z_{t-}) - D_z \hat{V}(t, Z_{t-}) \, \eta(y) \right\} v(dy) \, dt \\
+ D_z \hat{V}(t, Z_{t-}) \, \sigma_t \, dW_t \\
+ \int_E \left[ \hat{V}(t, Z_{t-} + \eta(y)) - \hat{V}(t, Z_{t-}) \right] \tilde{M}(dy, dt).
$$

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

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

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. Similar 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 critical reading of the manuscript.

**REFERENCES**


