

# Numerical Hedging of Electricity Contracts Using Dimension Reduction

Peter Hepperger \*

The basic contracts traded on energy exchanges involve fixed-rate payments for the delivery of electricity over a certain period of time. It has been shown that options on these electricity swaps can be priced efficiently using a Hilbert space-valued time-inhomogeneous jump-diffusion model for the forward curve. We consider the mean-variance hedging problem for European options under this model. We use portfolios containing only traded contracts. The computation of hedging strategies leads to quadratic optimization problems whose parameters depend on the solution of an infinite-dimensional partial integro-differential equation. The main objective of this article is to find an efficient numerical algorithm for this task. Using proper orthogonal decomposition (a dimension reduction method), approximately optimal strategies are computed. We prove convergence of the corresponding hedging error to the minimal achievable error in the incomplete electricity market. Numerical experiments are performed to analyze the resulting hedging strategies.

**Keywords** electricity options, hedging, Hilbert space-valued jump-diffusion, partial integro-differential equation, proper orthogonal decomposition

**Mathematics Subject Classification** 91B25, 62H25, 35R15, 60H30

## 1 Introduction

There are several important differences between electricity and other traded commodities [2]. Most of them are related to the fact that electricity cannot be stored efficiently. Not only does this break the close relation of spot and forward prices, it also changes the type of contracts which are available on energy exchanges. Usually, electricity contracts are of swap type. These are agreements on the constant delivery of a certain

---

\*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.

amount of electricity (usually 1 MW) during a delivery period for fixed-rate payments. Options written on such swaps are known as electricity swaptions.

While this type of contracts bears some similarity to interest rate markets, there are major differences concerning the hedging problem for electricity. There are no contracts similar to bonds, which imply delivery at a single point in time. Moreover, the driving factors of the forward curve are not necessarily related to the short rates and thus the discounting factors. In particular, we will assume a constant interest rate for simplicity. Electricity options can be interpreted as basket options on a continuum of (correlated) assets. Each “asset” corresponds to a single time of maturity in the delivery period. The challenge here is that we cannot trade in each of the “assets” but only in various averages, the swaps. We may, for example, want to hedge a monthly option with the monthly swap and four weekly swaps. Consequently, the market is inherently incomplete, even under a pure diffusion model.

We will use a Hilbert space-valued stochastic process to model the forward curve of electricity prices. Such infinite-dimensional models have been studied in the context of interest rate markets in [3, 4, 7]. Forward curve models have also been proposed for the electricity market [1]. The jump-diffusion model we consider here has been presented in [10].

Based on this model, we propose an algorithm to solve the quadratic hedging problem for European options. In [9], it has been shown that the mean-variance hedging strategy in this setting is obtained by solving linear equation systems. The parameters of these equations, however, depend on the solution of an infinite-dimensional partial integro-differential equation (PIDE) for the swaption price. The main goal of this article is to present an efficient numerical method for the actual computation of an optimal hedging strategy.

**Outline of the Article** We start with a short introduction to electricity swaps and options in Section 2. We also state the Hilbert space-valued forward curve model there. For a more detailed analysis of the model, we refer to [9].

Section 3 is concerned with the quadratic hedging problem for a swaption. Given an arbitrary portfolio of swaps, we state the quadratic optimization problem which characterizes the optimal hedging strategy. The computational solution of this problem is based on dimension reduction. Using proper orthogonal decomposition, which is in many respects similar to principal component analysis, we derive a projected problem. Thus, we obtain a finite-dimensional approximation. We show convergence of the approximation error.

Finally, we demonstrate the effectiveness of the algorithm with numerical experiments in Section 4. We discuss the resulting hedging strategies for a test problem, using various sets of traded swaps. Residual hedging errors are evaluated by Monte Carlo simulations of the forward curve. The error is shown to decrease with the number of available underlyings. Given a sufficiently diverse portfolio of traded swaps, a very precise hedge is achieved.

## 2 A Hilbert Space-Valued Model for Electricity Swaps

This section contains a short introduction of forward curves, swap prices, and options in the electricity market. We also give a compact overview of the function-valued stochastic process which we use to model the forward curve. This model has been discussed in detail in [9].

### 2.1 Electricity Swaps and Options

The typical contract on an energy exchange is an electricity swap. At time  $t \geq 0$ , two parties agree on the 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. For every maturity  $u \in [T_1, T_2]$ , let

$$(1) \quad f(t, u) := \lim_{v \rightarrow u} F(t; u, v)$$

be the corresponding value of the forward curve at time  $t \leq u$ . Let  $r$  be the constant risk free interest rate. Due to no-arbitrage considerations, the swap rate  $F$  can be written as the weighted integral

$$(2) \quad F(t; T_1, T_2) = \int_{T_1}^{T_2} \omega(u) f(t, u) du,$$

with the nonnegative discounting factor

$$(3) \quad \omega(u) := \omega(u; T_1, T_2) := \frac{e^{-ru}}{\int_{T_1}^{T_2} e^{-rv} dv}.$$

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 on the German Energy Exchange EEX are displayed. 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., [1, 12].

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

$$(4) \quad E \left[ \left( \int_{T_1}^{T_2} e^{-r(u-T)} F(T; T_1, T_2) du - \int_{T_1}^{T_2} e^{-r(u-T)} K du \right)^+ \middle| \mathcal{F}_t \right] \\ = \kappa(T; T_1, T_2) E \left[ \left( F(T; T_1, T_2) - K \right)^+ \middle| \mathcal{F}_t \right],$$

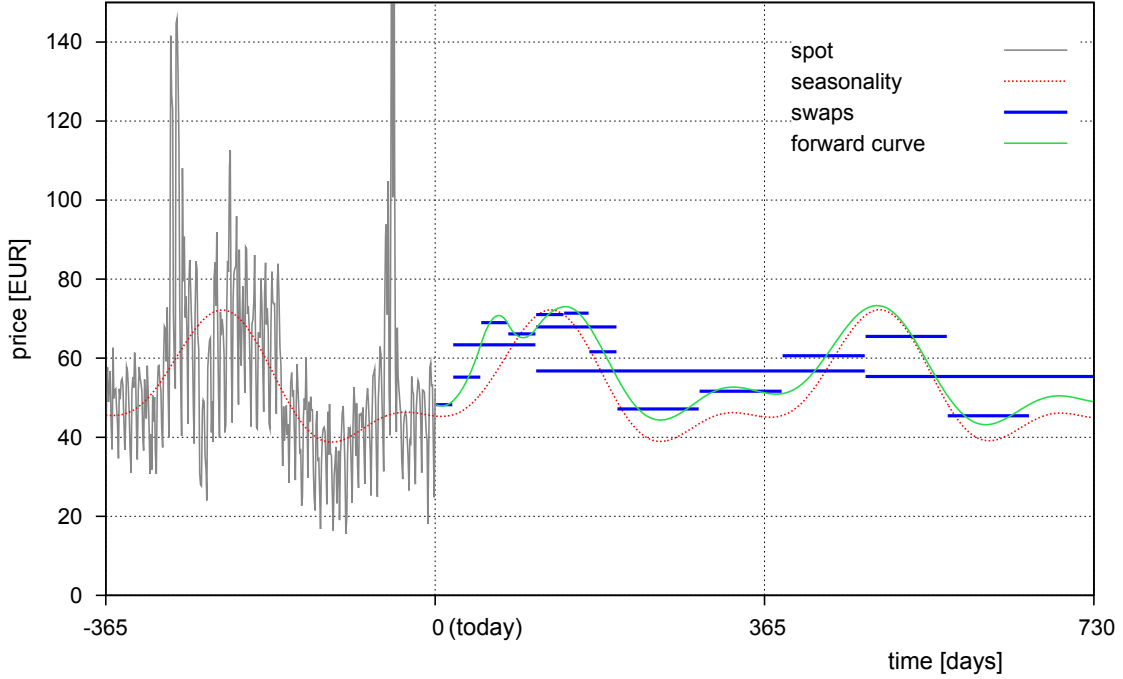


Fig. 1: EEX spot price data, fitted seasonality, traded swaps, and interpolated forward curve.

where

$$(5) \quad \kappa := \kappa(T; T_1, T_2) := \int_{T_1}^{T_2} e^{-r(u-T)} du.$$

## 2.2 Forward Curve Model

In order to model the forward curve, we employ an exponential additive process (also known as exponential time-inhomogeneous Lévy process) of jump-diffusion type. The model is assumed to be stated under the risk neutral pricing measure.

**Hilbert Space-Valued Exponential Jump-Diffusion** We consider forward curves which are elements of a separable Hilbert space  $H$ . Usually, we will choose functions defined on a settlement period  $D := [T_1, T_2]$ , setting

$$(6) \quad H := L^2([T_1, T_2]; \mu_D),$$

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

$$(7) \quad \|h\|_H := \sqrt{\int_{T_1}^{T_2} [h(u)]^2 \mu_D(u)}.$$

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

$$(8) \quad X_t := X(t) = \int_0^t \gamma(s) ds + \int_0^t \sigma(s) dW(s) + \int_0^t \int_H [\eta(s)](\xi) \tilde{M}(d\xi, ds), \quad t \geq 0.$$

The diffusion part is driven by an  $H$ -valued Wiener process  $W$  whose covariance is a symmetric nonnegative definite trace class operator  $Q$ . The jumps are characterized by  $\tilde{M}$ , the compensated random measure of an  $H$ -valued compound Poisson process

$$(9) \quad J(t) = \sum_{i=1}^{N(t)} Y_i, \quad t \geq 0,$$

which is independent of  $W$ . Here,  $N$  denotes a Poisson process with intensity  $\lambda$  and  $Y_i \sim P^Y$  ( $i = 1, 2, \dots$ ) are iid on  $H$  (and independent of  $N$ ). The corresponding Lévy measure is denoted by  $\nu = \lambda P^Y$ . We denote by  $L(H, H)$  the space of all bounded linear operators on  $H$ . We assume the drift  $\gamma : [0, T] \rightarrow H$ , the volatility  $\sigma : [0, T] \rightarrow L(H, H)$ , and the jump dampening factor  $\eta : [0, T] \rightarrow L(H, H)$  to be deterministic functions.

Subsequently, we will write  $f_t(u) := [f(t)](u)$  for every function  $f : [0, T] \rightarrow H$ ,  $u \in D$ , and similarly  $g_t(h) := [g(t)](h)$  for every  $g : [0, T] \rightarrow L(H, H)$ ,  $h \in H$ . The following hypothesis is assumed to hold.

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

$$(10) \quad E[e^{2\|Y\|_H}] = \int_H e^{2\|\xi\|_H} P^Y(d\xi) < \infty.$$

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

$$(11) \quad \gamma \in L^2(0, T; H), \quad \text{and} \quad \sigma \in L^2(0, T; L(H, H)).$$

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

$$(12) \quad \mathcal{C}_{X(T)} : \begin{cases} H \rightarrow H' \cong H, \\ h \mapsto E[\langle X_T - E[X_T], h \rangle_H \langle X_T - E[X_T], \cdot \rangle_H] \end{cases}$$

is a symmetric nonnegative definite trace class operator by [10, Thm. 2.4].

The forward curve  $f_t \in H$  is modeled as the exponential of the driving process  $X$  in the following sense. Let  $\{e_k\}_{k \in \mathbb{N}}$  be an arbitrary orthonormal basis of  $H$  and set

$$(13) \quad f_t := \sum_{k \in \mathbb{N}} \langle f_0, e_k \rangle_H e^{\langle X_t, e_k \rangle_H} e_k$$

for  $t \geq 0$ , with  $f_0 \in H$ . By [9, Thm. 2.3],  $(f_t)_{0 \leq t \leq T}$  is an  $H$ -martingale in the sense of Kunita [13] if and only if

$$(14) \quad \gamma_t = \sum_{k \in \mathbb{N}} \left[ -\frac{1}{2} \langle [\sigma_t Q \sigma_t^*](e_k), e_k \rangle_H - \int_H \left( e^{\langle \eta_t(\xi), e_k \rangle_H} - 1 - \langle \eta_t(\xi), e_k \rangle_H \right) \nu(d\xi) \right] e_k$$

for a.e.  $t \in [0, T]$ . Since forward prices are martingales under the risk neutral measure, we will subsequently assume that  $\gamma$  is defined by (14).

**Swap Rates and Option Prices** The swap rate  $F(t; T_1, T_2)$  (for a fixed delivery period  $[T_1, T_2]$ ) defined in (2) is the real-valued martingale

$$(15) \quad (F(t; T_1, T_2))_{0 \leq t \leq T} = (\langle \omega, f_t \rangle_H)_{0 \leq t \leq T}$$

The forward curve is a deterministic function of the process  $X$ , defined in (13). We may thus define

$$(16) \quad F : \begin{cases} H \rightarrow \mathbb{R}, \\ x \mapsto \langle \omega, \sum_{k \in \mathbb{N}} \langle f_0, e_k \rangle_H \exp(\langle x, e_k \rangle_H) e_k \rangle_H \end{cases}$$

and obtain (with slight abuse of notation)

$$(17) \quad F(X_t) = F(t; T_1, T_2), \quad t \in [0, T].$$

The function  $F$  is twice continuously Fréchet differentiable by [9, Thm. 4.1].

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

$$(18) \quad Z_t := X_t - E[X_t] = \int_0^t \sigma_s dW_s + \int_0^t \int_H \eta_s(\zeta) \tilde{M}(d\zeta, ds).$$

The price of the option at time  $t \leq T$ , discounted to time 0, is

$$(19) \quad \hat{V}(t, z) := e^{-rT} E[G(Z_T) | Z_t = z],$$

where  $G$  is the payoff function of the option in terms of  $Z_T$ . We denote the trace of a nuclear operator  $A$  by  $\text{tr}(A)$ , and the space of Hilbert-Schmidt operators defined on  $H$  by  $L_{\text{HS}}(H, H) \subset L(H, H)$ . The following assumptions are adopted from [9].

**Assumption 2.2.** *Suppose that the payoff function  $G$  is Lipschitz continuous on  $H$  with Lipschitz constant  $K_G$ . Suppose further 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  $D_z^2 \hat{V}(t, z) \in L_{\text{HS}}(H, H)$  for every  $(t, z) \in [0, T] \times H$ , and the mapping  $D_z^2 : (t, z) \rightarrow L_{\text{HS}}(H, H)$  is locally uniformly continuous.*

Applying Itô's formula on Hilbert spaces, we obtain the following proposition concerning  $\hat{V}$ . We denote by  $E_0(\mathcal{C}_{X(T)})$  the eigenspace of the covariance operator  $\mathcal{C}_{X(T)}$  corresponding to eigenvalue 0 (the kernel). Its orthogonal complement is  $E_0(\mathcal{C}_{X(T)})^\perp$ .

**Proposition 2.1.** *The discounted option price  $\hat{V}$  defined in (19) satisfies*

$$(20) \quad d\hat{V}(t, Z_t) = D_z \hat{V}(t, Z_{t-}) \sigma_t dW_t + \int_H [\hat{V}(t, Z_{t-} + \eta_t(\zeta)) - \hat{V}(t, Z_{t-})] \tilde{M}(d\zeta, dt).$$

Moreover, it is a classical solution of the partial integro-differential equation

$$(21) \quad \begin{aligned} -\partial_t \hat{V}(t, z) &= \frac{1}{2} \text{tr} [D_z^2 \hat{V}(t, z) \sigma_t Q \sigma_t^*] \\ &+ \int_H \left\{ \hat{V}(t, z + \eta_t(\zeta)) - \hat{V}(t, z) - [D_z \hat{V}(t, z)] \eta_t(\zeta) \right\} \nu(d\zeta), \end{aligned}$$

with terminal condition

$$(22) \quad \widehat{V}(T, z) = e^{-rT} G(z),$$

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

*Proof.* These results were shown in [9, Thm. 3.5]. □

### 3 Computing an Optimal Hedging Strategy

In this section, we propose a numerical algorithm which can be used to compute the mean-variance hedging strategy for an option. We start with a formal description of the portfolio of swaps and the hedging problem. It has been shown in [9] that every optimal strategy solves a linear equation system for each point in time.

There are two numerical issues when solving these equations. The first is their dependence on the solution of the infinite-dimensional PIDE (21), which cannot be solved directly. Secondly, the matrix corresponding to the linear equations is possibly semi-definite and thus singular. We solve the first and more important issue by applying a dimension reduction technique called proper orthogonal decomposition. This is closely related to principal component analysis. The method enables us to compute an approximate hedging strategy. We show that the corresponding hedging error converges to the minimal possible error (the market is inherently incomplete). The issue of singularity is easier to remedy. Depending on the portfolio, either another projection, or an additional regularization can be used to obtain a solution.

#### 3.1 Optimal Hedging Strategies

We consider a European option with delivery period  $[T_1, T_2]$ , maturity  $T \leq T_1$ , and strike rate  $K$ . We are looking for a hedging strategy using a portfolio of swaps with various delivery periods. We may for example want to hedge a quarterly option by trading in the quarterly swap itself as well as three monthly swaps. More generally, there are  $n$  swap contracts available for trading whose delivery periods are given by  $[T_1^i, T_2^i]$ ,  $i = 1, \dots, n$ . The corresponding swap rates satisfy

$$(23) \quad F(t; T_1^i, T_2^i) = \int_{T_1^i}^{T_2^i} \omega_i(u) f(t, u) du,$$

where

$$(24) \quad \omega_i(u) := \omega(u; T_1^i, T_2^i) = \frac{e^{-ru}}{\int_{T_1^i}^{T_2^i} e^{-ru} du}$$

is the discounting factor defined in (3). Since we cannot hedge with swaps whose delivery periods start before maturity of the option, we will assume  $T \leq T_1^i$  for every  $i = 1, \dots, n$ . Similar to (16), we define  $F_i(X_t) := F(t; T_1^i, T_2^i)$ ,  $t \in [0, T]$ .

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  $\theta(t) = (\theta_1(t), \dots, \theta_n(t)) \in \mathbb{R}^n$  describes the investment in each swap at time  $t$ . The value  $S_0$  of the risk free asset solves the differential equation

$$(25) \quad dS_0(t) = rS_0(t)dt.$$

The value of the portfolio at time  $t$  is denoted by  $V^\theta(t)$ . Since a swap has no inherent value (you can enter the contract without paying anything), we have

$$(26) \quad V^\theta(t) = \theta_0(t)S_0(t).$$

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:

$$(27) \quad d\widehat{V}^\theta(t) = \sum_{i=1}^n \theta_i(t)e^{-rt}\kappa_i dF_i(t),$$

where

$$(28) \quad \kappa_i := \kappa(T; T_1^i, T_2^i) = \int_{T_1^i}^{T_2^i} e^{-r(u-T)} du.$$

The discounting factor  $\kappa$  has been introduced in (5). A strategy  $(\theta_0(t), \theta(t))$  is admissible, if it is predictable, càglàd, and satisfies

$$(29) \quad E \left| \int_0^T \sum_{i=1}^n \theta_i(t)e^{-rt}\kappa_i dF_i(t) \right|^2 < \infty.$$

Mean-variance hedging consists in minimizing the expected global quadratic hedging error

$$(30) \quad J(\theta) := E \left| \widehat{V}^\theta(T) - \widehat{V}(T) \right|^2.$$

The following hypothesis guarantees that the driving stochastic processes do have influence on at least one of the traded swaps. To this end, we define the matrix valued process  $M$ , which will play an essential role in our algorithm.

$$(31) \quad \begin{aligned} m_{ij}(t) &:= e^{-2rt}\kappa_i\kappa_j \left( D_x F_i \sigma_t Q \sigma_t^* D_x F_j + \int_H \delta F_i \delta F_j \nu(d\zeta) \right), \quad i, j = 1, \dots, n, \\ M(t) &:= (m_{ij}(t))_{i,j=1}^n \in \mathbb{R}^{n \times n}. \end{aligned}$$

**Assumption 3.1.** *We assume that  $M(t) \neq 0$  for almost every  $t \in [0, T]$ .*

The following proposition states a representation of the optimal hedging strategy which will be the basis of the numerical algorithm. In order to simplify and shorten notation, we define abbreviations for the jumps of swap rates and option price:

$$(32) \quad \begin{aligned} \delta F_i(t, \zeta) &:= F_i(X_{t-} + \eta_t(\zeta)) - F_i(X_{t-}), \quad i = 1, \dots, n, \\ \delta \widehat{V}(t, \zeta) &:= \widehat{V}(t, Z_{t-} + \eta_t(\zeta)) - \widehat{V}(t, Z_{t-}). \end{aligned}$$



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  $D_z \widehat{V}(t, Z_{t-})$ .

**Proposition 3.1.** *Let  $M(t) \in \mathbb{R}^{n \times n}$  be the matrix valued process defined in (31). Define further*

$$(33) \quad b_i(t) := e^{-rt} \kappa_i \left( D_x F_i \sigma_i Q \sigma_i^* D_z \widehat{V} + \int_H \delta F_i \delta \widehat{V} \nu(d\zeta) \right), \quad i = 1, \dots, n.$$

*An investment strategy  $\bar{\theta}$  minimizes the hedging error if and only if it solves the quadratic program*

$$(34) \quad \bar{\theta} = \operatorname{argmin}_{\theta(t)} \theta(t)^T M(t) \theta(t) - 2b(t)^T \theta(t), \quad \text{for a.e. } t \in [0, T].$$

*This is equivalent to the linear equation system*

$$(35) \quad M(t) \bar{\theta}(t) = b(t) \quad \text{for a.e. } t \in [0, T].$$

*There is at least one solution to this equation, i.e.  $b(t)$  is in the range of  $M(t)$ .*

*Proof.* This result is taken from [9, Thm. 3.9]. □

Note that the solution of the equation system (35) is a vector  $\bar{\theta} \in \mathbb{R}^n$ . In contrast to hedging strategies for single stocks, it does not necessarily satisfy  $\bar{\theta} \in [0, 1]^n$ . In fact, arbitrary positive and negative values may occur, which means that short-selling swaps may be part of an optimal strategy. It is of course possible to add constraints on  $\theta$  to the optimization problem. Instead of the linear system, one then solves, e.g., the quadratic program

$$(36) \quad \begin{aligned} \bar{\theta} &= \operatorname{argmin}_{\theta(t)} \theta(t)^T M(t) \theta(t) - 2b(t)^T \theta(t), & t \in [0, T], \\ &\text{subject to } \bar{\theta}(t) \geq 0, \end{aligned}$$

where the inequality is interpreted component-wise. As we will see in the numerical experiments in section 4, however, the restriction to nonnegative hedging strategies may lead to considerably higher hedging errors.

### 3.2 Proper Orthogonal Decomposition (POD)

We will now discuss the numerical approximation of the optimal hedge  $\bar{\theta}(t)$  in Proposition 3.1. Since the functions  $F_i$  and their derivatives are known analytically (compare [9, Thm. 4.1]), the main task here is to approximate  $D_z \widehat{V}$  and  $\delta \widehat{V}$ . These are elements of the Hilbert space  $H$ , so we will construct finite-dimensional approximations. To this end, we apply the dimension reduction method based on proper orthogonal decomposition (POD) presented in [10]. The basic idea is to find an orthonormal basis of  $H$  such that a small number of basis elements is sufficient to obtain a good approximation for the terminal value  $Z(T)$ .

**Definition 3.2.** A sequence of orthonormal elements  $\{p_l\}_{l \in \mathbb{N}} \subset H$  is called a POD-basis for  $Z(T)$ , if it solves the minimization problem

$$(37) \quad \min_{\langle p_i, p_j \rangle_H = \delta_{ij}} E \left[ \left\| Z(T) - \sum_{l=1}^d p_l \langle Z(T), p_l \rangle_H \right\|_H^2 \right]$$

for every  $d \in \mathbb{N}$ .

The following theorem shows how to obtain such an orthonormal POD basis by computing eigenvectors of the covariance operator  $\mathcal{C}_{X(T)}$ .

**Theorem 3.3.** A sequence of orthonormal eigenvectors  $(p_l)_{l \in \mathbb{N}}$  of operator  $\mathcal{C}_{X(T)}$ , ordered by the size of the corresponding eigenvalues  $\mu_1 \geq \mu_2 \geq \dots \geq 0$ , solves the maximization problem

$$(38) \quad \max_{\langle p_i, p_j \rangle_H = \delta_{ij}} \sum_{l=1}^d \langle \mathcal{C}_{X(T)} p_l, p_l \rangle_H$$

for every  $d \in \mathbb{N}$ . The maximum value is

$$(39) \quad \sum_{l=1}^d \langle \mathcal{C}_{X(T)} p_l, p_l \rangle_H = \sum_{l=1}^d \mu_l.$$

Moreover, the eigenvectors are a POD basis in the sense of Definition 3.2, and the expectation of the projection error is

$$(40) \quad E \left[ \left\| Z_T - \sum_{l=1}^d p_l \langle Z_T, p_l \rangle_H \right\|_H^2 \right] = \sum_{l=d+1}^{\infty} \mu_l.$$

*Proof.* This is an application of [15, Thm. 2.7 and Prop. 2.8].  $\square$

Subsequently, let  $(p_l)_{l \in \mathbb{N}}$  and  $(\mu_l)_{l \in \mathbb{N}}$  denote the orthonormal basis and eigenvalues from Theorem 3.3. Further, let

$$(41) \quad U_d := \text{span}\{p_1, p_2, \dots, p_d\} \subset H$$

be the  $d$ -dimensional subspace spanned by the eigenvectors corresponding to the largest eigenvalues. We will assume that  $\mu_1 \geq \dots \geq \mu_d > 0$ , as there is no need to include eigenvectors of the covariance operator corresponding to eigenvalue 0. We define the projection operator

$$(42) \quad \mathcal{P}_d : \begin{cases} H & \rightarrow U_d \cong \mathbb{R}^d, \\ z & \mapsto x := \sum_{l=1}^d \langle z, p_l \rangle_H p_l \end{cases}$$

and identify  $U_d$  with  $\mathbb{R}^d$  via the isometry

$$(43) \quad \iota : \begin{cases} (U_d, \|\cdot\|_H) & \rightarrow (\mathbb{R}^d, \|\cdot\|), \\ x & \mapsto (\langle x, p_l \rangle_H)_{l=1}^d. \end{cases}$$

We introduce

$$(44) \quad \widehat{V}_d(t, x) := e^{-rT} E[G(x + \mathcal{P}_d(Z(T) - Z(t)))] ,$$

a finite-dimensional approximation of the discounted option price for  $x \in U_d \cong \mathbb{R}^d$ . It is possible to show that  $\widehat{V}_d$  is the unique solution of a  $d$ -dimensional PIDE, which can be solved numerically using sparse grid finite difference methods. The following theorem summarizes some results from [10]. Note that PIDEs used only for pricing can be constructed to have time constant coefficients. This can be achieved by introducing a Lévy process with the same terminal distribution as the driving time-inhomogeneous jump-diffusion. Here, we state a PIDE with time dependent coefficients, since hedging is path dependent.

**Theorem 3.4.** *Suppose that there is a constant  $c > 0$  such that*

$$(45) \quad \langle \sigma_t Q \sigma_t^* p, p \rangle_H \geq c \|p\|_H^2 \quad \text{for every } p \in \text{span}\{p_1, \dots, p_d\}, \text{ a.e. } t \in [0, T].$$

*Then  $\widehat{V}_d$  satisfies  $\widehat{V}_d \in C^{1,2}((0, T) \times U_d, \mathbb{R}) \cap C([0, T] \times U_d, \mathbb{R})$ . It is the unique solution of the PIDE*

$$(46) \quad \begin{aligned} -\partial_t \widehat{V}_d(t, x) &= \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t) \partial_i \partial_j \widehat{V}_d(t, x) \\ &+ \int_H \left\{ \widehat{V}_d(t, x + \mathcal{P}_d \eta_t(\zeta)) - \widehat{V}_d(t, x) - \sum_{i=1}^d \langle \eta_t(\zeta), p_i \rangle_H \partial_i \widehat{V}_d(t, x) \right\} \nu(d\zeta), \end{aligned}$$

*with time-dependent coefficients*

$$(47) \quad a_{ij}(t) := \langle \sigma_t Q \sigma_t^* p_i, p_j \rangle_H, \quad i, j = 1, \dots, d,$$

*and terminal condition*

$$(48) \quad \widehat{V}_d(T, x) = \widehat{V}(T, x) = e^{-rT} G(x),$$

*for  $t \in (0, T)$ ,  $x \in U_d$ . Moreover, there exists a constant  $C > 0$  such that*

$$(49) \quad \left| \widehat{V}_d(0, 0) - \widehat{V}(0, 0) \right| \leq C \sqrt{\sum_{l=d+1}^{\infty} \mu_l}.$$

*Proof.* The proofs can be taken almost without modification from [10, Thms. 3.6-3.13] and are therefore omitted. The only difference is that we are working with time varying coefficients  $a_{ij}(t)$  here. The uniform positive definiteness of  $\sigma_t Q \sigma_t^*$  in the hypothesis of the present theorem, however, enables us to show Gårding's inequality in exactly the same way as with time constant coefficients.  $\square$

### 3.3 Approximate Hedging with POD

The matrix  $M(t)$  in the linear equation system (35) is determined by the swap rate functions  $F_i$ ,  $i = 1, \dots, n$ , and their derivatives, which are known in closed form. Hence, they can be calculated directly, using a (possibly high-dimensional) discretization of the Hilbert space  $H$ . In contrast, the right-hand side  $b(t)$  of (35) depends on the unknown option price  $\widehat{V}$  and its derivatives. Since it is not feasible to solve the corresponding PIDE with a high-dimensional discretization, we will apply the dimension reduction method presented in the previous section.

The goal of this section is to compute the optimal hedge with equation (35) by replacing all expressions depending on  $\widehat{V}$  with similar expressions depending on  $\widehat{V}_d$ . Note that applying the chain rule yields

$$(50) \quad D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) = \frac{\partial}{\partial Z_{t-}} \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) = D_x \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) \circ \mathcal{P}_d \in L(H, \mathbb{R}).$$

The derivative  $D_x \widehat{V}_d$  is a finite-dimensional object, which we can approximate numerically by solving the PIDE for  $\widehat{V}_d$ . Similarly, the jump term

$$(51) \quad \delta \widehat{V}_d(t, \mathcal{P}_d \zeta) := \widehat{V}_d(t, \mathcal{P}_d(Z_{t-} + \zeta)) - \widehat{V}_d(t, \mathcal{P}_d Z_{t-})$$

can be evaluated numerically for every  $\zeta \in H$ .

To simplify subsequent notation, we further introduce

$$(52) \quad \langle h_1, h_2 \rangle_{\sigma_t Q \sigma_t^*} := h_1 \sigma_t Q \sigma_t^* h_2 \quad \text{and} \quad \|h_1\|_{\sigma_t Q \sigma_t^*} := \sqrt{\langle h_1, h_1 \rangle_{\sigma_t Q \sigma_t^*}},$$

for every  $h_1, h_2 \in H$ . Because of the fact that the Cauchy–Schwarz inequality holds not only for scalar products, but also for symmetric nonnegative definite bilinear forms (see, e.g., [6, Thm. 6.2.1]),  $\|\cdot\|_{\sigma_t Q \sigma_t^*}$  defines a semi-norm on  $H$ .

In order to approximate  $b$ , we define

$$(53) \quad \widetilde{b}_i(t) := e^{-rt} \kappa_i \left( \langle D_x F_i, D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) \rangle_{\sigma_t Q \sigma_t^*} + \int_H \delta F_i \delta \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) \nu(d\zeta) \right),$$

for  $i = 1, \dots, n$ , where  $\widehat{V}_d$  and  $\delta \widehat{V}_d$  are the approximations defined in (44) and (51), respectively. For  $t \in [0, T]$ , we consider a (not necessarily unique) solution  $\widetilde{\theta}$  of

$$(54) \quad M(t) \widetilde{\theta}(t) = \widetilde{b}(t),$$

which is an approximation for an optimal strategy  $\widetilde{\theta}(t)$ . Similar to (35), the linear equation for  $\widetilde{\theta}(t)$  has at least one solution, although  $M$  might be singular.

**Lemma 3.5.** *The vector  $\widetilde{b}(t)$  defined in (53) satisfies*

$$(55) \quad \forall y \in \mathbb{R}^n : (y^T M(t) = 0 \Rightarrow y^T \widetilde{b}(t) = 0)$$

for every  $t \in [0, T]$ . In particular,  $\widetilde{b}(t)$  is an element of the range of  $M(t)$ .

*Proof.* The proof is identical to the proof of [9, Lem. 3.8]. Compare also the proof of Lemma 3.6 below for a very similar argument.  $\square$

In contrast to  $\bar{\theta}$ , the expression for  $\tilde{\theta}$  contains only expressions which are either analytically known or can be approximated numerically. We will show, that the *additional* hedging error introduced by using  $\tilde{\theta}$  converges to 0 for increasing dimension  $d$  of the approximating problem. (The unhedgeable part of the hedging risk remains unchanged by definition.)

Before we can show the convergence result, we need to analyze the relation of  $b(t)$  and  $\tilde{b}(t)$  in more detail. For all subsequent computations, let  $t \in [0, T]$  be arbitrary, but fixed. Since the matrix  $M(t)$  is positive semi-definite, we can find an orthonormal basis of eigenvectors  $A = (a_1 | \dots | a_n) \in \mathbb{R}^{n \times n}$  with corresponding eigenvalues  $\lambda_i$ ,  $i = 1, \dots, n$ , such that

$$(56) \quad M(t) = A \operatorname{diag}(\lambda_1, \dots, \lambda_n) A^T,$$

with  $\operatorname{diag}(\dots)$  denoting a diagonal matrix. The following lemma states an upper bound for the projection of the difference  $b(t) - \tilde{b}(t)$  onto each of the eigenvectors  $a_i$  of  $M(t)$ .

**Lemma 3.6.** *For every  $l \in \{1, 2, \dots, n\}$ , the following estimate holds:*

$$(57) \quad \begin{aligned} & \left[ (\tilde{b}(t) - b(t))^T a_l \right]^2 \\ & \leq 2\lambda_l \left[ \left\| D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - D_z \widehat{V} \right\|_{\sigma_t Q \sigma_t^*}^2 + \int_H (\delta \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - \delta \widehat{V})^2 \nu(d\zeta) \right]. \end{aligned}$$

*Proof.* By definition of  $b(t)$  and  $\tilde{b}(t)$ , we have

$$(58) \quad \begin{aligned} \left[ (\tilde{b}(t) - b(t))^T a_l \right]^2 &= \left[ e^{-rt} \left[ \left\langle \sum_{i=1}^n a_l(i) \kappa_i D_x F_i, D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - D_z \widehat{V} \right\rangle_{\sigma_t Q \sigma_t^*} \right. \right. \\ & \quad \left. \left. + \int_H \left( \sum_{i=1}^n a_l(i) \kappa_i \delta F_i \right) (\delta \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - \delta \widehat{V}) \nu(d\zeta) \right] \right]^2 \end{aligned}$$

Applying the inequalities by Young and Cauchy–Schwarz yields

$$(59) \quad \begin{aligned} & \left[ (\tilde{b}(t) - b(t))^T a_l \right]^2 \\ & \leq 2e^{-2rt} \left\| \sum_{i=1}^n a_l(i) \kappa_i D_x F_i \right\|_{\sigma_t Q \sigma_t^*}^2 \left\| D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - D_z \widehat{V} \right\|_{\sigma_t Q \sigma_t^*}^2 \\ & \quad + 2e^{-2rt} \left( \int_H \left( \sum_{i=1}^n a_l(i) \kappa_i \delta F_i \right)^2 \nu(d\zeta) \right) \left( \int_H (\delta \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - \delta \widehat{V})^2 \nu(d\zeta) \right) \end{aligned}$$

On the other hand, the eigenvalues of  $M(t)$  satisfy

$$(60) \quad \lambda_l = a_l^T M(t) a_l = e^{-2rt} \left\| \sum_{i=1}^n a_l(i) \kappa_i D_x F_i \right\|_{\sigma_t Q \sigma_t^*}^2 + e^{-2rt} \int_H \left( \sum_{i=1}^n a_l(i) \kappa_i \delta F_i \right)^2 \nu(d\zeta).$$

Combined, we obtain (57).  $\square$

We can now prove our main convergence result. It gives a bound on the additional hedging error caused by the approximation of the optimal strategy.

**Theorem 3.7.** *There is a constant  $C > 0$  such that the additional hedging error when using  $\tilde{\theta}$ , defined in (54), instead of an optimal hedging strategy  $\bar{\theta}$ , defined in (35), satisfies*

$$(61) \quad 0 \leq J(\tilde{\theta}) - J(\bar{\theta}) \leq Ce^{-2rT} \sum_{l=d+1}^{\infty} \mu_l.$$

As before,  $\mu_l$  are the eigenvalues of the covariance operator  $C_{X(T)}$ .

*Proof.* Since  $\bar{\theta}$  is an optimal strategy, we have  $0 \leq J(\tilde{\theta}) - J(\bar{\theta})$ . Using [9, Thm. 3.7], we obtain

$$(62) \quad J(\tilde{\theta}) - J(\bar{\theta}) = E \int_0^T [\tilde{\theta}^T M \tilde{\theta}(t) - 2b^T \tilde{\theta}(t) - \bar{\theta}^T M \bar{\theta}(t) + 2b^T \bar{\theta}(t)] dt.$$

By construction,  $\tilde{\theta}$  and  $\bar{\theta}$  are solutions of

$$(63) \quad M(t)\tilde{\theta}(t) = \tilde{b}(t) \quad \text{and} \quad M(t)\bar{\theta}(t) = b(t),$$

respectively. Since  $M(t)$  might be singular, we cannot invert it. Instead, we introduce the Moore-Penrose pseudo-inverse  $M^+(t)$  of  $M(t)$ . Suppose without loss of generality that the eigenvalues of  $M(t)$  satisfy  $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ . Suppose further that  $\lambda_m > 0$  is the lowest non-zero eigenvalue. Then  $M^+(t)$  is given by

$$(64) \quad M^+(t) = A \operatorname{diag} \left( 0, \dots, 0, \frac{1}{\lambda_m}, \dots, \frac{1}{\lambda_n} \right) A^T.$$

Every pair of solutions  $(\tilde{\theta}(t), \bar{\theta}(t))$  to (63) yields the same additional hedging error. Thus, we may replace  $\tilde{\theta}(t)$  with the specific solution  $M^+(t)\tilde{b}(t)$  and  $\bar{\theta}(t)$  with  $M^+(t)b(t)$  in (62). Since the pseudo-inverse satisfies  $M^+MM^+(t) = M^+(t)$ , we obtain

$$(65) \quad \begin{aligned} J(\tilde{\theta}) - J(\bar{\theta}) &= E \int_0^T [\tilde{b}^T(t)M^+(t)\tilde{b}(t) - 2b^T(t)M^+\tilde{b}(t) + b^T(t)M^+(t)b(t)] dt \\ &= E \int_0^T (\tilde{b}(t) - b(t))^T M^+(t) (\tilde{b}(t) - b(t)) dt. \end{aligned}$$

Plugging in the definition of  $M^+(t)$  yields

$$(66) \quad J(\tilde{\theta}) - J(\bar{\theta}) = E \int_0^T \sum_{i=m}^n \frac{1}{\lambda_i} [(\tilde{b}(t) - b(t))^T a_i]^2 dt.$$

Applying Lemma 3.6, we find

$$\begin{aligned}
(67) \quad J(\tilde{\theta}) - J(\bar{\theta}) &\leq 2(n-m)E \int_0^T \left[ \|D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - D_z \widehat{V}\|_{\sigma_t, \mathbb{Q}\sigma_t^*}^2 \right. \\
&\quad \left. + \int_H (\delta \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - \delta \widehat{V})^2 \nu(d\zeta) \right] dt. \\
&= 2(n-m)E \left| \int_0^T (D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - D_z \widehat{V}) \sigma_t dW_t \right. \\
&\quad \left. + \int_0^T \int_H (\delta \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) - \delta \widehat{V}) \widetilde{M}(d\zeta, dt) \right|^2,
\end{aligned}$$

where, for the last equality, we have used [5, Cor. 4.14] for the Brownian part, [14, Thm. 8.23] for the jump part, and in addition the independence of  $W$  and  $\widetilde{M}$ .

The very same arguments as in the proof of Proposition 2.1 yield

$$(68) \quad d\widehat{V}_d(t, \mathcal{P}_d Z_t) = D_z \widehat{V}_d(t, \mathcal{P}_d Z_{t-}) \sigma_t dW_t + \int_H \delta \widehat{V}_d(t, \mathcal{P}_d \zeta) \widetilde{M}(d\zeta, dt).$$

Hence, we have

$$\begin{aligned}
(69) \quad J(\tilde{\theta}) - J(\bar{\theta}) &\leq 2nE \left| \int_0^T d\widehat{V}_d(t, \mathcal{P}_d Z_t) dt - \int_0^T d\widehat{V}(t, Z_t) dt \right|^2 \\
&= 2nE \left| [\widehat{V}_d(T, \mathcal{P}_d Z_T) - \widehat{V}_d(0, 0)] - [\widehat{V}(T, Z_T) - \widehat{V}(0, 0)] \right|^2.
\end{aligned}$$

Using Young's inequality and the definition of  $\widehat{V}$  and  $\widehat{V}_d$ , we see that

$$(70) \quad J(\tilde{\theta}) - J(\bar{\theta}) \leq 4n \left| \widehat{V}(0, 0) - \widehat{V}_d(0, 0) \right|^2 + 4ne^{-2rT} E |G(Z_T) - G(\mathcal{P}_d Z_T)|^2.$$

From Assumption 2.2 and Theorem 3.3 we finally obtain

$$(71) \quad J(\tilde{\theta}) - J(\bar{\theta}) \leq 8ne^{-2rT} K_G^2 E \|Z_T - \mathcal{P}_d Z_T\|_H^2 = 8ne^{-2rT} K_G^2 \sum_{l=d+1}^{\infty} \mu_l.$$

□

Theorem 3.7 shows that every solution of (54) is a good approximation to an optimal hedging strategy, if the dimension  $d$  is chosen appropriately. Note that although hedging errors are path-dependent, choosing a POD basis which minimizes the projection error for the terminal value  $Z_T$  is sufficient for the computation of  $\tilde{\theta}$ . We do *not* need to approximate the whole path of the process.

### 3.4 Semi-Definite Quadratic Optimization

Due to the results in the previous section, the numerical hedging algorithm will be based on solving the linear equation

$$(72) \quad M(t)\tilde{\theta}(t) = \tilde{b}(t).$$

By Lemma 3.5, there is always at least one solution. In practice, however, finding such a solution yields numerical issues, since the matrix  $M(t)$  is not *strictly* positive definite. Even small discretization or rounding errors in the computation of  $M(t)$  and  $\tilde{b}(t)$  may result in an unsolvable equation, where  $\tilde{b}(t)$  is no longer an element of the range of  $M(t)$ . For the remainder of the section, let once again  $t \in [0, T]$  be arbitrary but fixed.

Depending on the eigenvalues of  $M(t)$ , there are two possible remedies for this situation. Suppose as before that the eigenvalues satisfy  $0 \leq \lambda_1 \leq \dots \leq \lambda_n$ , with  $\lambda_m > 0$  being the lowest non-zero eigenvalue. The corresponding matrix of orthonormal eigenvectors is  $A = (a_1 | \dots | a_n)$ . We first consider the case when the ratio  $\frac{\lambda_n}{\lambda_m}$  is small, i.e.  $\lambda_m$  is well separated from 0. Instead of looking for an arbitrary solution of (72), we look for the one with the smallest norm  $\|\tilde{\theta}(t)\|_{\mathbb{R}^n}$ . Equation 72 is equivalent to

$$(73) \quad \text{diag}(0, \dots, 0, \lambda_m, \dots, \lambda_n) A^T \tilde{\theta}(t) = A^T \tilde{b}(t).$$

Since

$$(74) \quad \|\tilde{\theta}(t)\|_{\mathbb{R}^n}^2 = \sum_{i=1}^{m-1} \langle \tilde{\theta}(t), a_i \rangle_{\mathbb{R}^n}^2 + \sum_{i=m}^n \langle \tilde{\theta}(t), a_i \rangle_{\mathbb{R}^n}^2,$$

minimizing the norm of  $\tilde{\theta}(t)$  is equivalent to  $\tilde{\theta}$  being orthogonal to every  $a_i$ ,  $i = 1, \dots, m-1$ . Consequently, we may project (72) to the orthogonal complement of the kernel of  $M(t)$ . Hence, we solve

$$(75) \quad \text{diag}(\lambda_m, \dots, \lambda_n) \varphi(t) = (a_m | \dots | a_n)^T \tilde{b}(t)$$

for  $\varphi \in \mathbb{R}^{n-m+1}$ , and set

$$(76) \quad \tilde{\theta}(t) = (a_m | \dots | a_n) \varphi(t).$$

Note that the condition number of the linear system (75) is given by  $\frac{\lambda_n}{\lambda_m}$ , which is small by our hypotheses. Thus, the unique solution  $\varphi$  can be computed in a numerically stable way.

We consider now the case  $\frac{\lambda_n}{\lambda_m} \gg 1$ . The projected equation (75) is then badly conditioned. Moreover, it is difficult to identify the eigenvalues of  $M(t)$  which are equal to 0 numerically. In this case, we can either change our portfolio and include swaps which are less correlated, or we can apply regularization to the equation. Tikhonov regularization proved to be very effective in the numerical experiments. We replace (34) with the minimization problem

$$(77) \quad \bar{\theta}(t) = \operatorname{argmin}_{\theta(t)} \theta(t)^T (M(t) + \delta \|M\| I_n) \theta(t) - 2b(t)^T \theta(t), \quad \text{for a.e. } t \in [0, T],$$

where  $I_n \in \mathbb{R}^{n \times n}$  denotes the unit matrix, and  $\delta \ll 1$  is a small regularization parameter. Just like the projection method discussed above, the regularization approach also gives preference to solutions with smaller norms. For convergence results concerning the Tikhonov regularization, we refer to [8, Ch. 6.4.3].



## 4 Numerical Experiments

The computation of the optimal mean-variance hedging strategy relies on the dimension reduction presented in Section 3.2 as well as on the numerical solution of the PIDE (46) and the optimization problem (34). In order to analyze the resulting hedging strategies, these algorithms have been implemented in C++. In this final section, we will discuss the results obtained from numerical experiments with different sets of traded swaps. The test problem is described in detail below.

For each set of swaps, the average hedging error is computed by Monte Carlo (MC) simulations. To this end, a large number of stochastic paths for the forward curve are simulated. Along each path, the option prices and hedging strategies are calculated by solving the corresponding PIDEs. The values of the portfolios (defined as the initial option value plus the increments given by the dynamics (27)) are then compared to the true terminal option values. The average of their differences is a good approximation for the unhedgeable component of the option price.

### 4.1 Test Problem

We consider the test problem presented in [10]. We would like to hedge a European electricity call option with a delivery period of 28 days maturing in one year, i.e.  $T = T_1 = 1$ ,  $T_2 = T_1 + 28/365$ . We use the exponential additive model described in Section 2. The corresponding Hilbert space is  $H = L^2([T_1, T_2], \mu_D)$ , where  $D = [T_1, T_2]$  is the delivery period of the swap and  $\mu_D$  is the Lebesgue measure. The underlying stochastic process is

$$(78) \quad X(t, u) = \int_0^t \gamma(s, u) ds + \sum_{k=1}^{n_W} \int_0^t \sigma_k(s, u) dW_k(s) + \int_0^t \int_{\mathbb{R}} \eta_k(s, u) y \tilde{M}(dy, ds),$$

$u \in [T_1, T_2]$ ,  $t \in [0, T]$ . Here,  $W_k$  ( $k = 1, \dots, n_W \in \mathbb{N}$ ) denote scalar Brownian motions and  $\tilde{M}$  denotes a compensated random jump measure, all of them independent. For the diffusive part of the model, we use two factors similar to [11]. The volatilities are given by

$$(79) \quad \sigma_1(s, u) \equiv 0.15, \quad \sigma_2(s, u) = 0.3 e^{-1.4(u-s)}.$$

Moreover, we use a compound Poisson process with intensity  $\lambda = 12$  and  $\mathcal{N}(0.1, 0.1)$ -distributed jump heights, which yields a *Merton model*. Note that, since we consider the exponential process, a mean of 0.1 for the jump distribution corresponds to an average jump of slightly over 10 percent of the current value. The additional factor for dampening the jumps is

$$(80) \quad \eta(s, u) = 0.5 - 0.5 \frac{u - T_1}{T_2 - T_1}.$$

The drift term is

$$(81) \quad \gamma(s, u) = -\frac{1}{2} \sum_{k=1}^2 \sigma_k^2(s, u) - \lambda \int_{\mathbb{R}} (e^{\eta(s, u)y} - 1 - \eta(s, u)y) P^Y(dy).$$

It remains to specify the discretization of the delivery period  $[T_1, T_2]$ . Since we have a continuous forward curve model, we may use an arbitrary number of discretization points. On energy markets, monthly swaps on electricity are usually based on daily prices. Consequently, we will use exactly  $n = 28$  components. The forward curve process is defined by

$$(82) \quad f(t, u_i) = \exp[X(t, u_i)], \quad t \in [0, T],$$

where  $u_i = T_1 + (i - 1) \frac{T_2 - T_1}{28}$ ,  $i = 1, \dots, 28$ . Since an electricity swap requires no payment before the delivery period starts, it is a martingale under the risk neutral measure. Therefore, the drift  $\gamma$  defined in (81) was chosen such that  $f(t, u_i)$ ,  $i = 1, \dots, 28$ , are martingales (compare (14)). We will further assume that there are 8 delivery hours per day. This implies

$$(83) \quad \widehat{V}(t) = e^{-rT} \cdot 8 \cdot 28 \cdot E \left[ \left( \sum_{i=1}^{28} w(u_i; T_1, T_2) f(T, u_i) - K \right)^+ \middle| \mathcal{F}_t \right]$$

for the price of the option discounted to time 0, without any discretization error. The risk free interest rate is assumed to be constant at  $r = 0.02$ . The initial forward curve at time  $t = 0$  is  $f_0(u_i) \equiv 50$ ,  $i = 1, \dots, 28$ , the strike is  $K = 50$ .

In order to make use of the easily parallelized methods (MC simulations as well as PIDE solver), the experiments were run on a Linux workstation with 6 Opteron processors at 2.7 GHz. The number of POD components was set to  $d = 2$ , which is sufficient for precise calculation of the option prices in this test setting according to [10]. Models containing more driving factors need more POD components to achieve similar accuracy.

Since we are interested in the residual risk of the hedging strategy rather than the discretization error of the PIDE solver, a sufficiently small mesh width was chosen. With at most  $2^{11}$  grid points per coordinate (this number varies due to the application of a sparse grid combination method), the discretization error is negligible compared to the residual hedging error. The optimal hedging strategy presented in Section 3 is based on continuous rebalancing. Hence the time steps were also chosen to be very small ( $\Delta t = \frac{365}{3000}$ ). Because of the fine discretization, the solution of the time dependent PIDE takes up to 40 minutes. Note that if one is only interested in the hedging strategy at a finite number of time steps (and possibly also on a less fine spatial grid), computations can be sped up considerably. The current hedging portfolio at time  $t = 0$  can in fact be obtained from the solution of the PIDE with time independent coefficients which is also used for pricing. This can be done within a few seconds.

## 4.2 Results

The monthly option described above can be hedged with any set of traded swaps. We start with a single swap with a delivery period covering the whole month (28 days). Then we add further swaps whose delivery periods equal the weeks of the month. To simplify the notation, the corresponding strategies are denoted by capital letters as

shown in Table 1. Finally, we add the index “n-n” to the letter, if we consider only nonnegative strategies. Thus, e.g.,  $B_{n-n}$  denotes the nonnegative strategy using the monthly and the first weekly swap.

	delivery periods of included swaps				
	month	week 1	week 2	week 3	week 4
<i>A</i>	x				
<i>B</i>	x	x			
<i>C</i>	x	x	x		
<i>D</i>	x	x	x	x	
<i>E</i>		x	x	x	x
<i>F</i>	x	x	x	x	x

Table 1: Hedging portfolios used for numerical experiments.

The corresponding average hedging errors are computed from 10000 MC paths for each portfolio. The Tikhonov regularization parameter described in Section 3.4 is set to  $\delta = 10^{-10}$ . The results are shown in Table 2. These errors equal the unhedgeable part of the risk, which is caused by two effects: the jumps in the model (which render even a one-dimensional market incomplete) and the fact that we use only a finite number of assets to hedge the infinite-dimensional forward curve. As was to be expected, the error decreases substantially, if we add more swaps to the portfolio. A set of 4 swaps, however, is sufficient to obtain a very small error. This is of course due to the rather low number of driving factors in the test model. We also compare the terminal error with the initial option value for our test problem. (A relative error comparing with the option value at time  $T$  is ill-defined, since the option might be out of the money, yielding a division by zero.) Apparently, the cost for the additional nonnegativity constraint on the portfolio is high for portfolios which contain several swaps.

portfolio	<i>A</i>	$A_{n-n}$	<i>B</i>	$B_{n-n}$	<i>C</i>	$C_{n-n}$
absolute error	215.81	217.58	173.01	178.75	30.90	175.74
error / initial value	0.132	0.133	0.106	0.109	0.019	0.108

portfolio	<i>D</i>	$D_{n-n}$	<i>E</i>	$E_{n-n}$	<i>F</i>	$F_{n-n}$
absolute error	18.86	176.70	18.77	160.47	18.87	161.88
error / initial value	0.012	0.108	0.011	0.098	0.012	0.099

Table 2: Average hedging errors for different portfolios. The bottom rows show the ratio of the terminal hedging error at time  $T$  and the option value (1633.64) at time 0.

We now have a look at the influence of the regularization parameter  $\delta$ . Table 3 shows the average hedging error with portfolio *E* for different values of  $\delta$ . The error decreases

monotonously with decreasing  $\varepsilon$ , until the regularizing effect is no longer sufficient and the error increases again at  $\delta = 10^{-15}$  (which is close to machine precision). The error when using the projection method described in Section 3.4 is 18.25, which is about the same size as the best result achievable with Tikhonov regularization. We will continue to use regularization, however, since it is the more general concept, which also works for very bad conditioned matrices  $M$ . We set  $\delta = 10^{-10}$  for the remaining experiments.

$\delta$	$10^{-6}$	$10^{-7}$	$10^{-8}$	$10^{-9}$	$10^{-10}$	$10^{-11}$	$10^{-12}$	$10^{-13}$	$10^{-14}$	$10^{-15}$
error	30.24	28.14	21.36	18.77	18.78	18.71	18.56	18.47	18.38	18.78

Table 3: Average hedging errors for different regularization parameters in portfolio  $E$ .

We will now examine the hedging strategies in more detail. In order to illustrate some of their features, we pick two sample paths from the MC simulation: one where the option is in the money (ITM) at maturity and one where it ends up out of the money (OTM). We first analyze the ITM case. The hedging strategy for portfolio  $A$  is shown in Figure 2, together with the swap rate of the underlying monthly contract. This case is similar to that of hedging a single stock in that the quantity of swaps held in the portfolio is a number in  $[0, 1]$ . The investment jumps simultaneously with the swap rate and in the same directions. There is, e.g., a jump of the swap rate at day 243. Since the option changes from OTM to ITM with this jump, this has a large impact on the optimal investment. The strategy approaches 1 when time goes to  $T$ , since the option ends in the money.

Figure 3 displays the strategy for portfolio  $E$ . The optimal investments in the weekly swaps are now real numbers, sometimes negative and of large absolute value. When approaching maturity, however, the strategy satisfies  $|\theta_i| < 1, i = 1, \dots, 5$ .

The chronological development of the hedging errors corresponding to different portfolios in the ITM case is shown in Figure 4. When adding more swaps to the portfolio, two effects can be seen. The error gets smaller, and it evolves more smoothly. Both the diffusion and the jump part of the error decrease for more diverse portfolios. While the value of the hedge portfolio is visibly different from the true option value for portfolio  $A$ , these two curves are almost indistinguishable for portfolio  $E$ . One can also observe the Samuelson effect of increasing volatility in the option price closer to maturity. The volatility of the hedging error, however, is very small during the last month before maturity. This is due to the fact that after a significant jump at day 318 the swap rate is far above the strike. The option will thus very likely end in the money. The hedging investment is then almost equal to 1 and no further hedging error is accumulated.

The strategies in the OTM case approach 0 close to maturity. This holds for the investment in the monthly swap as well as any weekly swaps. Once again, the cor-

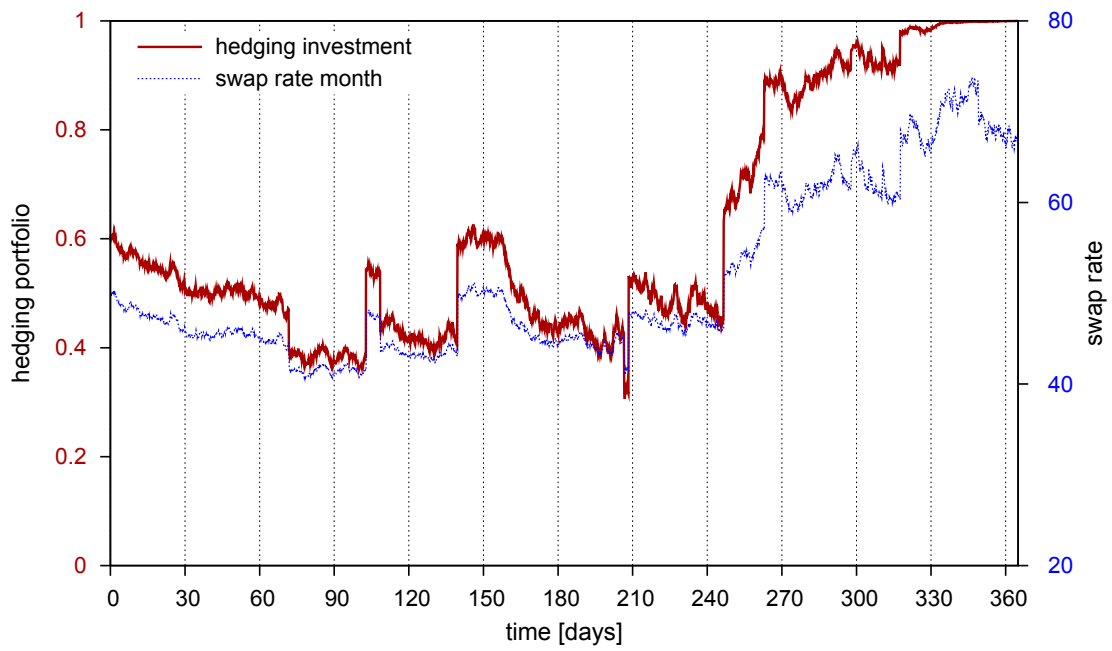


Fig. 2: Swap rate and hedging strategy for portfolio *A* when the option ends ITM.

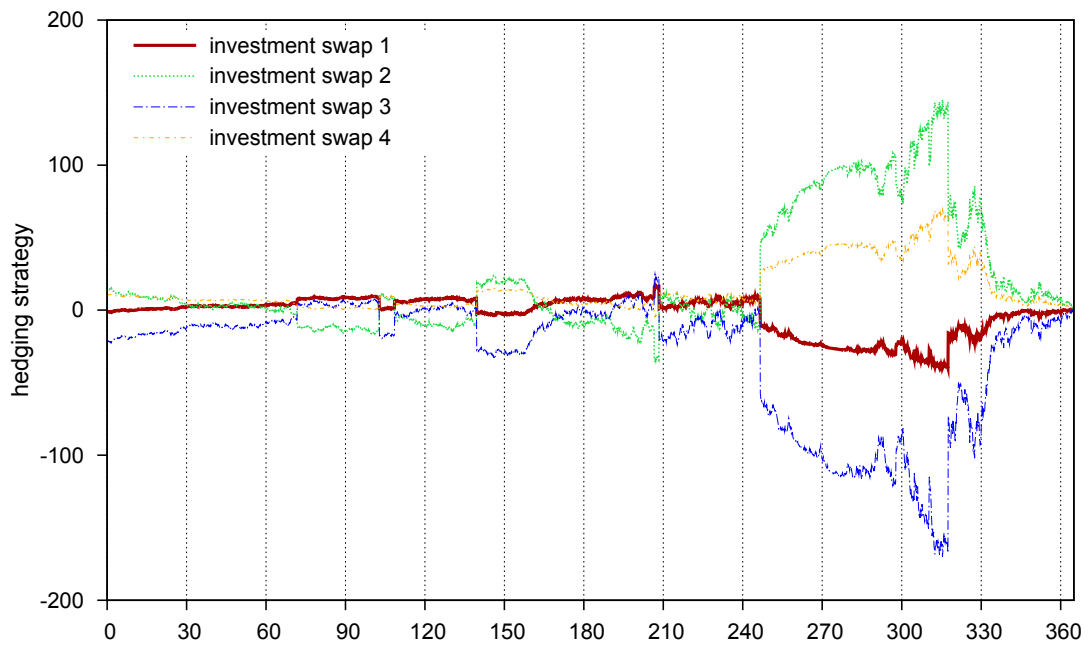


Fig. 3: Hedging strategy for portfolio *E* when the option ends ITM.

relation between hedging strategy  $\theta_0$  and swap rate is clearly visible, as shown in Figures 5 and 6. The terminal value of portfolio  $A$  is negative and thus slightly below the targeted value 0 (Figure 7). The behavior of the hedging errors is very similar to the ITM case. In particular, the error gets smaller and smoother when more swaps are included and is almost constant 0 for portfolio  $E$ .

## 5 Conclusion

This article is concerned with the numerical computation of hedging strategies for European electricity options. Trading in a portfolio of electricity contracts with various delivery periods, we look for an optimal mean-variance strategy. We consider an infinite-dimensional forward curve model, driven by an exponential time-inhomogeneous jump-diffusion process. Similar to a classical delta hedge, the strategy depends on partial derivatives of the option price, which can be obtained from the solution of a partial integro-differential equation (PIDE). Using a dimension reduction method, the infinite-dimensional problem can be approximated with finite-dimensional, possibly semi-definite, quadratic optimization problems. Tikhonov regularization is used to overcome numerical issues corresponding to the semi-definiteness. Convergence of the hedging error corresponding to the approximating strategy to the error of an optimal strategy is shown.

Numerical experiments are performed to analyze the resulting hedging strategies. The unhedgeable part of the risk, which is due to the inherent incompleteness of the market, is evaluated by Monte Carlo simulations. The hedging strategy along each simulated path is computed from a finite-dimensional PIDE solution.

## Acknowledgments

I would like to thank Fred Espen Benth and Claudia Klüppelberg for helpful comments and critical reading of the manuscript.

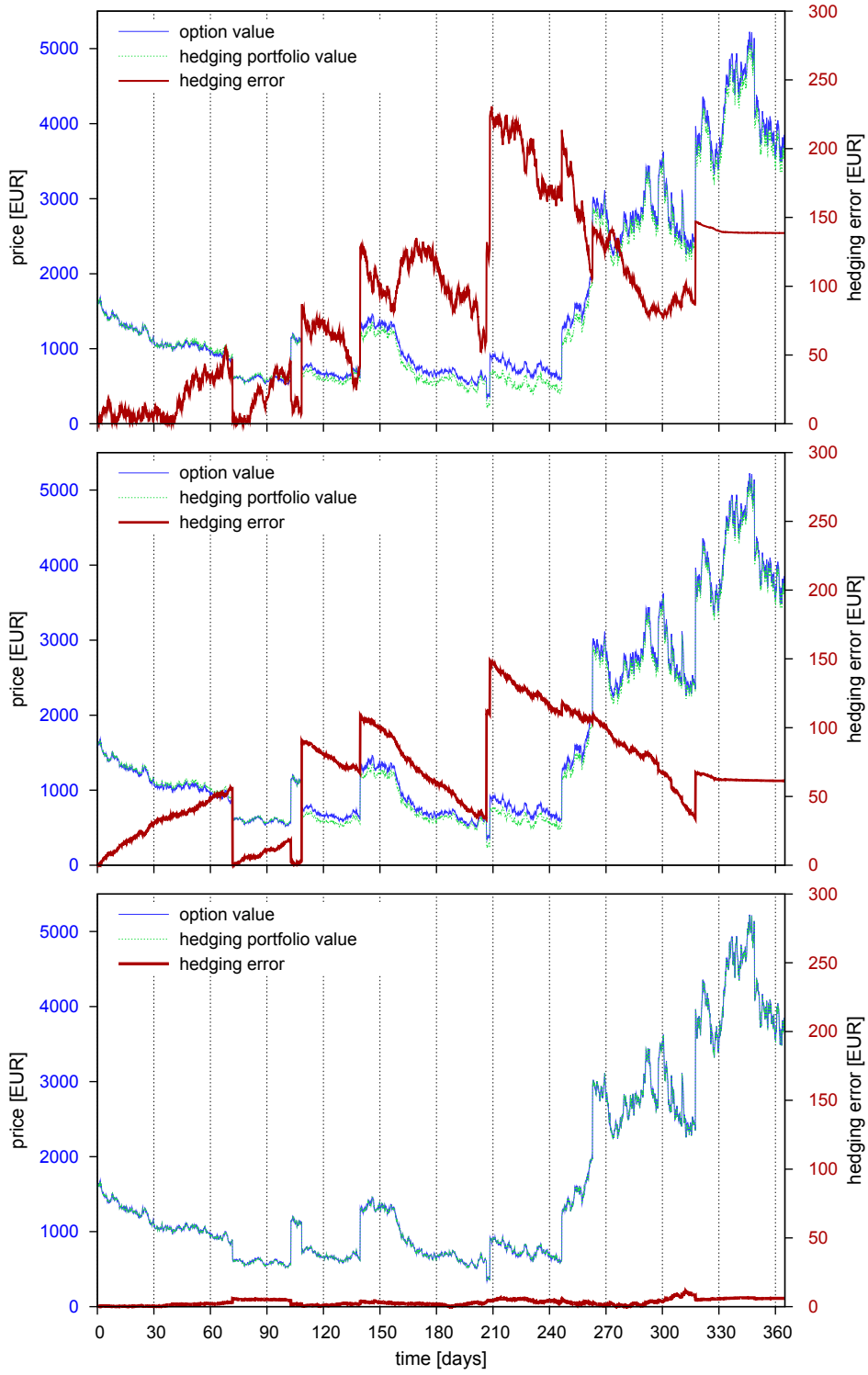


Fig. 4: Option value, hedging portfolio value, and absolute hedging error for portfolios A, B, and E (top to bottom) when the option ends ITM.

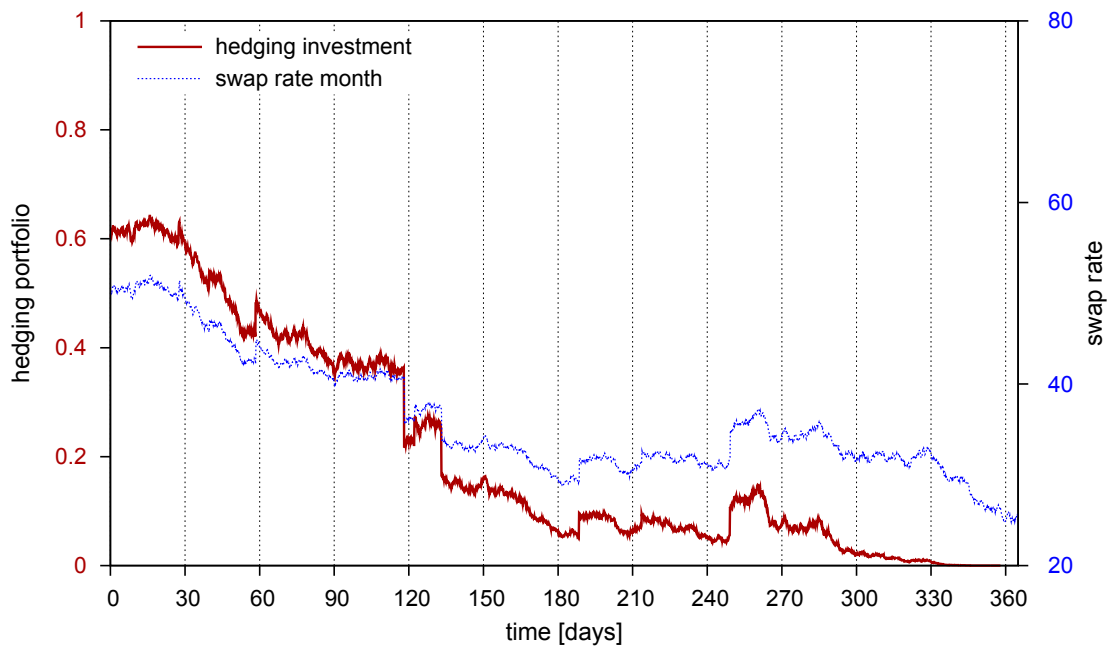


Fig. 5: Swap rate and hedging strategy for portfolio *A* when the option ends OTM.

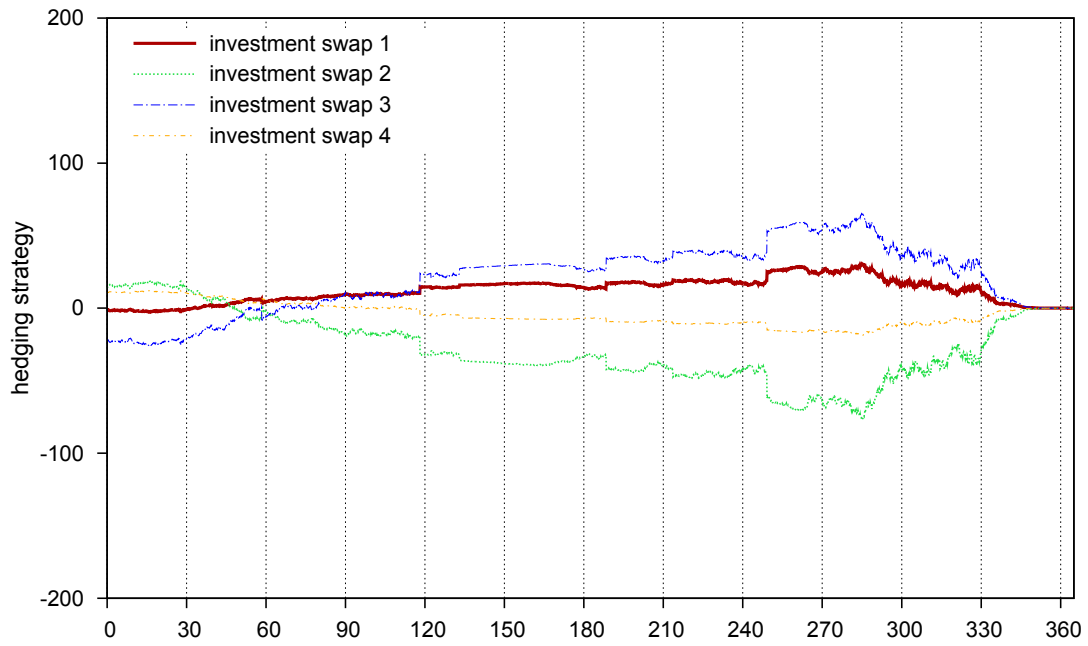


Fig. 6: Hedging strategy for portfolio *E* when the option ends OTM.



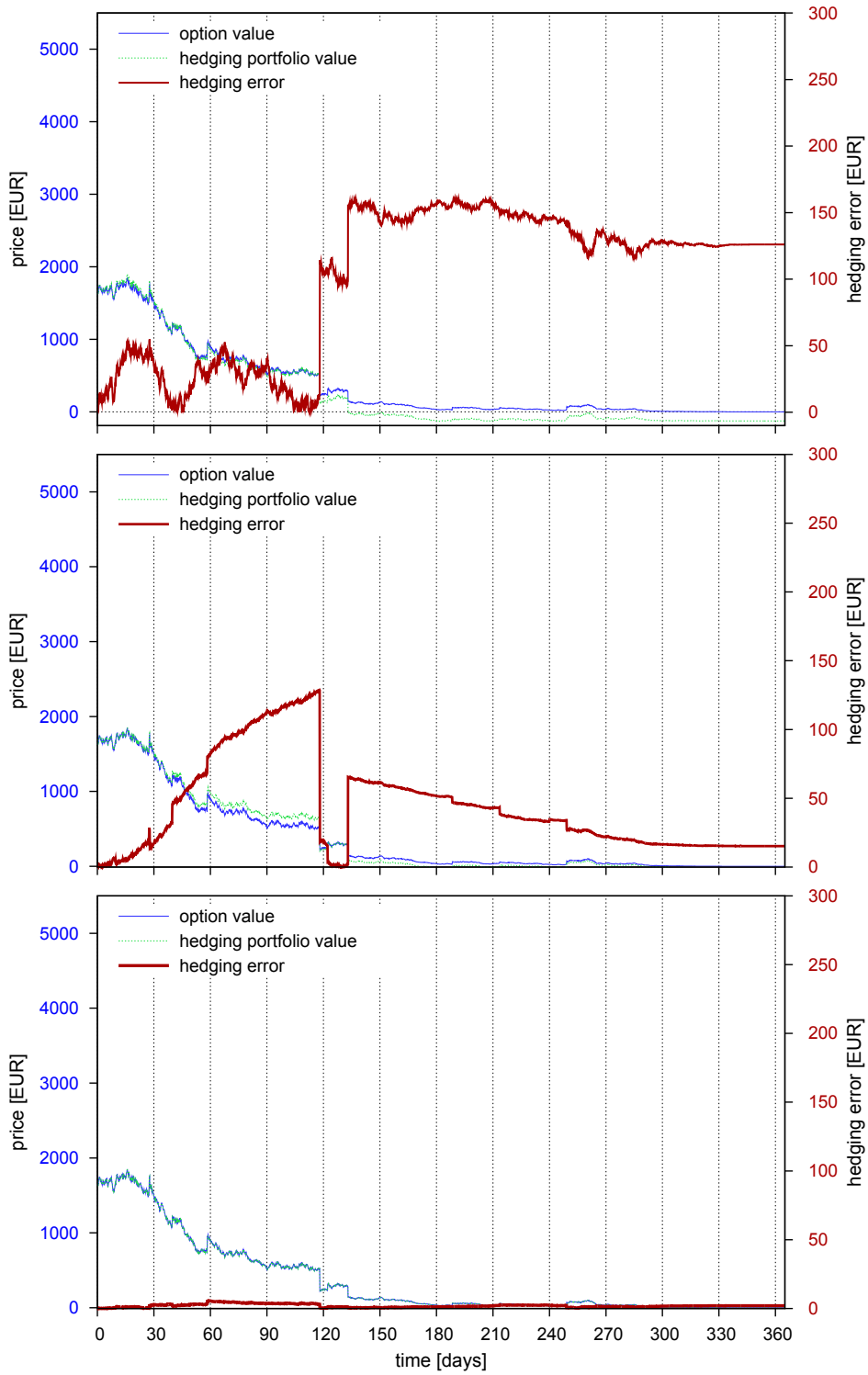


Fig. 7: Option value, hedging portfolio value, and absolute hedging error for portfolio A, B, and E (top to bottom) when the option ends OTM.

## References

- [1] F. E. Benth, S. Koekebakker, and F. Ollmar. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *Journal of Derivatives*, 52(15):62–66, 2007.
- [2] F. E. Benth, J. Sältytė Benth, and S. Koekebakker. *Stochastic Modeling of Electricity and Related Markets*. World Scientific, Singapore, 2008.
- [3] R. Carmona and M. Tehranchi. *Interest Rate Models: An Infinite-Dimensional Stochastic Analysis Perspective*. Springer, Berlin, 2006.
- [4] R. Cont. Modeling term structure dynamics: an infinite dimensional approach. *Int J. Theor. and Appl. Finance*.
- [5] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, Cambridge, 1992.
- [6] J. Dieudonné. *Foundations of Modern Analysis*. Academic Press, New York, 1960.
- [7] E. Eberlein and W. Kluge. Exact pricing fomulae for caps and swaptions in a Lévy term structure model. *Journal of Computational Finance*, (9):99–125, 2006.
- [8] C. Geiger and C. Kanzow. *Theorie und Numerik Restringierter Optimierungsaufgaben*. Springer, Berlin, 2002.
- [9] P. Hepperger. Hedging electricity swaptions using partial integro-differential equations. *Submitted for publication, preprint on <http://www-m4.ma.tum.de>*, 2010.
- [10] P. Hepperger. Option pricing in Hilbert space valued jump-diffusion models using partial integro-differential equations. *SIAM Journal on Financial Mathematics*, 1:454–489, 2010.
- [11] R. Kiesel, G. Schindlmayr, and R. Börger. A two-factor model for the electricity forward market. *Quant. Finance*, 9(3):279–287, 2009.
- [12] T. Kluge. *Pricing Swing Options and other Electricity Derivatives*. PhD thesis, University of Oxford, 2006.
- [13] H. Kunita. Stochastic integrals based on martingales taking values in Hilbert space. *Nagoya Math. J.*, 38:41–52, 1970.
- [14] S. Peszat and J. Zabczyk. *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach*. Cambridge University Press, Cambridge, 2007.
- [15] C. Schwab and R. A. Todor. Karhunen-Loève approximation of random fields by generalized fast multipole methods. *J. Comput. Phys.*, 217:100–122, 2006.