

# A SEQUENTIAL QUADRATIC PROGRAMMING METHOD FOR VOLATILITY ESTIMATION IN OPTION PRICING

B. DÜRING, A. JÜNGEL, AND S. VOLKWEIN

**ABSTRACT.** Our goal is to identify the volatility function in Dupire’s equation from given option prices. Following an optimal control approach in a Lagrangian framework, we propose a globalized sequential quadratic programming (SQP) algorithm with a modified Hessian – to ensure that every SQP step is a descent direction – and implement a line search strategy. In each level of the SQP method a linear–quadratic optimal control problem with box constraints is solved by a primal–dual active set strategy. This guarantees  $L^\infty$  constraints for the volatility, in particular assuring its positivity. The proposed algorithm is founded on a thorough first– and second–order optimality analysis. We prove the existence of local optimal solutions and of a Lagrange multiplier associated with the inequality constraints. Furthermore, we prove a sufficient second-order optimality condition and present some numerical results underlining the good properties of the numerical scheme.

## 1. INTRODUCTION

Financial derivatives, in particular options, became very popular financial contracts in the last few decades. Options can be used, for instance, to hedge assets and portfolios in order to control the risk due to movements in the share price. We recall that a European Call (Put) option provides the right to buy (to sell) a fixed number of assets at the fixed exercise price  $E$  at the expiry time  $T$ ; see, e.g. [16].

In an idealized financial market the price of a European option  $V(t, S)$  on an underlying asset  $S$  at time  $t$  can be obtained as the solution of the celebrated Black-Scholes equation (see, e.g., [8, 41])

$$V_t(t, S) + \frac{\sigma^2}{2} S^2 V_{SS}(t, S) + rSV_S(t, S) - rV(t, S) = 0, \quad t \in (0, T), \quad S > 0, \quad (1.1a)$$

where  $r > 0$  is the riskless interest rate and  $T > 0$  the time of maturity, with the final condition

$$V(T, S) = P(S), \quad S > 0, \quad (1.1b)$$

with given pay-off  $P(S)$  and appropriate boundary conditions.

---

*Date:* March 29, 2006.

*2000 Mathematics Subject Classification.* 35Kxx, 49J20, 49K20, 90C55.

*Key words and phrases.* Dupire equation, parameter identification, optimal control, optimality conditions, SQP method, primal-dual active set strategy.

The first and second author acknowledge partial support from the Deutsche Forschungsgemeinschaft, grant JU 359/6 (Forschergruppe 518). The first author was supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung under the Special Research Center F003 “Optimization and Control”.

The Black–Scholes equation has been derived under several assumptions, in particular the asset price  $S(t)$  is supposed to follow a stochastic process

$$dS(t) = \mu S(t) dt + \sigma S(t) dW(t),$$

where  $\mu \in \mathbb{R}$ ,  $\sigma > 0$  are the constant drift and constant volatility of the underlying asset, respectively, and  $W(t)$  denotes a Brownian motion. The drift and the volatility are not directly observable. The drift is removed from the model by a hedging argument [16] and does not enter explicitly in the Black–Scholes equation. Obtaining values for  $\sigma$  is often done by computing the so-called *implied volatility* out of observed option prices by inverting the closed-form solution to (1.1), the so-called Black–Scholes formula. A widely observed phenomenon is that these computed volatilities are not constant.

The pattern of implied volatilities for different exercise prices sometimes forms a *smile* shape, i.e., implied volatilities of in-the-money and out-of-the-money options are generally higher than that of at-the-money options. This is observed, for example, in coffee option markets. In equity option markets, typically, one observes a so-called volatility *skew*, i.e. the implied volatility for in-the-money calls is significantly higher than the implied volatility of at-the-money calls and out-of-the-money calls. Additionally, often variation with respect to time to maturity is present as well. This is usually referred to as the *volatility term structure*.

These observations lead to a natural generalization of the Black–Scholes model replacing the constant volatility  $\sigma$  in the model by a (deterministic) *local volatility function*  $\sigma = \sigma(T, E)$ , where  $T$  denotes the time to maturity and  $E$  the exercise price. It arises the question of how to determine this volatility function from option prices observed in markets, such that the generalized Black–Scholes model replicates the market prices. This problem is often referred to as the *calibration problem*.

As first observed by Dupire [18], the option price  $V = V(T, E)$  as a function of the exercise time  $T$  and the exercise price  $E$  satisfies the (forward) differential equation

$$V_T(T, E) - \frac{1}{2}\sigma^2(T, E)E^2V_{EE}(T, E) + rEV_E(T, E) = 0, \quad T > 0, E > 0 \quad (1.2a)$$

with the initial condition

$$V(0, E) = V_0(E) = \max(S_0 - E, 0), \quad E > 0, \quad (1.2b)$$

and boundary conditions

$$V(T, 0) = S_0, \quad \lim_{E \rightarrow \infty} V(T, E) = 0, \quad T > 0. \quad (1.2c)$$

It is derived from a Fokker–Planck equation integrated twice with respect to the space variable  $E$  and using the (formal) identity  $(S_0 - E)_{EE}^+ = \delta_{S_0}(E)$ , where  $\delta_{S_0}$  denotes the Dirac mass at  $S_0$ . Solving (1.2a) for the volatility leads to *Dupire’s formula*

$$\sigma(E, T) = \left( \frac{2[V_T(T, E) + rEV_E(T, E)]}{E^2V_{EE}(T, E)} \right)^{1/2}. \quad (1.3)$$

Note that typical option prices are strictly convex in  $E$  which implies positivity of the denominator.

Dupire’s local volatility function model has received great attention as well as some criticism [17]. It was extended in [19, 36] by defining the local variance as the expectation of the future instantaneous variance conditional on a given asset price.

Therein, the (stochastic) instantaneous variance can be quite general, such that this approach is consistent with (univariate diffusion) stochastic volatility models, see for example [29]. However, if one stays within the completely deterministic setting, (1.2) is the most elaborate model up to our knowledge.

The problem of determining the volatility in (1.2) from observed option prices is an ill-posed optimization problem in the sense of the lack of continuous dependence of the minimizers with respect to perturbations of the problem. In the mathematical literature, there are two main approaches to address the *calibration problem*. The first is to apply equation (1.3) on interpolated data sets of option prices observed in the market [12, 18, 24]. This approach depends largely on the interpolation method but it is computationally cheap.

The second approach is to use a regularization technique. For instance, the problem is reformulated as a stochastic optimal control problem and a so-called entropic regularization [2] is performed or a Tikhonov regularized cost functional is used in a (deterministic) inverse problem [38]. The last approach has been adopted in many works, see, e.g., in [1, 32]. For a complete review of the literature we refer to [14], for a survey on Tikhonov regularization see [21].

Most of the references mentioned above focus on the numerical results obtained by standard methods without analyzing in-depth the employed algorithms. A theoretical foundation of the approach with Tikhonov regularized cost functional is given in [13, 14]. In [14] a trinomial tree method using Tikhonov regularization and a probabilistic interpretation of the cost function's gradient is analyzed and numerical results are shown. Convergence rates for Tikhonov regularization under interpretable conditions have been derived in [20].

Our goal is to identify from given option prices  $V(T, E)$  the volatility function  $\sigma$  in (1.2). We follow the optimal control approach using a Lagrangian framework. The proposed algorithm is based on a sequential quadratic programming method (SQP) and on a primal-dual active set strategy that guarantees pointwise bilateral constraints for the volatility, in particular assuring its positivity. The algorithm proposed is founded on a thorough analysis of first- and second-order optimality conditions. Furthermore, we prove the existence of a Lagrange multiplier associated with the inequality constraints.

SQP methods have been widely applied to optimization problems of the form

$$\text{minimize } J(x) \text{ subject to } e(x) = 0,$$

where the cost functional  $J : X \rightarrow \mathbb{R}$  and the constraint  $e : X \rightarrow Y$  are sufficiently smooth functions and  $X, Y$  are real Hilbert spaces. Such problems occur frequently in optimal control of systems described by partial differential equations [3]. SQP methods for constrained optimal control of partial differential equations have been studied widely. For a general survey on SQP methods we refer to [9], for instance, and the references therein.

The basic idea of SQP methods is to minimize at each iteration a quadratic approximation of the Lagrangian associated with the cost functional over an affine subspace of solutions of the linearized constraint. In each level of the SQP method a linear-quadratic subproblem has to be solved. In the presence of bilateral coefficient constraints, this subproblem involves linear inequality constraints. For the solution of the subproblems we use a primal-dual active set method based on a generalized Moreau-Yosida approximation of the indicator function of the admissible control set [6, 27].

This paper is organized in the following manner: In Section 2 we formulate the parameter estimation as an optimal control problem and prove the existence of local optimal solutions. Moreover, any optimal solution is characterized by an optimality system involving an adjoint equation for the Lagrange multiplier. The optimization method is proposed in Section 3. We apply a globalized SQP method with a modified Hessian matrix to ensure that every SQP step is a descent direction and implement a line search strategy. In each level of the SQP method a linear-quadratic optimal control problem with box constraints is solved by a primal-dual active set strategy. In Section 4 numerical examples are presented and discussed.

## 2. THE OPTIMAL CONTROL PROBLEM

In this section the parameter identification problem is introduced as an optimal control problem. We prove the existence of at least one optimal solution and present first-order necessary optimality conditions. Furthermore, we investigate sufficient second-order optimality conditions.

**2.1. Formulation of the optimal control problem.** We start by introducing some notation. For  $R > E > M > 0$  and  $T > 0$  let  $\Omega = (M, R)$  be the one-dimensional spatial domain and  $Q = (0, T) \times \Omega$  the time-spatial domain. Concerning the error inflicted by introducing artificial boundary conditions we refer to [4, 35].

We define the Hilbert space

$$V = \{\varphi \in H^1(\Omega) : \varphi(R) = 0\}$$

endowed with the inner product

$$\langle \varphi, \psi \rangle_V = \int_{\Omega} \varphi_x \psi_x \, dx \quad \text{for all } \varphi, \psi \in V.$$

By  $L^2(0, T; V)$  we denote the space of (equivalence classes) of measurable functions  $\varphi : [0, T] \rightarrow V$ , which are square integrable, i.e.,

$$\int_0^T \|\varphi(t)\|_V^2 \, dt < \infty.$$

Analogously, the spaces  $L^2(0, T; H^1(\Omega))$  and  $L^2(0, T; L^\infty(\Omega))$  are defined. In particular,  $L^2(0, T; L^2(\Omega))$  can be identified with  $L^2(Q)$ . Moreover we make use of the space

$$W(0, T) = \{\varphi \in L^2(0, T; V) : \varphi_t \in L^2(0, T; V')\},$$

which is a Hilbert space endowed with the common inner product; see [15, p. 473]. Let us recall the Hilbert space

$$\begin{aligned} H^{2,1}(Q) &= H^1(0, T; L^2(\Omega)) \cap L^2(0, T; H^2(\Omega)) \\ &= \{\varphi : Q \rightarrow \mathbb{R} \mid \varphi, \varphi_t, \varphi_x, \varphi_{xx} \in L^2(Q)\}, \end{aligned}$$

supplied with the inner product

$$\langle \varphi, \psi \rangle_{H^{2,1}(Q)} = \int_0^T \int_{\Omega} \varphi_t \psi_t + \varphi_{xx} \psi_{xx} + \varphi_x \psi_x + \varphi \psi \, dx dt \quad \text{for } \varphi, \psi \in H^{2,1}(Q)$$

and the induced norm  $\|\cdot\|_{H^{2,1}(Q)} = \langle \cdot, \cdot \rangle_{H^{2,1}(Q)}^{1/2}$ . Recall that from  $\Omega \subset \mathbb{R}$  it follows that  $H^{2,1}(Q)$  is continuously embedded into  $L^\infty(Q)$ ; see, e.g. [39, p. 24].

When  $t$  is fixed, the expression  $\varphi(t)$  stands for the function  $\varphi(t, \cdot)$  considered as a function in  $\Omega$  only.

Next we specify the set of admissible coefficient functions. Suppose that  $q_{\min}$  and  $q_{\max}$  are given functions in  $H^{2,1}(Q) \cap L^\infty(0, T; H^2(\Omega))$  satisfying  $\bar{q}_{\min} \leq q_{\min} < q_{\max}$  in  $Q$  almost everywhere (a.e.) with  $\bar{q}_{\min} = \text{essinf}\{q(t, x) : (t, x) \in Q\} > 0$ . In particular, there exists  $C_{\text{ad}} > 0$  such that

$$\max \{ \|q_{\min}\|_{L^\infty(0, T; H^2(\Omega))}, \|q_{\max}\|_{L^\infty(0, T; H^2(\Omega))} \} \leq C_{\text{ad}}.$$

We introduce the set for the admissible coefficient functions by

$$\mathcal{Q}_{\text{ad}} = \{q \in H^{2,1}(Q) : \|q\|_{L^\infty(0, T; H^2(\Omega))} \leq C_{\text{ad}}, q_{\min} \leq q \leq q_{\max} \text{ in } Q \text{ a.e.}\}, \quad (2.1)$$

which is a closed, bounded and convex set in  $H^{2,1}(Q)$ . Note, that the bound  $C_{\text{ad}} > 0$  is purely technical, and can be chosen arbitrarily large.

The goal of the parameter identification is to determine the volatility in (1.2a). For streamlining the presentation we restrict ourselves to the case  $r = 0$  of zero interest rate in the analytical part of the paper. Therefore, we need to determine the coefficient function  $q = q(t, x) = \frac{1}{2}E^2\sigma^2(T, E)$  in the parabolic problem

$$u_t(t, x) - q(t, x)u_{xx}(t, x) = 0 \quad \text{for all } (t, x) \in Q, \quad (2.2a)$$

$$u(t, M) = u_D(t) \quad \text{for all } t \in (0, T), \quad (2.2b)$$

$$u(t, R) = 0 \quad \text{for all } t \in (0, T), \quad (2.2c)$$

$$u(0, x) = u_0(x) \quad \text{for all } x \in \Omega \quad (2.2d)$$

from given, observed option data  $u_T \in L^2(\Omega)$  for the solution  $u$  of (2.2) at the final time  $T$ .

**Definition 2.1.** For given  $q \in \mathcal{Q}_{\text{ad}}$ ,  $u_D \in H^1(0, T)$  and  $u_0 \in L^2(\Omega)$  a function  $u$  is called a weak solution to (2.2) if  $u \in W(0, T)$ ,  $u(\cdot, M) = u_D$  in  $L^2(0, T)$ ,  $u(0) = u_0$  in  $L^2(\Omega)$  and

$$\int_0^T \langle u_t, \varphi \rangle_{H^{-1}, H_0^1} + \left( \int_\Omega q u_x \varphi_x + q_x u_x \varphi \, dx \right) dt = 0, \quad (2.3)$$

for all  $\varphi \in L^2(0, T; H_0^1(\Omega))$ . In (2.3)  $\langle \cdot, \cdot \rangle_{H^{-1}, H_0^1}$  denotes the duality pairing between  $H_0^1(\Omega)$  and its dual space  $H^{-1}(\Omega)$ .

**Remark 2.2.** Recall that  $H_0^1(\Omega) \hookrightarrow V$  and  $H_0^1(\Omega)$  is dense in  $V$ . Consequently,  $V' \hookrightarrow H^{-1}(\Omega)$  and  $u \in W(0, T) \subset H^1(0, T; H^{-1}(\Omega))$ . Furthermore,  $q \in H^{2,1}(Q) \hookrightarrow L^\infty(Q)$  and  $q_x \in H^{1,1}(Q) \hookrightarrow C([0, T]; L^2(\Omega))$ . Thus, the integral in (2.3) is well-defined for every  $\varphi \in L^2(0, T; H_0^1(\Omega))$ .

The following theorem ensures existence of a weak solution to (2.2) for positive coefficient functions. Its proof follows from standard arguments [37].

**Theorem 2.3.** Suppose that  $u_0 \in L^2(\Omega)$  and  $u_D \in H^1(0, T)$ . Then, for every  $q \in \mathcal{Q}_{\text{ad}}$  there exists a unique weak solution  $u$  to (2.2) and a constant  $C > 0$  such that

$$\|u\|_{W(0, T)} \leq C (\|u_0\|_{L^2(\Omega)} + \|u_D\|_{H^1(0, T)}). \quad (2.4)$$

If the initial condition  $u_0$  is more regular, we have the following corollary. Its proof is omitted, because it is standard.

**Corollary 2.4.** *If  $u_0 \in V$  holds with the compatibility condition  $u_0(M) = u_D(0)$ , it follows that  $u \in H^{2,1}(Q)$  and there exists a constant  $C > 0$  such that*

$$\|u\|_{H^{2,1}(Q)} \leq C (\|u_0\|_V + \|u_D\|_{H^1(0,T)}). \quad (2.5)$$

To write the state equations (2.2) in an abstract form we define the two Hilbert spaces

$$X = H^{2,1}(Q) \times W(0, T) \quad \text{and} \quad Y = L^2(0, T; H_0^1(\Omega)) \times L^2(0, T) \times L^2(\Omega)$$

endowed with their product topologies. Moreover, let

$$K_{\text{ad}} = \mathcal{Q}_{\text{ad}} \times W(0, T)$$

which is closed and convex. In the sequel we identify the dual  $Y'$  of  $Y$  with the product space  $L^2(0, T; H^{-1}(\Omega)) \times L^2(0, T) \times L^2(\Omega)$ .

Next we introduce the bilinear operator  $e = (e_1, e_2, e_3) : X \rightarrow Y'$  by

$$e_1(\omega) = u_t - qu_{xx}, \quad (2.6a)$$

$$e_2(\omega) = u(\cdot, M) - u_D, \quad (2.6b)$$

$$e_3(\omega) = u(0) - u_0, \quad (2.6c)$$

where  $\omega = (q, u)$  holds and the identity  $e_1(\omega) = u_t - qu_{xx}$  in  $L^2(0, T; H^{-1}(\Omega))$  stands for

$$\begin{aligned} & \langle e_1(\omega), \varphi \rangle_{L^2(0,T;H^{-1}(\Omega)), L^2(0,T;H_0^1(\Omega))} \\ &= \int_0^T \langle u_t, \varphi \rangle_{H^{-1}, H_0^1} dt + \int_0^T \int_{\Omega} qu_x \varphi_x + q_x u_x \varphi \, dx dt \quad \text{for } \varphi \in L^2(0, T; H_0^1(\Omega)). \end{aligned}$$

**Remark 2.5.** From  $q \in H^{2,1}(Q)$  we infer that  $q_x \in C([0, T]; L^2(\Omega))$ . Thus, for  $\varphi \in L^2(0, T; H_0^1(\Omega))$

$$\begin{aligned} \left| \int_0^T \int_{\Omega} qu_x \varphi_x + q_x u_x \varphi \, dx dt \right| &\leq \|q\|_{L^\infty(Q)} \|u_x\|_{L^2(Q)} \|\varphi_x\|_{L^2(Q)} \\ &\quad + \|q_x\|_{C([0,T];L^2(\Omega))} \|u_x\|_{L^2(Q)} \|\varphi\|_{L^2(0,T;L^\infty(\Omega))}. \end{aligned}$$

It follows that the bilinear operator  $e_1$  is well-defined for every  $\omega \in X$ .

Now we address the properties of the operator  $e$ . In particular, we prove that  $e$  is Fréchet differentiable and its linearization  $e'(\omega)$  is surjective at any point  $\omega \in K_{\text{ad}}$ . The latter condition guarantees a constraint qualification, so that there exists a (unique) Lagrange multiplier  $\lambda^*$  satisfying the first-order necessary optimality condition (see Theorem 2.10). The Fréchet derivatives with respect to  $\omega$  are denoted by primes, where subscripts denote as usual the associated partial Fréchet derivative.

**Proposition 2.6.** *The bilinear operator  $e : X \rightarrow Y'$  is twice continuously Fréchet differentiable and the mapping  $\omega \mapsto e''(\omega)$  is Lipschitz continuous on  $X$ . Moreover, its linearization  $e'(\omega) : X \rightarrow Y'$  at any point  $\omega = (q, u) \in K_{\text{ad}}$  is surjective. Furthermore, we have*

$$\|\delta u\|_{W(0,T)} \leq C_1 \|\delta q\|_{H^{2,1}(Q)} \quad \text{for all } \delta \omega = (\delta q, \delta u) \in N(e'(\omega)), \quad (2.7)$$

where  $N(e'(\omega)) \subset X$  denotes the null space of  $e'(\omega)$ .

*Proof.* First we prove that  $e$  is twice continuously Fréchet differentiable at any point  $\omega = (q, u) \in K_{\text{ad}}$ . For arbitrary directions  $\delta\omega = (\delta q, \delta u)$ ,  $\widetilde{\delta\omega} = (\widetilde{\delta q}, \widetilde{\delta u}) \in X$  we compute the directional derivatives as

$$e'(\omega)\delta\omega = \begin{pmatrix} \delta u_t - q\delta u_{xx} - \delta q u_{xx} \\ \delta u(\cdot, M) \\ \delta u(0) \end{pmatrix} \quad (2.8)$$

and

$$e''(\omega)(\delta\omega, \widetilde{\delta\omega}) = \begin{pmatrix} -\widetilde{\delta q}\delta u_{xx} - \delta q\widetilde{\delta u}_{xx} \\ 0 \\ 0 \end{pmatrix}. \quad (2.9)$$

These equalities hold in the  $Y'$ -sense. Using Young's inequality we obtain

$$\begin{aligned} & \int_0^T \int_{\Omega} \delta q \delta u_x \varphi_x + \delta q_x \delta u_x \varphi \, dx dt \\ & \leq \left( \|\delta q\|_{L^\infty(Q)} \|\delta u_x\|_{L^2(Q)} + \|\delta q_x\|_{L^\infty(0,T;L^2(\Omega))} \|\delta u_x\|_{L^2(Q)} \right) \|\varphi\|_{L^2(0,T;H_0^1(\Omega))} \\ & \leq C \left( \|\delta q\|_{H^{2,1}(Q)}^2 + \|\delta u\|_{W(0,T)}^2 \right) \|\varphi\|_{L^2(0,T;H_0^1(\Omega))} \leq C \|\delta\omega\|_X^2 \|\varphi\|_{L^2(0,T;H_0^1(\Omega))} \end{aligned}$$

for all  $\varphi \in L^2(0,T;H_0^1(\Omega))$  and  $\delta\omega = (\delta q, \delta u) \in X$ . Hence,

$$\begin{aligned} & \|e_1(\omega + \delta\omega) - e_1(\omega) - e'_1(\omega)(\delta q, \delta u)\|_{L^2(0,T;H^{-1}(\Omega))} \\ & = \sup_{\|\varphi\|_{L^2(0,T;H_0^1(\Omega))}=1} \int_0^T \int_{\Omega} \delta q \delta u_x \varphi_x + \delta q_x \delta u_x \varphi \, dx dt \leq C \|\delta\omega\|_X^2 \end{aligned}$$

and thus

$$\lim_{\|\delta\omega\|_X \searrow 0} \frac{\|e_1(\omega + \delta\omega) - e_1(\omega) - e'_1(\omega)\delta\omega\|_{L^2(0,T;H^{-1}(\Omega))}}{\|\delta\omega\|_X} = 0. \quad (2.10)$$

Notice that — due to the linearity of the operators  $e_2$  and  $e_3$  — we have

$$\|e_2(\omega + \delta\omega) - e_2(\omega) - e'_2(\omega)\delta\omega\|_{L^2(0,T)} = 0 \quad (2.11)$$

and

$$\|e_3(\omega + \delta\omega) - e_3(\omega) - e'_3(\omega)\delta\omega\|_{L^2(\Omega)} = 0. \quad (2.12)$$

Consequently, we infer from (2.10), (2.11) and (2.12) that the operator  $e$  is Fréchet differentiable with Fréchet derivative (2.8). Now we turn to the second derivative. In view of (2.9)

$$\|e'_1(\omega + \widetilde{\delta\omega})\delta\omega - e'_1(\omega)\delta\omega - e''_1(\omega)(\delta\omega, \widetilde{\delta\omega})\|_{L^2(0,T;H^{-1}(\Omega))} = 0,$$

and  $e''_2(\omega) = e''_3(\omega) = 0$  holds. Hence, we infer that  $e$  is twice Fréchet differentiable and the directional derivative, given in (2.9), is the second Fréchet derivative of  $e$ . Since  $e''(\omega)$  does not depend on  $\omega \in X$ , the Lipschitz-continuity on  $X$  is obvious.

It remains to prove that  $e'(\omega)$  is surjective and that the estimate (2.7) is satisfied for all  $\delta\omega \in N(e'(\omega))$ . Suppose that  $r = (r_1, r_2, r_3) \in Y'$  is arbitrary. Then the

operator  $e'(\omega)$  is surjective, if there exists a pair  $\delta\omega = (\delta q, \delta u) \in X$  such that  $e'(\omega)\delta\omega = r$ , which is equivalent to

$$\delta u_t - q\delta u_{xx} = r_1 + \delta q u_{xx} \quad \text{in } L^2(0, T; H^{-1}(\Omega)), \quad (2.13a)$$

$$\delta u(\cdot, M) = r_2 \quad \text{in } L^2(0, T), \quad (2.13b)$$

$$\delta u(0) = r_3 \quad \text{in } L^2(\Omega). \quad (2.13c)$$

Choosing  $\delta q = 0$  there exists a unique  $\delta u \in W(0, T)$ , which solves (2.13). Hence  $e'(\omega)$  is surjective.

Let  $\delta\omega = (\delta q, \delta u) \in N(e'(\omega))$ . Estimate (2.7) follows from standard arguments. For that reason we only estimate the additional right-hand side in (2.13a), namely the term  $\delta q u_{xx}$ . We infer from Hölder's and Young's inequalities

$$\begin{aligned} \int_0^t \int_{\Omega} (\delta q \delta u)_x u_x \, dx \, ds &\leq \int_0^t \|\delta q\|_{L^\infty(\Omega)} \|\delta u_x\|_{L^2(\Omega)} \|u_x\|_{L^2(\Omega)} \, ds \\ &\quad + \int_0^t \|\delta q_x\|_{L^2(\Omega)} \|\delta u\|_{L^\infty(\Omega)} \|u_x\|_{L^2(\Omega)} \, ds \\ &\leq C(\varepsilon) \|\delta q\|_{H^{2,1}(Q)}^2 + \varepsilon \|\delta u\|_{W(0,T)}^2 \end{aligned}$$

for almost all  $t \in [0, T]$  and for every  $\varepsilon > 0$ , where the constant  $C(\varepsilon) > 0$  depends on  $\|u\|_{L^2(0,T;V)}$  and  $\varepsilon$ . Choosing  $\varepsilon$  appropriately and using standard arguments the estimate follows.  $\square$

**Remark 2.7.** It follows from the proof of Proposition 2.6 that at any point  $\omega \in K_{\text{ad}}$  the operator  $e_u : W(0, T) \rightarrow Y'$  is even bijective.

Next we introduce the cost functional  $J : X \rightarrow [0, \infty)$  by

$$J(\omega) = \frac{1}{2} \int_{\Omega} |u(T) - u_T|^2 \, dx + \frac{\beta}{2} \|q\|_{H^{2,1}(Q)}^2 \quad \text{for } \omega = (q, u) \in X, \quad (2.14)$$

where  $u_T$  is a given observed option price at the end-time  $T$ , and  $\beta > 0$  is a regularization parameter.

**Lemma 2.8.** *The cost functional  $J : X \rightarrow [0, \infty)$  is twice Fréchet differentiable and its Fréchet derivatives are given by*

$$J'(\omega)\delta\omega = \int_{\Omega} (u(T) - u_T)\delta u(T) \, dx + \beta \langle q, \delta q \rangle_{H^{2,1}(Q)} \quad (2.15)$$

and

$$J''(\omega)(\delta\omega, \widetilde{\delta\omega}) = \int_{\Omega} \delta u(T) \widetilde{\delta u}(T) \, dx + \beta \langle \delta q, \widetilde{\delta q} \rangle_{H^{2,1}(Q)} \quad (2.16)$$

for arbitrary directions  $\delta\omega = (\delta q, \delta u)$ ,  $\widetilde{\delta\omega} = (\widetilde{\delta q}, \widetilde{\delta u}) \in X$ . In particular, the mapping  $\omega \mapsto J''(\omega)$  is Lipschitz-continuous on  $X$ .

*Proof.* For all  $\delta u \in W(0, T)$  we have  $\delta u(T) \in L^2(\Omega)$  (see, e.g. [15, p. 480]) so that the integrals are well-defined. It follows by standard arguments that the first and second Fréchet derivative are given by (2.15) and (2.16), respectively. Since  $\omega \mapsto J''(\omega)$  does not depend on  $\omega$ , the mapping  $\omega \mapsto J''(\omega)$  is clearly Lipschitz-continuous on  $X$ .  $\square$



The parameter identification problem is given by a constrained optimal control problem in the following form

$$\min J(\omega) \quad \text{s.t.} \quad \omega \in K_{\text{ad}} \text{ and } e(\omega) = 0. \quad (\mathbf{P})$$

Note that in our formulation, both the state variable  $u$  and the coefficient  $q$  are considered as independent variables while the realization of (2.2) is an explicit constraint. Alternatively, one could use the equality constraint to treat  $u = u(q)$  as a variable depending on the unknown coefficient  $q$  and solve the nonlinear least-squares problem by the Gauss–Newton method.

In this paper, we choose the SQP approach with independent variables. SQP methods can be viewed as a natural extension of Newton methods, and are hence expected to inherit its fast local convergence property. Indeed, the iterates of the SQP method are identical to those generated by Newton’s method when applied to the system composed of the first-order necessary conditions for the Lagrangian associated with  $(\mathbf{P})$  and the equality constraint. Note that SQP methods are not feasible-point methods, i.e. its iterates need not be points satisfying the constraints.

**2.2. Existence of optimal solutions.** The next theorem guarantees that  $(\mathbf{P})$  possesses an optimal solution.

**Theorem 2.9.** *Problem  $(\mathbf{P})$  has at least one (global) solution  $\omega^* = (q^*, u^*) \in K_{\text{ad}}$ .*

*Proof.* In view of Theorem 2.3 the admissible set

$$E = \{\omega = (q, u) \in X : e(\omega) = 0 \text{ in } Y' \text{ and } \omega \in K_{\text{ad}}\} \quad (2.17)$$

is non-empty (from  $q_{\min} \in \mathcal{Q}_{\text{ad}}$  follows that  $(q_{\min}, u(q_{\min})) \in E$ ). Moreover,  $J(\omega) \geq 0$  holds for all  $\omega \in E$ . Thus there exists a  $\zeta \geq 0$  such that

$$\zeta = \inf\{J(\omega) : \omega \in E\}. \quad (2.18)$$

We infer that there exists a minimizing sequence  $(\omega^n)_{n \in \mathbb{N}} \subset E$ ,  $\omega^n = (q^n, u^n)$ , with

$$\lim_{n \rightarrow \infty} J(\omega^n) = \zeta.$$

Due to (2.4) and

$$J(\omega^n) \geq \frac{\beta}{2} \|q^n\|_{H^{2,1}(Q)}^2 \quad \text{for all } n,$$

we infer that the sequence  $(\omega^n)_{n \in \mathbb{N}}$  is bounded in  $X$ . Thus, there exist subsequences, again denoted by  $(\omega^n)_{n \in \mathbb{N}}$ , and a pair  $\omega^* = (q^*, u^*) \in X$  satisfying

$$\begin{aligned} q^n &\rightharpoonup q^* \text{ in } H^{2,1}(Q) \text{ as } n \rightarrow \infty, \\ u^n &\rightharpoonup u^* \text{ in } W(0, T) \text{ as } n \rightarrow \infty. \end{aligned} \quad (2.19)$$

Furthermore, since  $q^n \in \mathcal{Q}_{\text{ad}}$  and it holds  $L^\infty(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)) \hookrightarrow C([0, T]; H^1(\Omega))$  compactly due to Aubin’s lemma [43], we obtain

$$q^n \rightarrow q^* \text{ in } C([0, T]; H^1(\Omega)) \text{ as } n \rightarrow \infty. \quad (2.20)$$

In view of (2.19) and (2.20) it holds

$$\int_0^T \int_\Omega q^n u_x^n \varphi_x + q_x^n u_x^n \varphi \, dx dt \rightarrow \int_0^T \int_\Omega q^* u_x^* \varphi_x + q_x^* u_x^* \varphi \, dx dt$$

as  $n \rightarrow \infty$  for every  $\varphi \in L^2(0, T; H_0^1(\Omega))$ . Therefore,

$$\lim_{n \rightarrow \infty} e_1(\omega^n) = e_1(\omega^*) \text{ in } L^2(0, T; H^{-1}(\Omega)).$$

From  $e_1(\omega^n) = 0$  for all  $n \in \mathbb{N}$  we conclude that  $e_1(\omega^*) = 0$ . Since the operators  $e_2$  and  $e_3$  are linear, we find  $e(\omega^*) = 0$ . Since  $J$  is convex and continuous, and therefore weakly lower semi-continuous, we obtain  $J(\omega^*) \leq \lim_{n \rightarrow \infty} J(\omega^n) = \zeta$ . Finally, since  $\mathcal{Q}_{\text{ad}}$  is convex and closed in  $H^{2,1}(Q)$ , and therefore weakly closed, we have  $q^* \in \mathcal{Q}_{\text{ad}}$ , and the claim follows.  $\square$

**2.3. First-order necessary optimality conditions.** Problem **(P)** is a non-convex programming problem so that different local minima might occur. A numerical method will produce a local minimum close to its starting value. Hence, we do not restrict our investigations to global solutions of **(P)**. We will assume that a fixed reference solution  $\omega^* = (q^*, u^*) \in K_{\text{ad}}$  is given satisfying certain first- and second-order optimality conditions (ensuring local optimality of the solution).

In this section we introduce the Lagrange functional associated with **(P)** and derive first-order necessary optimality conditions. Furthermore, we show that there exists a unique Lagrange multiplier associated with the inequality constraints for the optimal coefficient  $q^*$ .

To formulate the optimality conditions we introduce the Lagrange functional  $L : X \times Y \rightarrow \mathbb{R}$  associated with **(P)** by

$$\begin{aligned} L(\omega, p) &= J(\omega) + \langle e(\omega), (\lambda, \mu, \nu) \rangle_{Y', Y} \\ &= \frac{1}{2} \|u(T) - u_T\|_{L^2(\Omega)}^2 + \frac{\beta}{2} \|q\|_{H^{2,1}(Q)}^2 + \int_{\Omega} (u(0) - u_0) \nu \, dx \\ &\quad + \int_0^T \langle u_t, \lambda \rangle_{H^{-1}, H_0^1} \, dt + \int_0^T \int_{\Omega} (q\lambda)_x u_x \, dx \, dt + \int_0^T (u(\cdot, M) - u_D) \mu \, dt, \end{aligned}$$

with  $\omega = (q, u) \in X$  and  $p = (\lambda, \mu, \nu) \in Y$ . Due to Proposition 2.6 and Lemma 2.8 the Lagrangian is twice continuously Fréchet differentiable with respect to  $\omega \in X$  for each fixed  $p \in Y$  and its second Fréchet derivative is Lipschitz-continuous.

An optimal solution to **(P)** can be characterized by first-order necessary optimality conditions. This is formulated in the next theorem. Recall that the set  $E$  has been introduced in (2.17). Moreover, let

$$B_{\rho}(\omega) = \{\tilde{\omega} \in X : \|\tilde{\omega} - \omega\|_X < \rho\}$$

be the open ball in  $X$  with radius  $\rho > 0$  and mid point  $\omega \in X$ .

**Theorem 2.10.** *Suppose that  $\omega^* = (q^*, u^*) \in K_{\text{ad}}$  is a local solution to **(P)**, i.e.,  $\omega^* \in E$  and there exists a constant  $\rho > 0$  such that*

$$J(\omega^*) \leq J(\omega) \quad \text{for all } \omega \in E \cap B_{\rho}(\omega^*).$$

*Then there is a unique Lagrange multiplier  $p^* = (\lambda^*, \mu^*, \nu^*) \in Y$  satisfying the adjoint equations*

$$-\lambda_t^* - (q^* \lambda^*)_{xx} = 0 \quad \text{in } Q, \quad (2.21a)$$

$$\lambda^*(\cdot, M) = \lambda^*(\cdot, R) = 0 \quad \text{in } (0, T), \quad (2.21b)$$

$$\lambda^*(T) = -(u^*(T) - u_T) \quad \text{in } \Omega \quad (2.21c)$$

*in the weak sense and the identities*

$$\mu^* = (q^* \lambda^*)_x(\cdot, M) \quad \text{in } L^2(0, T), \quad (2.22)$$

$$\nu^* = \lambda^*(0) \quad \text{in } L^2(\Omega) \quad (2.23)$$

hold. Moreover, the variational inequality

$$\langle \beta q^* - \mathcal{R}(\lambda^* u_{xx}^*), q - q^* \rangle_{H^{2,1}(Q)} \geq 0 \quad \text{for all } q \in \mathcal{Q}_{\text{ad}} \quad (2.24)$$

holds, where  $\mathcal{R} : (H^{2,1}(Q))' \rightarrow H^{2,1}(Q)$  denotes the Riesz isomorphism, i.e.,  $q = \mathcal{R}(f) \in H^{2,1}(Q)$  solves

$$\int_0^T \int_{\Omega} q_t \varphi_t + q_{xx} \varphi_{xx} + q_x \varphi_x + q \varphi \, dx dt = \langle f, \varphi \rangle_{(H^{2,1}(Q))', H^{2,1}(Q)} \quad \text{for all } \varphi \in H^{2,1}(Q)$$

with  $f \in (H^{2,1}(Q))'$ . Here,  $\langle \cdot, \cdot \rangle_{(H^{2,1}(Q))', H^{2,1}(Q)}$  denotes the duality pairing between  $H^{2,1}(Q)$  and its dual.

*Proof.* We infer from Proposition 2.6 and Remark 2.7 that a standard constraint qualification holds at  $(q^*, u^*)$  [40]. Therefore, there exists a unique Lagrange multiplier  $p^* = (\lambda^*, \mu^*, \nu^*) \in Y$  such that

$$L_q(\omega^*, p^*)(q - q^*) \geq 0 \quad \text{for all } q \in \mathcal{Q}_{\text{ad}}, \quad (2.25)$$

$$L_u(\omega^*, p^*)u = 0 \quad \text{for all } u \in W(0, T), \quad (2.26)$$

$$L_p(\omega^*, p^*)p = 0 \quad \text{for all } p \in Y. \quad (2.27)$$

Equation (2.27) is equivalent to the equality constraint  $e(\omega^*) = 0$  and is fulfilled since  $\omega^*$  solves **(P)**. Next we turn to (2.26), which is equivalent to

$$\begin{aligned} 0 &= \int_{\Omega} (u^*(T) - u_T)u(T) \, dx + \int_0^T \langle u_t, \lambda^* \rangle_{H^{-1}, H_0^1(\Omega)} \, dt \\ &\quad + \int_0^T \int_{\Omega} (q^* \lambda^*)_x u_x \, dx dt + \int_0^T u(\cdot, M) \mu^* \, dt + \int_{\Omega} u(0) \nu^* \, dx \end{aligned} \quad (2.28)$$

for all  $u \in W(0, T)$ . In particular, (2.28) holds for all  $u(t, x) = \chi(t)\psi(x)$  with  $\chi \in C_0^1(0, T)$  and  $\psi \in H_0^1(\Omega) \subset V$ . Consequently,

$$\int_0^T \int_{\Omega} \chi_t \psi \lambda^* + (q^* \lambda^*)_x \chi \psi' \, dx dt = 0 \quad (2.29)$$

for all  $\chi \in C_0^1(0, T)$  and  $\psi \in H_0^1(\Omega)$ . Notice that

$$\int_0^T \int_{\Omega} \chi_t \psi \lambda^* \, dx dt = \int_{\Omega} \left( \int_0^T \chi_t \lambda^* \, dt \right) \psi \, dx = \left\langle - \int_0^T \lambda_t^* \chi \, dt, \psi \right\rangle_{H^{-1}, H_0^1}, \quad (2.30)$$

where  $\lambda_t^*$  denotes the distributional derivative of  $\lambda^*$  with respect to  $t$ . The remaining term in (2.29) leads to

$$\int_0^T \int_{\Omega} (q^* \lambda^*)_x \psi' \chi \, dx dt = \left\langle - \int_0^T (q^* \lambda^*)_{xx} \chi \, dt, \psi \right\rangle_{H^{-1}, H_0^1}. \quad (2.31)$$

Inserting (2.30) and (2.31) into (2.29) we get

$$\left\langle \int_0^T (-\lambda_t^* - (q^* \lambda^*)_{xx}) \chi \, dt, \psi \right\rangle_{H^{-1}, H_0^1} = 0, \quad (2.32)$$

for all  $\chi \in C_0^1(0, T)$  and  $\psi \in H_0^1(\Omega)$ . Notice that  $q^* \in \mathcal{Q}_{\text{ad}}$  implies  $q^* \in L^\infty(Q)$  as well as  $q_x^* \in L^\infty(0, T; L^2(\Omega))$ . Therefore, it follows  $(q^* \lambda^*)_x \in L^2(Q)$  and, consequently,  $(q^* \lambda^*)_{xx} \in L^2(0, T; H^{-1}(\Omega))$ . The set

$$\{\varphi \in L^2(0, T; H_0^1(\Omega)) : \varphi(t, x) = \chi(t)\psi(x) \text{ with } \chi \in C_0^1(0, T) \text{ and } \psi \in H_0^1(\Omega)\}$$

is dense in  $L^2(0, T; H_0^1(\Omega))$  so that  $\lambda_t^* \in L^2(0, T; H^{-1}(\Omega))$  and (2.21a) hold. Moreover,

$$\begin{aligned} \int_0^T \frac{d}{dt} \langle \lambda^*, u \rangle_{L^2(\Omega)} dt &= \langle \lambda_t^*, u \rangle_{L^2(0, T; H^{-1}(\Omega)), L^2(0, T; H_0^1(\Omega))} \\ &\quad + \langle u_t, \lambda^* \rangle_{L^2(0, T; H^{-1}(\Omega)), L^2(0, T; H_0^1(\Omega))} \end{aligned} \quad (2.33)$$

for  $u \in W(0, T)$ . Hence, we may apply (2.28), (2.32), and (2.33) to obtain

$$\begin{aligned} 0 &= \int_{\Omega} (u^*(T) - u_T) u(T) dx + \int_0^T \frac{d}{dt} \langle \lambda^*, u \rangle_{L^2(\Omega)} dt \\ &\quad + \int_0^T \langle -\lambda_t^* - (q^* \lambda^*)_{xx}, u \rangle_{H^{-1}, H_0^1(\Omega)} + \int_0^T (q^* \lambda^*)_x u|_{x=M}^{x=R} dt \\ &\quad + \int_0^T \mu^* u(\cdot, M) dt + \int_{\Omega} \nu^* u(0) dx \\ &= \langle (u^*(T) - u_T + \lambda^*(T), u(T)) \rangle_{L^2(\Omega)} + \langle \nu^* - \lambda^*(0), u(0) \rangle_{L^2(\Omega)} \\ &\quad + \langle \mu^* - (q^*(\cdot, M) \lambda^*(\cdot, M))_x, u(\cdot, M) \rangle_{L^2(0, T)}. \end{aligned}$$

Choosing appropriate test functions in  $W(0, T)$ , we find (2.21c), (2.22), and (2.23). Finally, we consider (2.25). We compute

$$L_q(q^*, u^*, p^*)q = \int_0^T \int_{\Omega} \beta(q_t^* q_t + q^* q + q_x^* q_x + q_{xx}^* q_{xx}) + (q \lambda^*)_x u_x^* dx dt \quad (2.34)$$

for all  $q \in \mathcal{Q}_{\text{ad}}$ . For  $\lambda^* \in L^2(0, T; H_0^1(\Omega))$  and  $u_x^* \in L^2(Q)$  the integral

$$\int_0^T \int_{\Omega} (q \lambda^*)_x u_x^* dx dt$$

is bounded for all  $q \in \mathcal{Q}_{\text{ad}}$ . Moreover,  $(\lambda^* u_x^*)(\cdot, M) = (\lambda^* u_x^*)(\cdot, R) = 0$  holds. Thus, the function  $g = -\lambda^* u_{xx}^*$  can be identified with an element in  $(H^{2,1}(Q))'$  and we derive from (2.34)

$$L_q(q^*, u^*, p^*)q = \beta \langle q^*, q \rangle_{H^{2,1}(Q)} + \langle g, q \rangle_{(H^{2,1}(Q))', H^{2,1}(Q)} \quad (2.35)$$

for all  $q \in \mathcal{Q}_{\text{ad}}$ . Employing the Riesz isomorphism  $\mathcal{R}$ , inserting (2.35) into (2.25) we find

$$\langle \beta q^* - \mathcal{R}(\lambda^* u_{xx}^*), q - q^* \rangle_{H^{2,1}(Q)} \geq 0 \quad \text{for all } q \in \mathcal{Q}_{\text{ad}},$$

which is the variational inequality (2.24).  $\square$

**Remark 2.11.** The usage of the Riesz operator  $\mathcal{R} : (H^{2,1}(Q))' \rightarrow H^{2,1}(Q)$  in (2.24) requires to solve a problem of the form

$$-u_{tt} + u_{xxxx} - u_{xx} + u = f \quad \text{in } Q,$$

including initial and boundary conditions. Hence, in our numeric realization we will employ the ‘weaker’ norm in  $L^2(0, T; H^1(\Omega))$ , see Section 3. Then  $\mathcal{R}$  can be replaced by the Riesz operator  $\tilde{\mathcal{R}} : (H^1(\Omega))' \rightarrow H^1(\Omega)$ , that requires only the solution of the Neumann problem

$$-u(t)_{xx} + u(t) = f(t) \quad \text{in } \Omega, \quad u(t)_x|_{\partial\Omega} = 0,$$

for a.e.  $t \in (0, T)$ .

Utilizing variational techniques we can prove the following error estimate for the adjoint variable  $\lambda^*$ .

**Corollary 2.12.** *Let all hypotheses of Theorem 2.10 hold. Then there exists a constant  $C_2 > 0$  depending on  $\|q^*\|_{L^\infty(0,T;L^4(\Omega))}$  and  $\bar{q}_{\min}$  such that*

$$\|\lambda^*\|_{L^2(0,T;H_0^1(\Omega))} \leq C_2 \|u^*(T) - u_T\|_{L^2(\Omega)}$$

Hence, if the residual  $\|u^*(T) - u_T\|_{L^2(\Omega)}$  becomes small the norm of the Lagrange multiplier  $\lambda^*$  is small. We will make use of this estimate in the next section.

From Theorem 2.10 we infer the existence of a Lagrange multiplier associated with the constraint  $q^* \in \mathcal{Q}_{\text{ad}}$ . To formulate the result we introduce the following sets.

**Definition 2.13.** *Let  $K$  be a convex subset of a (real) Banach space  $Z$  and  $z^* \in K$ . The cone of feasible directions  $R_K$  at the point  $z^*$ , the tangent cone  $T_K$  at the point  $z^*$  and the normal cone  $N_K$  at the point  $z^*$  are defined by*

$$\begin{aligned} R_K(z^*) &= \{z \in Z : \exists \sigma > 0 : z^* + \sigma z \in K\}, \\ T_K(z^*) &= \{z \in Z : \exists z^*(\sigma) = z^* + \sigma z + o(\sigma) \in K, \sigma \geq 0\}, \\ N_K(z^*) &= \{z \in Z' : \langle z, \tilde{z} - z^* \rangle_{Z',Z} \leq 0 \text{ for all } \tilde{z} \in K\}. \end{aligned}$$

In case of  $z^* \notin K$  the normal cone  $N_K(z^*)$  is set equal to the empty set.

Let us recall the concept of polyhedricity.

**Definition 2.14.** *Let  $K$  be a closed convex subset of the Hilbert space  $Z$ ,  $z \in Z$  and  $v \in N_K(z)$ . Then  $K$  is called polyhedric at  $z$  for the normal direction  $v$  provided*

$$T_K(z) \cap \{v\}^\perp = \overline{R_K(z) \cap \{v\}^\perp}.$$

If  $K$  is polyhedric at each  $z \in K$  for all directions  $v \in N_K(z)$ , we call  $K$  polyhedric.

In the following we choose  $Z = H^{2,1}(Q)$ ,  $K = \mathcal{Q}_{\text{ad}}$  and  $z^* = q^*$ . Then, the following proposition follows directly from [11, Prop. 4.3].

**Proposition 2.15.** *The closed convex set  $\mathcal{Q}_{\text{ad}}$  is polyhedric.*

**Corollary 2.16.** *Let all hypotheses of Theorem 2.10 be satisfied. Then there exists a Lagrange multiplier  $\tilde{\xi}^* \in N_{\mathcal{Q}_{\text{ad}}}(q^*)$  associated with the inequality constraints such that*

$$L_q(\omega^*, p^*) + \tilde{\xi}^* = 0 \quad \text{in } (H^{2,1}(Q))'. \quad (2.36)$$

*Proof.* Defining  $\tilde{\xi}^* = -\beta q^* + \lambda^* u_{xx}^* \in (H^{2,1}(Q))'$  and using (2.24) we obtain  $\tilde{\xi}^* \in N_{\mathcal{Q}_{\text{ad}}}(q^*)$ . In particular, (2.36) follows.  $\square$

**Remark 2.17.** Using the Riesz isomorphism  $\mathcal{R}$  introduced in Theorem 2.10 we can identify  $\tilde{\xi}^* \in (H^{2,1}(Q))'$  with an element in the Hilbert space  $H^{2,1}(Q)$  by setting  $\xi^* = -\beta q^* + \mathcal{R}(\lambda^* u_{xx}^*)$ .

Let  $\omega^* = (q^*, u^*) \in K_{\text{ad}}$  denote a local solution to  $(\mathbf{P})$ . If the solution  $q^* \in \mathcal{Q}_{\text{ad}}$  is inactive with respect to the norm constraint, i.e.,  $\|q^*\|_{L^\infty(0,T;H^2(\Omega))} < C_{\text{ad}}$ , then  $(\mathbf{P})$  is locally equivalent to

$$\min J(\omega) \quad \text{s.t.} \quad \omega \in \hat{K}_{\text{ad}} \text{ and } e(\omega) = 0, \quad (\hat{\mathbf{P}})$$

where  $\hat{K}_{\text{ad}} = \hat{\mathcal{Q}}_{\text{ad}} \times W(0, T)$  and

$$\hat{\mathcal{Q}}_{\text{ad}} = \{q \in H^{2,1}(Q) : q_{\min} \leq q \leq q_{\max} \text{ in } Q \text{ a.e.}\},$$

which is a closed, convex and bounded subset in  $L^2(Q)$ . We define by

$$\hat{E} = \{\omega \in \hat{K}_{\text{ad}} : e(\omega) = 0\}$$

the admissible set of  $(\hat{\mathbf{P}})$ . Monitoring the sequence  $\|q^n\|_{L^\infty(0,T;H^2(\Omega))}$  we solve  $(\hat{\mathbf{P}})$  in our numerical experiments, see Section 4 below. For that reason we focus on  $(\hat{\mathbf{P}})$  in the remainder of this section.

The Lagrange multiplier  $\xi^*$  associated with the inequality constraints for the optimal coefficient  $q^*$  is characterized by the following corollary.

**Corollary 2.18.** *Let all hypotheses of Theorem 2.10 be satisfied. Suppose that  $\|q^*\|_{L^\infty(0,T;H^2(\Omega))} < C_{\text{ad}}$ . Then  $\xi^*$  satisfies*

$$\xi^*|_{A_-^*} \leq 0, \quad \xi^*|_{A_+^*} \geq 0, \quad \xi^*|_{I^*} = 0, \quad (2.37)$$

where

$$\begin{aligned} A_-^* &= \{(t, x) \in \overline{Q} : q^*(t, x) = q_{\min}(t, x)\}, \\ A_+^* &= \{(t, x) \in \overline{Q} : q^*(t, x) = q_{\max}(t, x)\}, \\ I^* &= \{(t, x) \in \overline{Q} : q_{\min}(t, x) < q^*(t, x) < q_{\max}(t, x)\} \end{aligned}$$

are the active and inactive sets for the optimal coefficient  $q^*$ .

*Proof.* The proof uses similar arguments as the proof of Theorem 2.3 in [27]. Therefore, we give only the proof of  $\xi^*|_{A_-^*} \leq 0$ . Define

$$\begin{aligned} A_-^\geq &= \{(t, x) \in \overline{Q} : (q^*(t, x) = q_{\min}(t, x)) \wedge (\xi^* > 0)\}, \\ A_-^{\geq, l} &= \left\{ (t, x) \in A_-^\geq : \xi^* > \frac{1}{l} \right\}, \\ C_-^l &= \left\{ (t, x) \in A_-^{\geq, l} : q_{\max}(t, x) - q_{\min}(t, x) > \frac{1}{l} \right\}. \end{aligned}$$

Assume that  $A_-^\geq$  has positive measure  $\mu(A_-^\geq) > \varepsilon > 0$ . Since

$$\mu\{(t, x) \in \overline{Q} : q_{\max}(t, x) = q_{\min}(t, x)\} = 0$$

and  $A_-^{\geq, l} \uparrow A_-^\geq$  for  $l \rightarrow \infty$ , it follows  $\mu(C_-^l) > 0$  for  $l$  sufficiently large and  $C_-^l \uparrow A_-^\geq$ . Hence there exists  $l > 0$  such that  $\mu(C_-^l) > \varepsilon$  because of the lower continuity of  $\mu$ . Define  $\delta \in (H^{2,1}(Q))'$  by  $\varphi \mapsto \int_0^T \int_\Omega (q_{\max} - q_{\min}) \chi_{C_-^l} \varphi \, dx dt$  and its Riesz representative by  $\mathcal{R}(\delta) \in H^{2,1}(Q)$ . Recall that  $\xi^* = -\beta q^* + \mathcal{R}(\lambda^* u_{xx}^*)$  by Remark 2.17 and consider the directional derivative (see (2.35))

$$\begin{aligned} L_q(q^*, u^*, p^*) \mathcal{R}(\delta) &= \langle \mathcal{R}(\delta), \beta q^* - \mathcal{R}(\lambda^* u_{xx}^*) \rangle_{H^{2,1}(Q)} \\ &= \langle \delta, -\xi^* \rangle_{(H^{2,1})', H^{2,1}} \\ &= - \int_0^T \int_\Omega (q_{\max} - q_{\min}) \chi_{C_-^l} \xi^* \, dx dt < -\frac{\varepsilon}{l^2} < 0. \end{aligned}$$

This contradicts the optimality of  $q^*$ . Hence,  $\mu(A_-^\geq) = 0$ .  $\square$

The primal–dual active set algorithm used below makes use of the following result from convex analysis [26, 31]. Using the generalized Moreau–Yosida regularization

of the indicator function  $\chi_{\hat{\mathcal{Q}}_{\text{ad}}}$  of the convex set  $\hat{\mathcal{Q}}_{\text{ad}}$  of admissible controls, i.e.,

$$\chi_{\hat{\mathcal{Q}}_{\text{ad}}}(q^*) = \inf_{q \in H^{2,1}(Q)} \left\{ \chi_{\hat{\mathcal{Q}}_{\text{ad}}}(q^* - q) + \langle \xi^*, q \rangle_{H^{2,1}(Q)} + \frac{c}{2} \|q\|_{H^{2,1}(Q)}^2 \right\}$$

with  $c > 0$ , one can replace  $q^* \in \hat{\mathcal{Q}}_{\text{ad}}$  and condition (2.37) by

$$q^* = \mathcal{P}_{\text{ad}} \left( q^* + \frac{\xi^*}{c} \right) \quad \text{for every } c > 0, \quad (2.38)$$

where

$$\mathcal{P}_{\text{ad}} : L^2(Q) \rightarrow \{q \in L^2(Q) \mid q_{\min} \leq q \leq q_{\max} \text{ in } Q \text{ a.e.}\}$$

by

$$\mathcal{P}_{\text{ad}}(q)(t, x) = \begin{cases} q_{\min}(t, x) & \text{if } q(t, x) < q_{\min}(t, x), \\ q(t, x) & \text{if } q_{\min}(t, x) \leq q(t, x) \leq q_{\max}(t, x), \\ q_{\max}(t, x) & \text{if } q(t, x) > q_{\max}(t, x) \end{cases}$$

for almost all  $(t, x) \in Q$ . It can be proved that (2.38) is equivalent to the differential inclusion  $\xi^* \in \partial \chi_{\hat{\mathcal{Q}}_{\text{ad}}}(q^*)$  (see [3]), where  $\partial \chi_{\hat{\mathcal{Q}}_{\text{ad}}}$  denotes the subdifferential of the indicator function  $\chi_{\hat{\mathcal{Q}}_{\text{ad}}}$ .

The primal–dual active set method uses the identification (2.38) as a prediction strategy, i.e. for a current primal–dual iteration pair  $(q_k, \xi_k)$  and arbitrarily fixed  $c > 0$  the next active and inactive sets are given by

$$\begin{aligned} \mathcal{A}_-^k &= \{(t, x) \in Q \mid q_k + \frac{\xi_k}{c} < q_{\min}\}, \quad \mathcal{A}_+^k = \{(t, x) \in Q \mid q_k + \frac{\xi_k}{c} > q_{\max}\}, \\ \mathcal{I}^k &= Q \setminus (\mathcal{A}_-^k \cup \mathcal{A}_+^k). \end{aligned}$$

**2.4. Second-order analysis.** In Section 2.3 we have investigated first-order necessary optimality conditions for  $(\hat{\mathbf{P}})$ . To ensure that a solution  $(\omega^*, p^*)$  satisfying  $\omega^* = (q^*, u^*) \in \hat{E}$ ,  $q^* \in \hat{\mathcal{Q}}_{\text{ad}}$ , (2.21) and (2.37) indeed solves  $(\hat{\mathbf{P}})$ , we have to guarantee second-order sufficient optimality. This is the focus of this section. We review different second-order optimality conditions and set them into relation. Then, we prove that the second-order sufficient optimality condition holds, provided the residual  $\|u^*(T) - u_T\|_{L^2(\Omega)}$  is sufficiently small.

For any directions  $\delta\omega = (\delta q, \delta u)$ ,  $\widetilde{\delta\omega} = (\widetilde{\delta q}, \widetilde{\delta u}) \in X$  the second Fréchet derivative of the Lagrangian is given by

$$\begin{aligned} L''(\omega, p)(\delta\omega, \widetilde{\delta\omega}) &= \beta \int_0^T \int_{\Omega} \delta q_t \widetilde{\delta q}_t + \delta q \widetilde{\delta q} + \delta q_x \widetilde{\delta q}_x + \delta q_{xx} \widetilde{\delta q}_{xx} \, dx dt \\ &\quad + \int_{\Omega} \delta u(T) \widetilde{\delta u}(T) \, dx + \int_0^T \int_{\Omega} (\delta q \lambda)_x \widetilde{\delta u}_x + (\widetilde{\delta q} \lambda)_x \delta u_x \, dx dt \end{aligned}$$

with  $\omega = (q, u) \in X$  and  $p = (\lambda, \mu, \nu) \in Y$ . In particular, we set

$$\begin{aligned} \mathbf{Q}(\delta\omega) &= L''(\omega, p)(\delta\omega, \delta\omega) \\ &= \|\delta u(T)\|_{L^2(\Omega)}^2 + \beta \|\delta q\|_{H^{2,1}(Q)}^2 + 2 \int_0^T \int_{\Omega} (\delta q \lambda)_x \delta u_x \, dx dt \end{aligned}$$

for  $\delta\omega \in X$ . From the boundedness of the second derivative of the Lagrangian we infer that  $\mathbf{Q}$  is continuous.

**Lemma 2.19.** *The quadratic form  $\mathbf{Q}$  is weakly lower semi-continuous. Moreover, let  $(\delta\omega^n)_{n \in \mathbb{N}}$  be a sequence in  $N(e'(\omega))$ ,  $\omega = (q, u) \in X$ , with  $\delta\omega^n \rightharpoonup 0$  in  $X$  and  $\mathbf{Q}(\delta\omega^n) \rightarrow 0$  as  $n \rightarrow \infty$ . Then it follows that  $\delta\omega^n \rightarrow 0$  strongly in  $X$ .*

*Proof.* Note that for  $\delta\omega = (\delta q, \delta u) \in X$  it holds

$$Q(\delta\omega) = J''(\omega)(\delta\omega, \delta\omega) + 2 \int_0^T \int_{\Omega} (\delta q \lambda)_x \delta u_x \, dx dt,$$

and  $\delta\omega \mapsto J''(\omega)(\delta\omega, \delta\omega)$  is weakly lower semi-continuous. Since the integral is even weakly continuous (see the proof of Theorem 2.9), it follows that  $Q$  is weakly lower semi-continuous on  $X$ . Now assume that  $(\delta\omega^n)_{n \in \mathbb{N}} = (\delta q^n, \delta u^n)_{n \in \mathbb{N}}$  is a sequence in  $N(e'(\omega))$  with  $\delta\omega^n \rightharpoonup 0$  in  $X$  and  $Q(\delta\omega^n) \rightarrow 0$  as  $n \rightarrow \infty$ . Analogously as in the proof of Theorem 2.9 we derive that  $\delta q^n \rightarrow 0$  in  $C([0, T]; V)$  as  $n \rightarrow \infty$ . Thus,

$$\lim_{n \rightarrow \infty} \int_0^T \int_{\Omega} (\delta q^n \lambda)_x \delta u_x^n \, dx dt = 0.$$

Since  $Q(\delta\omega^n)$  converges to zero, it follows that for every  $\varepsilon > 0$  there exists an  $n_\varepsilon \in \mathbb{N}$  such that

$$0 \leq J''(\omega)(\delta\omega^n, \delta\omega^n) < \varepsilon \quad \text{for all } n \geq n_\varepsilon.$$

In particular, this implies that

$$\beta \|\delta q^n\|_{H^{2,1}(Q)}^2 < \varepsilon \quad \text{for all } n \geq n_\varepsilon,$$

which gives  $\delta q^n \rightarrow 0$  in  $H^{2,1}(Q)$  as  $n \rightarrow \infty$ . Here we use that  $\beta > 0$  holds. Since  $\delta\omega^n \in N(e'(\omega))$  holds, we infer from Proposition 2.6 that  $\delta u^n \rightarrow 0$  in  $W(0, T)$  as  $n \rightarrow \infty$ .  $\square$

Let us recall the following definition, see [11].

**Definition 2.20.** Let  $\omega^* = (q^*, u^*) \in \hat{E}$ .

- a) The point  $\omega^*$  is a local solution to  $(\hat{\mathbf{P}})$  satisfying the quadratic growth condition if there exists a  $\rho > 0$  satisfying

$$J(\omega) \geq J(\omega^*) + \rho \|\omega - \omega^*\|_X^2 + o(\|\omega - \omega^*\|_X^2) \quad \text{for all } \omega \in \hat{E}. \quad (2.39)$$

- b) Suppose that  $\omega^*$  satisfies the first-order necessary optimality conditions with associated unique Lagrange multipliers  $p^* \in Y$  and  $\xi^* \in N_{\hat{Q}_{\text{ad}}}(q^*)$ . At  $(\omega^*, p^*)$  the second-order sufficient optimality condition holds if there exists a constant  $\kappa > 0$  such that

$$L''(\omega^*, p^*)(\delta\omega, \delta\omega) \geq \kappa \|\delta\omega\|_X^2 \quad \text{for all } \delta\omega \in C(\omega^*), \quad (2.40)$$

where

$$C(\omega^*) = \left\{ \delta\omega \in (T_{\hat{Q}_{\text{ad}}}(q^*) \cap \{\xi^*\}^\perp) \times W(0, T) : \delta\omega \in N(e'(\omega^*)) \right\}$$

denotes the critical cone at  $\omega^*$ ,  $^\perp$  denotes the orthogonal complement in  $H^{2,1}(Q)$  and  $T_{\hat{Q}_{\text{ad}}}(q^*)$  the tangential cone at  $q^*$  (introduced in Def. 2.13).

The critical cone  $C(\omega^*)$  is the set of directions that are tangent to the feasible set. It turns out that (2.39) and (2.40) are related to the weaker condition

$$L''(\omega^*, p^*)(\delta\omega, \delta\omega) > 0 \quad \text{for all } \delta\omega \in C(\omega^*) \setminus \{0\}, \quad (2.41)$$

which is very close to the necessary optimality condition. In particular, the following theorem holds.

**Theorem 2.21.** The quadratic growth condition (2.39), the second-order sufficient optimality condition (2.40), and (2.41) are equivalent.



*Proof.* The proof is similar to the proof of Theorem 2.7 in [11]. We show (2.39)  $\implies$  (2.40)  $\implies$  (2.41)  $\implies$  (2.39). Assume that (2.39) holds,  $\omega^*$  satisfies the first-order necessary optimality conditions with associated Lagrange multipliers  $p^* \in W(0, T)$  and  $\xi^* \in N_{\hat{\mathcal{Q}}_{\text{ad}}}(q^*)$ . Then there exists a number  $\varrho > 0$  such that  $\omega^*$  is also a local solution to

$$\min_{\omega \in \hat{E}} J(\omega) - \frac{\varrho}{2} \|\omega - \omega^*\|_X^2.$$

Hence, due to the second-order necessary optimality conditions we have

$$L''(\omega^*, p^*)(\delta w, \delta w) - \varrho \|\delta w\|_X^2 \geq 0 \quad \text{for all } \delta w \in C(\omega^*).$$

This gives (2.40). From (2.40) we directly infer (2.41). Suppose now that (2.41) is satisfied, while (2.39) is not. Then there exists a sequence  $(\omega^n)_{n \in \mathbb{N}} \subset \hat{E}$  such that  $\omega^n \rightarrow \omega^*$  as  $n \rightarrow \infty$  and

$$J(\omega^n) < J(\omega^*) + \frac{1}{n} \|\omega^n - \omega^*\|_X^2. \quad (2.42)$$

Extracting if necessary a subsequence we can write  $\omega^n - \omega^* = t_n v^n$  with  $t_n \searrow 0$ ,  $\|v^n\|_X = 1$ , and  $v^n \rightharpoonup v$  in  $X$ . Next we prove that  $v \in C(\omega^*)$  holds. From  $\omega^n = \omega^* + t_n v^n \in \hat{\mathcal{Q}}_{\text{ad}} \times W(0, T)$  and  $t_n > 0$  for all  $n$  we infer that  $v^n \in R_{\hat{\mathcal{Q}}_{\text{ad}}}(\xi^*) \times W(0, T)$ . Since  $\overline{R_{\hat{\mathcal{Q}}_{\text{ad}}}(\xi^*)}$  is weakly closed,  $v \in \overline{R_{\hat{\mathcal{Q}}_{\text{ad}}}(\xi^*)} \times W(0, T)$ . By Proposition 2.15 the set  $\hat{\mathcal{Q}}_{\text{ad}}$  is polyhedric. Thus,  $v \in T_{\hat{\mathcal{Q}}_{\text{ad}}}(\xi^*) \times W(0, T)$ . Moreover,  $e(\omega^n) = e(\omega^*) = 0$  leads to

$$e'(\omega^*)v = \lim_{n \rightarrow \infty} \frac{e(\omega^* + t_n v^n) - e(\omega^*)}{t_n} = 0,$$

i.e.,  $v \in N(e'(q^*))$ . By second order expansion,

$$L(\omega^n, p^*) = L(\omega^*, p^*) + t_n L'(\omega^*, p^*)v^n + \frac{t_n^2}{2} L''(\omega^*, p^*)(v^n, v^n) + o(t_n^2).$$

Using again  $e(\omega^n) = e(\omega^*) = 0$  and (2.42), we obtain

$$J(\omega^*) + \frac{t_n^2}{n} \|v^n\|_X^2 > J(\omega^n) = J(\omega^*) + t_n L'(\omega^*, p^*)v^n + \frac{t_n^2}{2} L''(\omega^*, p^*)(v^n, v^n) + o(t_n^2). \quad (2.43)$$

It follows that  $L'(\omega^*, p^*)v = 0$ . Together with (2.36) we conclude that  $v_q \in \{\xi^*\}^\perp$  for  $v = (v_q, v_u) \in (T_{\hat{\mathcal{Q}}_{\text{ad}}}(\xi^*) \times W(0, T)) \cap N(e'(\omega^*))$ . Therefore, we have proved that  $v \in C(\omega^*)$ . Due to (2.25) and (2.26) we know  $L'(\omega^*, p^*)v^n \geq 0$ . Using (2.43) we deduce

$$\begin{aligned} \frac{2}{n} \|v^n\|_X^2 &\geq \frac{2L'(\omega^*, p^*)v^n}{t_n} + L''(\omega^*, p^*)(v^n, v^n) + o(1) \\ &\geq L''(\omega^*, p^*)(v^n, v^n) + o(1) = Q(v^n) + o(1). \end{aligned}$$

Consequently,  $\limsup_{n \rightarrow \infty} Q(v^n) \leq 0$ . Due to Lemma 2.19,  $Q$  is weakly lower semi-continuous and therefore  $Q(v) \leq 0$ . By (2.41) we obtain  $v = 0$ . Hence,  $v^n \rightarrow 0$  and  $\lim_{n \rightarrow \infty} Q(v^n) = 0$ . Again using Lemma 2.19 it follows that  $\|v^n\|_X \rightarrow 0$ , but this is a contradiction to  $\|v^n\|_X = 1$  for all  $n$ .  $\square$

In the next theorem we present a sufficient condition for the second-order sufficient optimality condition (2.40).

**Theorem 2.22.** *Let all hypotheses of Theorem 2.10 be satisfied. Then (2.40) holds provided  $\|u^*(T) - u_T\|_{L^2(\Omega)}$  is sufficiently small.*

*Proof.* Applying estimate (2.7) and Hölder's inequality, we estimate for arbitrary  $\delta\omega = (\delta q, \delta u) \in N(e'(\omega^*))$

$$\begin{aligned} & L''(\omega^*, p^*)(\delta\omega, \delta\omega) \\ & \geq \beta \|\delta q\|_{H^{2,1}(Q)}^2 + 2 \int_0^T \int_{\Omega} (\delta q \lambda^*)_x \delta u_x \, dx dt \\ & \geq \frac{\beta}{2} \|\delta q\|_{H^{2,1}(Q)}^2 + \frac{\beta}{2C_1} \|\delta u\|_{W(0,T)}^2 - 2C \|\delta q\|_{H^{2,1}(Q)} \|\lambda^*\|_{L^2(0,T;H_0^1(\Omega))} \|\delta u\|_{W(0,T)} \\ & \geq \left( \frac{\beta}{2} \min\{1, 1/C_1\} - CC_2 \|u^*(T) - u_T\|_{L^2(\Omega)} \right) \left( \|\delta q\|_{H^{2,1}(Q)}^2 + \|\delta u\|_{W(0,T)}^2 \right), \end{aligned}$$

where we used Corollary 2.12 and Young's inequality in the last inequality. Supposing that

$$\|u^*(T) - u_T\|_{L^2(\Omega)} < \frac{\beta \min\{1, 1/C_1\}}{2CC_2},$$

we conclude (2.40).  $\square$

### 3. THE OPTIMIZATION METHOD

In this section we turn to the optimization algorithm used to solve the parameter identification problem  $(\mathbf{P})$ . We suppose that  $\|q^*\|_{L^\infty(0,T;H^2(\Omega))} < C_{\text{ad}}$ . Since empirical results suggest that the volatility is quite regular, this is not a severe restriction for our application. Hence we solve  $(\hat{\mathbf{P}})$  instead of  $(\mathbf{P})$ . To solve  $(\hat{\mathbf{P}})$  we apply a globalized SQP method. The globalization is realized by a modification of the Hessian matrix to ensure that every SQP step is a descent direction and by a line search strategy. Since in each level of the SQP method a linear-quadratic optimal control problem with box constraints has to be solved we utilize a primal-dual active set strategy.

The globalized SQP method is addressed in Section 3.1. Section 3.2 is concerned with the primal-dual active set method. The line search strategy is discussed in Section 3.3.

**3.1. The globalized SQP method.** In the numerical realization we initialize our SQP method by taking a function  $u^0 \in W(0, T)$ , which satisfies  $u^0(\cdot, M) = u_D$  in  $(0, T)$  and  $u^0(0) = u_0$  in  $\Omega$ . Hence, the next iterate  $u^{n+1} = u^n + \delta u^n$ ,  $n \geq 0$ , can be determined by choosing  $\delta u^n$  in the linear space

$$\mathfrak{U} = \{\delta u \in W(0, T) : \delta u(t, \cdot) \in H_0^1(\Omega) \text{ for } t \in (0, T) \text{ and } \delta u(0) = 0 \text{ in } \Omega\},$$

which is a Hilbert space endowed with the topology of  $W(0, T)$ . Thus, the constraints  $e_2(\omega^n) = 0$  and  $e_3(\omega^n) = 0$  are guaranteed by construction for any  $n \geq 0$ . Consequently, there is only one constraint that is  $e_1(\omega) = 0$  in  $L^2(0, T; H^{-1}(\Omega))$ . Therefore, the Lagrange variables  $\mu$  and  $\nu$  are redundant. It follows from

$$\mu^* = (q^* \lambda^*)_x(\cdot, M) \quad \text{and} \quad \nu^* = \lambda^*(0)$$

that  $\mu^*$  and  $\nu^*$  can be computed after determining the optimal coefficient  $q^*$  and the Lagrange multiplier  $\lambda^*$  associated with the constraint  $e_1(\omega^*) = 0$ . We set  $Y_1 = L^2(0, T; H_0^1(\Omega))$  and identify its dual space  $Y_1'$  with  $L^2(0, T; H^{-1}(\Omega))$ .

It turns out that by replacing the  $H^{2,1}(Q)$ -norm for the regularization term in (2.14) by the 'weaker' norm in  $L^2(0, T; H^1(\Omega))$ , relative to a prior  $q_d$ , we get good results while saving computational effort, since we only have to solve a Neumann

problem for  $q$  in each time step (see Remark 2.11). Hence, in the following  $X = X_1 \times X_2 = L^2(0, T; H^1(\Omega)) \times W(0, T)$ .

The linear-quadratic minimization problems that have to be solved in each step of the SQP method are well-defined provided  $L''(\omega^n, \lambda^n)$  is coercive on  $N(e'(\omega^n))$  and  $e'(\omega^n)$  on  $N(e'(\omega^n))$  is surjective for every iterate. Often these requirements hold only locally so we consider in the following a globalization strategy using a modified Hessian; compare [26]. For  $\gamma \in [0, 1]$  let us define the function  $\gamma L : X \times Y_1 \rightarrow \mathbb{R}$  by (see [27])

$$\gamma L(\omega, \lambda) = J(\omega) + \gamma \langle e_1(\omega), \lambda \rangle_{Y'_1, Y_1} \quad \text{for } (\omega, \lambda) \in X \times Y_1.$$

Notice that for  $\gamma = 1$  the function  $\gamma L$  is the usual Lagrangian associated with the single constraint  $e_1(\omega) = 0$  in  $Y'_1$ . Therefore, we set  $L = \gamma L$  for  $\gamma = 1$ .

**Algorithm 1** (Globalized SQP method).

- 1) Choose  $\omega^0 = (q^0, u^0) \in X = X_1 \times X_2$  that satisfies  $u^0(\cdot, M) = u_D$  in  $(0, T)$  and  $u^0(0) = u_0$  in  $\Omega$ ,  $\lambda^0 \in Y_1$  with  $\lambda^0(T) = u_T - u^0(T)$  in  $\Omega$ . Fix relative and absolute stopping tolerances  $1 > \varepsilon_{\text{rel}} \geq \varepsilon_{\text{abs}} > 0$ . Choose the maximal number of SQP iterations  $n_{\text{sqp}} \in \mathbb{N}$ . Set  $n := 0$  and  $\bar{\kappa} \in (0, 1]$ .

- 2) For  $\omega^n = (q^n, u^n)$  evaluate  $e_1(\omega^n)$ ,  $e'_1(\omega^n)$ ,  $J(\omega^n)$ ,  $J'(\omega^n)$ . If

$$\|\nabla L(\omega^n, \lambda^n)\|_{X' \times Y'_1}^2 = \|L'(\omega^n, \lambda^n)\|_{X'}^2 + \|e_1(\omega^n)\|_{Y'_1}^2 < \varepsilon_{\text{abs}}$$

or

$$\|\nabla L(\omega^n, \lambda^n)\|_{X' \times Y'_1}^2 < \varepsilon_{\text{rel}} \|\nabla L(\omega^0, \lambda^0)\|_{X' \times Y'_1}^2$$

or  $n = n_{\text{sqp}}$ , then STOP. Otherwise continue with step 3).

- 3) Set  $\gamma = 1$ .

- 4) Solve for  $\delta\omega^n = (\delta q^n, \delta u^n)$  the following linear-quadratic optimal control problem

$$\begin{aligned} & \min L(\omega^n, \lambda^n) + L'(\omega^n, \lambda^n)\delta\omega + \frac{1}{2} \gamma L''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n) \\ & \text{s.t.} \quad \begin{cases} e'_1(\omega^n)\delta\omega^n + e_1(\omega^n) = 0 & \text{in } Y'_1, \\ q_{\min} \leq q^n + \delta q^n \leq q_{\max} & \text{in } X_1. \end{cases} \end{aligned} \quad (\text{QP})$$

- 5) If  $\|\delta\omega\|_X > 0$  evaluate the quotient

$$\kappa^n = \frac{\gamma L''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n)}{\|\delta q^n\|_{X_1}^2 + \|\delta u^n\|_{L^2(0, T; L^2(\Omega))}^2}.$$

where

$$\gamma L''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n) = J''(\omega^n)(\delta\omega^n, \delta\omega^n) + \gamma \langle e''_1(\omega^n)(\delta\omega^n, \delta\omega^n), \lambda^n \rangle_{Y'_1, Y_1}$$

holds. If  $\kappa^n \leq 0$  then set  $\gamma = 0$  and go back to step 4). If  $\kappa^n \in (0, \bar{\kappa})$  then set  $\bar{\kappa} = \kappa^n$ .

- 6) Determine a step size parameter  $\alpha_n \in (0, 1]$  by a backtracking line search (see Section 3.3).

- 7) Set  $q^{n+1} = q^n + \alpha_n \delta q^n$ ,  $u^{n+1} = u^n + \alpha_n \delta u^n$ ,  $\lambda^{n+1} = \lambda^n + \alpha_n \delta \lambda^n$ , set  $n := n + 1$ , and go back to step 2).

**Remark 3.1.** 1) Alternatively,  $\gamma$  can be adjusted by using the following iteration strategy after step 3)

- a) Choose  $\eta \in (0, 1)$  and set  $i = 0$ .
- b) Perform steps 4) and 5).

c) If  $\kappa^n \leq 0$  set

$$\gamma := \min \left\{ \eta, \frac{\bar{\kappa} \|\delta\omega^n\|_X^2 - J''(\omega^n)(\delta\omega^n, \delta\omega^n)}{\langle e_1''(\omega^n)(\delta\omega^n, \delta\omega^n), \lambda^n \rangle_{Y_1', Y_1}} \right\}. \quad (3.1)$$

Set  $i := i + 1$  and go back to b).

From  $J''(\omega^n)(\delta\omega^n, \delta\omega^n) \geq 0$  and  $\kappa^n \leq 0$  it follows that

$$\langle e_1''(\omega^n)(\delta\omega^n, \delta\omega^n), \lambda^n \rangle_{Y_1', Y_1} < 0.$$

The procedure using (3.1) is less strict than setting directly  $\gamma = 0$  if  $\kappa^n \leq 0$ . However, it may involve solving **(QP)** several times, i.e., more often than at most two times as needed by the strategy that switches directly from  $\gamma = 1$  to  $\gamma = 0$ .

- 2) In Algorithm 1, the positive scalar  $\bar{\kappa}$  serves as an estimate for the coercivity constant in (2.40).

**3.2. The primal–dual active set method.** To solve the linear–quadratic optimal control problem **(QP)** in step 4) of Algorithm 1 we apply the primal–dual active set method [27]. It involves both primal and dual variables and is therefore different from conventional active set strategies that involve primal variables only, see, e.g. [42]. In practice, the algorithm behaves like an infeasible one, since its iterates violate the constraints up to the last–but–one iterate. The algorithm stops at a feasible and optimal solution. Based on the identification (2.38) for the inactive and active sets, the algorithm exhibits a low number of iterations to find the optimal solution and is very robust [27, 6]. Note also that the algorithm uses only one Lagrange multiplier to realize both inequality constraints, which reduces the number of variables and hence the amount of memory needed by its implementation.

Suppose that we are at level  $n$  of the SQP method. Thus, we have iterates  $(q^n, u^n, \lambda^n)$  and start our primal–dual active set strategy to determine the directions  $(\delta q^n, \delta u^n, \delta \lambda^n)$ . Let us define  $q_{\min}^n = q_{\min} - q^n$  and  $q_{\max}^n = q_{\max} - q^n$ . Then we are looking for a step  $\delta q^n$  satisfying the inequality constraints  $q_{\min}^n \leq \delta q^n \leq q_{\max}^n$  in  $Q$ . The method uses the identification (2.38) as a prediction strategy, i.e. given a current primal–dual iteration pair  $(\delta q_k, \xi_k)$  the next inactive and active sets are given by (3.2)–(3.4) (see below). The method stops at a feasible and optimal solution as soon as two consecutive tuples of active and inactive sets are equal. This solution is then used in Algorithm 1.

Instead of solving problem **(QP)** with inequality constraints we solve

$$\begin{aligned} & \min L(\omega^n, \lambda^n) + L'(\omega^n, \lambda^n)\delta\omega + \frac{1}{2} \gamma L''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n) \\ & \text{s.t.} \quad \begin{cases} e_1'(\omega^n)\delta\omega^n + e_1(\omega^n) = 0 & \text{in } Y_1', \\ \delta q_k = q_{\min}^n & \text{in } \mathcal{A}_-^k, \\ \delta q_k = q_{\max}^n & \text{in } \mathcal{A}_+^k, \\ q_{\min}^n \leq q^n + \delta q_k \leq q_{\max}^n & \text{in } X_1. \end{cases} \quad (\mathbf{QP}_{\mathcal{I}_k}) \end{aligned}$$

Hence, it is solved for  $\delta q_k$  only on the inactive set  $\mathcal{I}_k$ , but  $\delta q_k$  is fixed on the active set  $\mathcal{A}^k = \mathcal{A}_-^k \cup \mathcal{A}_+^k$ . We have the following algorithm.

**Algorithm 2** (Primal–dual active set method).

- 1) Choose initial values  $(\delta q_0, \xi_0) \in X_1 \times X_1$ , the parameter  $c > 0$  and set  $k := 0$ .

2) Determine the active and inactive sets

$$\mathcal{A}_-^k = \{(t, x) \in Q \mid \delta q_k + \frac{\xi_k}{c} < q_{\min}^n\}, \quad (3.2)$$

$$\mathcal{A}_+^k = \{(t, x) \in Q \mid \delta q_k + \frac{\xi_k}{c} > q_{\max}^n\}, \quad (3.3)$$

$$\mathcal{I}^k = Q \setminus \mathcal{A}^k. \quad (3.4)$$

with  $\mathcal{A}^k = \mathcal{A}_-^k \cup \mathcal{A}_+^k$ , and set

$$\widetilde{\delta q_k} = \begin{cases} q_{\min}^n & \text{in } \mathcal{A}_-^k, \\ q_{\max}^n & \text{in } \mathcal{A}_+^k, \\ \delta q_k & \text{in } \mathcal{I}^k. \end{cases}$$

3) If  $k > 0$  and  $\mathcal{A}^k = \mathcal{A}^{k-1}$  holds then STOP.

4) Solve for  $(\delta q_k, \delta u_k, \delta \lambda_k) \in X_1 \times \mathfrak{U} \times Y_1$  the linearized state equations

$$(\delta u_k)_t - q^n(\delta u_k)_{xx} - \widetilde{\delta q_k} u_{xx}^n = -(u_t^n - q^n u_{xx}^n) \quad \text{in } Q, \quad (3.5a)$$

$$\delta u_k(\cdot, M) = \delta u_k(\cdot, R) = 0 \quad \text{in } (0, T), \quad (3.5b)$$

$$\delta u_k(0) = 0 \quad \text{in } \Omega, \quad (3.5c)$$

the linearized adjoint system

$$-(\delta \lambda_k)_t - (q^n \delta \lambda_k)_{xx} - \gamma(\widetilde{\delta q_k} \lambda^n)_{xx} = -(-\lambda_t^n - (q^n \lambda^n)_{xx}) \quad \text{in } Q, \quad (3.5d)$$

$$\delta \lambda_k(\cdot, M) = \delta \lambda_k(\cdot, R) = 0 \quad \text{in } (0, T), \quad (3.5e)$$

$$\delta \lambda_k(T) + \delta u_k(T) = u_T - u^n(T) - \lambda^n(T) \quad \text{in } \Omega, \quad (3.5f)$$

and the linearized optimality condition on the inactive set

$$\begin{aligned} & \beta(-(\delta q_k)_{xx} + \delta q_k) - \delta \lambda_k u_{xx}^n - \gamma \lambda^n (\delta u_k)_{xx} \\ &= -\left(\beta(-(q^n - q_d)_{xx} + (q^n - q_d)) - \widetilde{\mathcal{R}}(\lambda^n u_{xx}^n)\right) \quad \text{in } \mathcal{I}^k, \end{aligned} \quad (3.5g)$$

where  $(q^n, u^n, \lambda^n)$  denotes the current SQP iterate and  $\widetilde{\mathcal{R}} = (-\Delta + \text{Id})^{-1}$  is the Riesz operator between  $(H^1(\Omega))'$  and  $H^1(\Omega)$ .

5) Set  $q_k^n = q^n + \delta q_k$ ,  $u_k^n = u^n + \delta u_k$ ,  $\lambda_k^n = \lambda^n + \delta \lambda_k$  and

$$\xi_k = -\beta(-(q_k^n - q_d)_{xx} + q_k^n - q_d) + \widetilde{\mathcal{R}}(\lambda_k^n (u_k^n)_{xx}) \quad \text{in } Q.$$

From (3.5g) we infer that  $\xi_k = 0$  on  $\mathcal{I}_k$ . Set  $k := k+1$  and go back to step 2).

Recall that we have chosen  $\lambda^0$  in such a way that  $\lambda^0(T) = u_T - u^0(T)$  holds in  $\Omega$ . Thus, by induction we have  $\lambda^n(T) = u_T - u^n(T)$  in  $\Omega$  and (3.5f) reads

$$\delta \lambda_k(T) + \delta u_k(T) = 0 \quad \text{in } \Omega.$$

**Remark 3.2.** In [30] a global and local convergence analysis is done for the non-linear primal-dual active set strategy in an abstract setting. It turns out that it converges globally as  $k \rightarrow \infty$  provided a certain merit function satisfies a sufficient decrease condition. This merit function depends on integrals of the positive part functions  $\max\{0, \delta q_k - q_{\max}^n\}$  and  $\max\{0, q_{\min}^n - \delta q_k\}$  over the set  $Q$  as well as of the positive part functions  $\max\{0, -\xi_k\}$  and  $\max\{0, \xi_k\}$  over the active sets. If, in addition, a smoothness and Lipschitz condition hold, the method converges locally superlinearly.

**3.3. The line search strategy.** In this section we address the line search strategy employed in step 6) of Algorithm 1. Notice that Algorithm 2 ensures feasibility of each solution to **(QP)** with respect of the solution to the inequality constraints. Thus, the goal is to find a compromise between the descent of the cost functional  $J$  and the reduction of the violation of the equality constraint  $e(\omega) = 0$ . This is realized by a line search method utilizing a suitable chosen merit function. Of course, there are many suggestions for merit functions in the literature; see, for instance, in [7, 42]. Here, we use the exact penalty functional

$$\varphi^n(\alpha^n) = J(\omega^n) + \mu \|e_1(\omega^n + \alpha^n \delta\omega^n)\|_{Y_1'}, \quad (3.6)$$

where  $\alpha^n \in (0, 1]$  is the step size parameter, which has to be determined, and  $\mu > 0$  is a parameter penalizing violations of the constraint  $e_1(\omega) = 0$ . The reason for our choice (3.6) comes from the fact that — apart from evaluating the norm — our merit function does not introduce an additional nonlinearity into our problem.

In a first-order variation we use the approximation

$$\begin{aligned} \bar{\varphi}^n(\alpha^n) &= J(\omega^n) + \alpha^n J'(\omega^n) \delta\omega^n + \mu \|e_1(\omega^n) + \alpha^n e_1'(\omega^n) \delta\omega^n\|_{Y_1'} \\ &= J(\omega^n) + \alpha^n J'(\omega^n) \delta\omega^n + \mu(1 - \alpha^n) \|e_1(\omega^n)\|_{Y_1'}, \end{aligned}$$

where we have used that  $e_1(\omega^n) + e_1'(\omega^n) \delta\omega^n = 0$  holds for the SQP step  $\delta\omega^n$ . Notice that  $\bar{\varphi}^n(1)$  does not depend on  $\mu$ . For appropriately chosen penalty parameter  $\mu > 0$ , our line search strategy is based on the well-known Armijo rule (see e.g. [10, 23])

$$\varphi^n(\alpha^n) - \varphi^n(0) \leq c\alpha^n (\bar{\varphi}^n(1) - \bar{\varphi}^n(0)) \quad \text{for } c \in \left(0, \frac{1}{2}\right). \quad (3.7)$$

In [26] sufficient conditions are given for the existence of a sufficiently large penalty parameter  $\bar{\mu} > 0$  such that

$$\bar{\varphi}^n(1) - \bar{\varphi}^n(0) < 0 \quad \text{for all } \mu \geq \bar{\mu}. \quad (3.8)$$

To find an appropriate value for  $\mu$  we check whether (3.8) holds and increase  $\mu$  if not and iterate. If (3.8) is fulfilled we determine  $\alpha^n$  from (3.7) by a backtracking strategy starting with  $\alpha^n = 1$ . If (3.7) is violated we decrease  $\alpha^n$  by setting  $\alpha^n := \zeta \alpha^n$  with  $\zeta \in (0, 1)$  and iterate until (3.7) holds or  $\alpha^n$  falls below a minimal step size parameter.

#### 4. NUMERICAL EXPERIMENTS

In this section we report the results of our numerical experiments. The algorithm proposed in Section 3 is discretized for the numerical realization. We make use of a finite element method with piecewise linear finite elements for the spatial discretization. To solve the discrete version of (3.5) we apply a preconditioned GMRES method.

A typical amount of data noise in option prices, that can be caused for example by bid-ask spreads, is  $\delta = 0.1\%$ . For the choice of the regularization parameter  $\beta$  we follow here the strategy proposed in [20], and define a decreasing sequence of admissible regularization parameters by

$$(\beta_1, \dots, \beta_6) = (10^2, 10^1, \dots, 10^{-3})\delta = (10^{-1}, 10^{-2}, \dots, 10^{-6}).$$

Strike $E$	95	97.5	100	102.5	105
True value	5.24433	3.23921	1.72734	0.77577	0.28866
Good guess	5.24429	3.23921	1.72734	0.77576	0.28861
Good guess & noise	5.24430	3.23921	1.72734	0.77576	0.28861
Bad guess	5.24435	3.23922	1.72733	0.77578	0.28866
Bad guess & noise	5.24435	3.23922	1.72733	0.77578	0.28866
Good guess, fine grid	5.24433	3.23921	1.72734	0.77577	0.28866

TABLE 4.1. Run 4.1: True option price and reconstructed option prices computed for different strikes  $E$  and different *a priori* guesses  $q_d$ . The true values were computed using the Black–Scholes formula with constant volatility  $\sigma = 0.15$ .

Strike $E$	95	97.5	100	102.5	105
True value	0.1500	0.1500	0.1500	0.1500	0.1500
Good guess	0.1454	0.1500	0.1517	0.1506	0.1470
Good guess & noise	0.1457	0.1500	0.1517	0.1506	0.1470
Bad guess	0.1458	0.1500	0.1517	0.1506	0.1472
Bad guess & noise	0.1460	0.1500	0.1517	0.1506	0.1472
Good guess, fine grid	0.1488	0.1500	0.1508	0.1502	0.1494

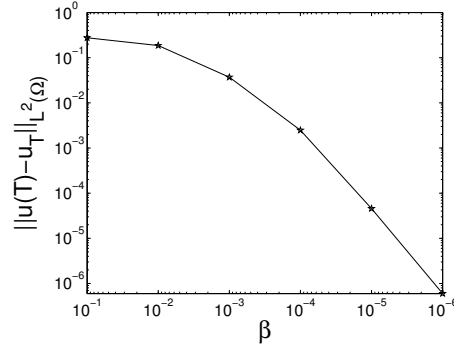
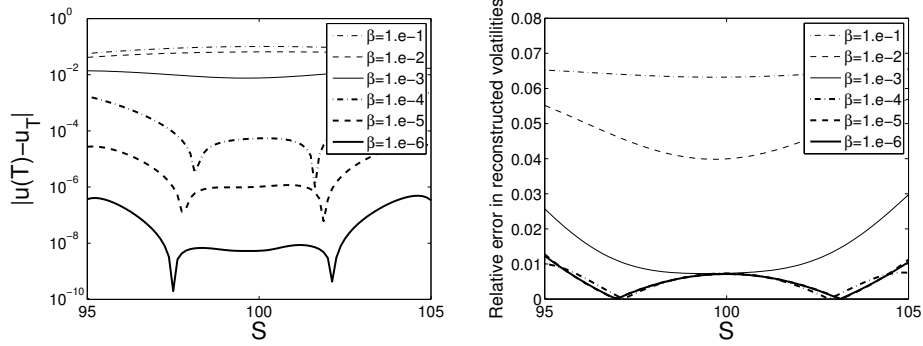
TABLE 4.2. Run 4.1: Reconstructed volatilities for different strikes  $E$  and different *a priori* guesses  $q_d$ .

We start minimizing the functional with the highest value  $\beta = \beta_1$  and subsequently decrease the regularization parameter, starting the method at the minimizers obtained in the previous step. For the SQP method we choose stopping tolerances  $\varepsilon_{\text{abs}} = 10^{-6}$  and  $\varepsilon_{\text{rel}} = 10^{-3}$  and a maximal number of iterations  $n_{\text{sqp}} = 20$ . We use a non-uniform grid with 140 nodes locally refined around  $x = S_0$  for the spatial discretization. In time, we employ a fixed, non-equidistant grid consisting of 35 points with small time steps close to  $t = 0$ .

**Run 4.1.** As a first example we apply our method to an artificial data set of Black–Scholes prices, i.e. prices computed with the Black–Scholes formula, with  $S_0 = 100$ ,  $r = 0$ , one month to maturity and constant volatility  $\sigma = 0.15$ . We consider four different cases with *a priori* guess  $q_d = \frac{1}{2}\sigma_d^2 x^2$ . In our first simulation we use the ‘good’ *a priori* guess  $\sigma_d = 0.16$ , in a second one we add 0.1% uniformly distributed noise. We compare these results to those from a third and fourth run using a ‘bad’ *a priori* guess  $\sigma_d = 0.1$ , where in the fourth run again we added 0.1% uniformly distributed noise. The resulting option prices are given in Table 4.1 and the corresponding volatilities are shown in Table 4.2. The error-free true values for the option prices were computed using the Black–Scholes formula. Table 4.3 displays the residuals  $\|u(T) - u_T\|_{L^2(\Omega)}$  remaining after reconstruction. In all four tests, the identified option prices correspond very well to the true values, the difference is neglectably small. The corresponding volatilities are well identified, with small differences remaining due to discretization errors, which can be reduced by using a finer grid. This can be seen from the results of a fifth run which was executed

	$\ u(T) - u_T\ _{L^2(\Omega)}$
Good guess	$6.76 \times 10^{-3}$
Good guess & noise	$3.56 \times 10^{-2}$
Bad guess	$6.50 \times 10^{-3}$
Bad guess & noise	$3.47 \times 10^{-2}$
Good guess, fine grid	$1.51 \times 10^{-3}$

TABLE 4.3. Run 4.1: Residuals remaining after reconstruction.

FIGURE 4.1. Run 4.1: Residual  $\|u(T) - u_T\|_{L^2(\Omega)}$  on  $[95, 105]$  for decreasing values of the regularization parameter  $\beta$ .FIGURE 4.2. Run 4.1: Error in reconstructed option prices (left figure) and relative error in the associated volatilities (right figure) for different values of the regularization parameter  $\beta$ .

on a grid with halved mesh size in space and time. Figures 4.1 and 4.2 show the influence of the regularization parameter  $\beta$ . Figure 4.1 is in double logarithmic scale and shows the residual  $\|u(T) - u_T\|_{L^2(\Omega)}$  for the fifth run on the domain  $[95, 105]$  for decreasing values of  $\beta$ . Figure 4.2 displays the error in reconstructed option prices (left figure) and the relative error in the associated volatilities (right figure) for different values of the regularization parameter  $\beta$ . Decreasing  $\beta$  further



Strike $E$	Price $C$
5825	469.5
6175	223.5
6225	195.5
6275	169.0
6325	144.5
6575	56.5
6875	10.5
7225	0.5

TABLE 4.4. Run 4.2: Call option prices for maturity 0.09589.

is not advisable since for very small values of  $\beta$ , the reconstructed  $\sigma$  shows oscillations, since we do not control the  $H^1$  norm (Remark 2.11), while the residual still decreases.

Overall, the method shows only a very small dependence on the chosen *a priori* guess and it is robust regarding to additional data noise. In all runs, Algorithm 1 needs very few iterations, typically one to three, to meet the prescribed stopping tolerances. We find that the identification process is very stable and the option prices and associated volatilities are very well recovered.

**Run 4.2.** In our second example we use market data from [14]. These data involve FTSE index call option prices from February 11, 2000. The option prices are given in Table 4.4. The spot price is  $S_0 = 6219$ , the constant interest rate is  $r = 0.061451$ , and the maturity is 0.095890. Note that meaningful empirical data, i.e., prices of options that are actually traded, are usually only available for strike prices  $E$  in a small region around the spot price  $S_0$ , typically  $S_0 \pm 10\%$  or even only  $S_0 \pm 5\%$  [33]. Restricting the computational domain to this small region is not advisable, hence one needs to extrapolate the data outside of this region. Here, we compute the implied volatilities of the options with the largest and the smallest strike price available by inverting the Black–Scholes formula, and use prices computed by the Black–Scholes formula with these volatilities in those regions in which no market data are available. The data are interpolated using a cubic spline. The resulting local volatility function is shown in Figure 4.3. It is skewed, as is typical for equity index options. The volatility function is higher for options in-the-money, i.e. for options with  $E < S_0$ , than for options at-the-money and out-of-the-money. Furthermore, it shows a term structure, with volatility decreasing as time approaches maturity. These characteristics are consistent with empirically observed patterns in equity index options [34].

## REFERENCES

- [1] Y. Achdou, O. Pironneau. Volatility smile by multilevel least square. *Int. J. Theor. Appl. Finance* **5**(6), 619-643, 2002.
- [2] M. Avellaneda, C. Friedman, R. Holmes, D. Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Appl. Math. Finance* **4**(1), 37-64, 1997.
- [3] V. Barbu. *Analysis and Control of Nonlinear Infinite Dimensional Systems*. Math. Sci. Engrg. 190. Boston Academic Press, 1993.
- [4] G. Barles, Ch. Daher, M. Romano. Convergence of numerical schemes for parabolic equations arising in finance theory. *Math. Models Meth. Appl. Sci.* **5**(1), 125-143, 1995.

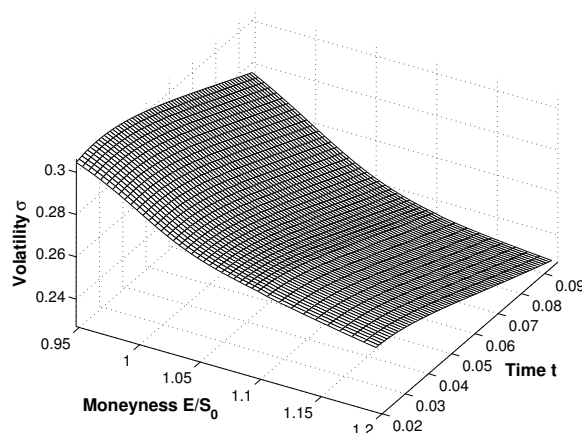


FIGURE 4.3. Run 4.2: Local volatility function computed from market data.

- [5] H. Berestycki, J. Busca, I. Florent. Asymptotics and calibration of local volatility models. *Quant. Finance* **2**, 61-69, 2002.
- [6] M. Bergounioux, K. Ito, K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.* **35**, 1524-1543, 1997.
- [7] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, Massachusetts, 1995.
- [8] F. Black, M. Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.* **81**, 637-659, 1973.
- [9] P.T. Boggs, J.W. Tolle. Sequential quadratic programming. In *Acta Numerica 1995*, A. Iserles (ed.), 1-51, Cambridge University Press, 1995.
- [10] J.F. Bonnans. *Optimisation Numérique*. Springer-Verlag, Paris, 1997.
- [11] J.F. Bonnans. Second-order analysis for control constrained optimal control problems of semi-linear elliptic systems. *Appl. Math. Optim.* **38**, 303-325, 1998.
- [12] I. Bouchouev, V. Isakov. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Inverse Probl.* **15**(3), R95-R116, 1999.
- [13] S. Crépey. Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization. *SIAM J. Math. Anal.* **34**(5), 1183-1206, 2003.
- [14] S. Crépey. Calibration of the local volatility in a trinomial tree using Tikhonov regularization. *Inverse Probl.* **19**(1), 91-127, 2003.
- [15] R. Dautray, J.-P. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems*. Berlin, Springer-Verlag, 1992.
- [16] J. Dewynne, S. Howison, P. Wilmott. *Option Pricing: Mathematical Models and Computation*. Oxford, Financial Press, 1995.
- [17] B. Dumas, J. Fleming, R.E. Whaley. Implied volatility functions: empirical tests. *J. Finance* **53**(6), 2059-2106, 1998.
- [18] B. Dupire. Pricing with a smile. *Risk* **7**, 18-20, 1994.
- [19] B. Dupire. A unified theory of volatility. In *Derivatives Pricing: The Classic Collection*, P. Carr (ed.), Risk Books, 2004.
- [20] H. Egger, H.W. Engl. Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Probl.* **21**, 1027-1045, 2005.
- [21] H.W. Engl, M. Hanke, A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996.
- [22] V. Girault, P.-A. Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*. Springer-Verlag, Berlin, 1980.

- [23] W.A. Gruver, E.W. Sachs. *Algorithmic Methods in Optimal Control*. Research Notes in Mathematics, vol. 47, Pitman, London, 1980.
- [24] M. Hanke, E. Rösler. Computation of local volatilities from regularized Dupire equations. *Int. J. Theor. Appl. Finance* **8**(2), 207-221, 2005.
- [25] M. R. Hestenes. Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math.* **1**, 525-581, 1951.
- [26] M. Hintermüller. On a globalized augmented Lagrangian-SQP algorithm for nonlinear optimal control problems with box constraints. In K.-H. Hoffmann, R. H. W. Hoppe, V. Schulz, editors, *Fast Solution of Discretized Optimization Problems*, International Series of Numerical Mathematics **138**, 139-153, 2001.
- [27] M. Hintermüller. A primal-dual active set algorithm for bilaterally control constrained optimal control problems. *Quarterly Appl. Math.* **61**, 131-161, 2003.
- [28] M. Hintermüller, K. Ito, K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM J. Optim.* **13**, 865-888, 2003.
- [29] J. Hull, A. White. The pricing of options on assets with stochastic volatilities. *J. Finance* **42**, 281-300, 1987.
- [30] K. Ito, K. Kunisch. The primal-dual active set method for nonlinear problems with bilateral constraints. *SIAM J. Control Optim.* **43**(1), 357-376, 2004.
- [31] K. Ito, K. Kunisch. Augmented Lagrangian methods for nonsmooth convex optimization in Hilbert spaces. *Nonlin. Anal. TMA* **41**, 591-616, 2000.
- [32] N. Jackson, E. Sli, S. Howison. Computation of deterministic volatility surfaces. *J. Comp. Finance* **2**, 5-32, 1999.
- [33] J.C. Jackwerth. Recovering risk aversion from option prices and realized returns. *Rev. Financ. Studies* **13**, 433-451, 2000.
- [34] J.C. Jackwerth, M. Rubinstein. Recovering probability distributions from contemporary security prices, *J. Finance* **51**, 347-369, 1996.
- [35] P. Kangro, R. Nicolaides. Far field boundary conditions for Black-Scholes equations. *SIAM J. Numer. Anal.* **38**, 1357-1368, 2000.
- [36] I. Kani, E. Derman, M. Kamal. Trading and hedging local volatility. *J. Financ. Engin.* **6**, 1233-1268, 1997.
- [37] O.A. Ladyzenskaja, V.A. Solonnikov, N.N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*, AMS Transl., Vol. 23, Providence, 1968.
- [38] R. Lagnado, S. Osher. A technique for calibrating derivative security pricing models: numerical solution of an inverse problem. *J. Comput. Finance* **1**, 13-25, 1997.
- [39] J.-L. Lions. *Control of Distributed Singular Systems*. Gauthier-Villars, Paris, 1983.
- [40] D.G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, New York, 1969.
- [41] R. C. Merton. Theory of rational option pricing. *Bell J. Econ. Manag. Sci.* **4**, 141-183, 1973.
- [42] K. Schittkowski. On the convergence of a sequential quadratic programming method with an augmented Lagrangian line search function. *Math. Operationsforsch. Stat., Ser. Optimization* **14**, 197-216, 1983.
- [43] J. Simon. Compact sets in the space  $L^p(0, T; B)$ . *Ann. Mat. Pura Appl., IV. Ser.* **146**, 65-96, 1987.

DR. B. DÜRING, INSTITUT FÜR MATHEMATIK, J. GUTENBERG-UNIVERSITÄT MAINZ, STAUDINGER WEG 9, D-55099 MAINZ, GERMANY  
*E-mail address:* [duering@uni-mainz.de](mailto:duering@uni-mainz.de)

UNIV.-PROF. DR. A. JÜNGEL, INSTITUT FÜR MATHEMATIK, J. GUTENBERG-UNIVERSITÄT MAINZ, STAUDINGER WEG 9, D-55099 MAINZ, GERMANY  
*E-mail address:* [juengel@mathematik.uni-mainz.de](mailto:juengel@mathematik.uni-mainz.de)

AO. UNIV.-PROF. DR. S. VOLKWEIN, INSTITUT FÜR MATHEMATIK UND WISSENSCHAFTLICHES RECHNEN, KARL-FRANZENS-UNIVERSITÄT GRAZ, HEINRICHSTRASSE 36, A-8010 GRAZ, AUSTRIA  
*E-mail address:* [stefan.volkwein@uni-graz.at](mailto:stefan.volkwein@uni-graz.at)