

SIMULTANEOUS IDENTIFICATION OF VOLATILITY AND INTEREST RATE FUNCTIONS—A TWO-PARAMETER REGULARIZATION APPROACH*

CHRISTOPHER HOFMANN[†], BERND HOFMANN[†], AND ALOIS PICHLER[†]

Abstract. This paper investigates a specific ill-posed nonlinear inverse problem that arises in financial markets. Precisely, as a benchmark problem in the context of volatility surface calibration, we consider the simultaneous recovery of implied volatility and interest rate functions over a finite time interval from corresponding call- and put-price functions for idealized continuous families of European vanilla options over the same maturity interval. We prove identifiability of the pair of functions to be identified by showing injectivity of the forward operator in L^2 -spaces. To overcome the ill-posedness we employ a two-parameter Tikhonov regularization with heuristic parameter choice rules and demonstrate chances and limitations by means of numerical case studies using synthetic data.

Key words. inverse option pricing, simultaneous identification, volatility, interest rate, regularization

AMS subject classifications. 65J20, 91G60, 47H30, 47A52

1. Introduction. Many authors with focus on analysis and numerics have considered the inverse problem arising in financial markets of recovering local volatility surfaces from option price data in the past years; see, e.g., [1, 2, 3, 4, 10, 11, 12, 13, 15, 23, 34]. In most cases, tools from regularization theory have been incorporated in the treatment of the inverse problem in order to overcome or at least to suppress the occurring ill-posedness phenomena. The associated forward operators are based on solutions to the corresponding partial differential equations named after Black, Scholes, and Dupire (cf. [9, 14]).

It has been observed soon that market prices deviate from computed prices predicted by the models. Initial attempts to connect the prices simply vary the volatility, a single model parameter. These models to calibrate the *implied volatility* σ based on Black-Scholes-type formulas seem at first glance too simplistic to extract significant market information. In the case of varying term structures $\sigma(t)$ and varying maturities $0 \leq t \leq T$, they leave the dependence of required volatility surfaces $\sigma(K, t)$ on the strike price K out of consideration.

To better match observed and computed prices, this paper extends the effort and considers the varying interest in addition. The numerically obtained results demonstrate that implied volatilities and interest rate functions can be reconstructed simultaneously. For this reason the methods outlined below allow reconstructing the fundamental economic parameters of option issuers by reverse engineering. To the best of our knowledge, the simultaneous identification of time-dependent volatility and interest rate functions has not been addressed in the literature so far.

The structure of the forward operators is straightforward if the simplified model of purely time-dependent functions can be exploited, which may serve as a benchmark situation for studying the *nature of ill-posedness* of such *nonlinear inverse problems*. This has been illustrated and worked out comprehensively in the paper [22] to identify the volatility term structure from noisy data of prices $V_c(t)$, $0 \leq t \leq T$, of corresponding European plain vanilla call options. Regularization approaches are required for the stable approximate solution of the term structure of the volatility σ . The subsequent paper [16] exploits the product ansatz $\sigma(K, t) = \sigma_1 \left(K \exp \left(- \int_0^t r(\tau) d\tau \right) \right) \cdot \sigma_2(t)$ with some interest rate r and clearly

*Received July 12, 2018. Accepted March 15, 2019. Published online on April 15, 2019. Recommended by Ronny Ramlau.

[†]Chemnitz University of Technology, Faculty of Mathematics, 09107 Chemnitz
(christopher.hofmann, bernd.hofmann, alois.pichler}@math.tu-chemnitz.de).

demonstrates that studies concerning the purely time-dependent case have a relevant impact on affirmative results to recover volatility surfaces.

The continuous but not necessarily constant term structure function $r(t)$ ($0 \leq t \leq T$) of the interest rate is not negligible. The present paper addresses—as a next step—the *simultaneous* identification of the term structure of $\sigma(t)$ ($0 \leq t \leq T$) and $r(t)$ ($0 \leq t \leq T$) from option price data. Since we thus have to recover a *pair* of continuous functions over a finite time interval, also a pair of associated data functions must be observed. We incorporate the associated put prices $V_p(t)$ ($0 \leq t \leq T$) for this purpose. In this idealized setting, it is assumed for the analysis in infinite-dimensional abstract spaces that data are available for a continuum of call and put prices over the time interval under consideration. Instead of one regularization parameter in [22], we now have to execute two-parameter regularization approaches (cf., e.g., [29, Chapt. 3] and [30]) to handle the ill-posedness phenomena for this new inverse problem.

The volatility and interest rate functions σ and r are apparently present in observed prices. The functions themselves, however, are hidden by the emitter of the derivative. To judge the issuing conditions of the option, it is of fundamental economic interest to understand the occurring parameters and their interplay.

Outline. The remainder of the manuscript is organized as follows: in Section 2 we introduce the relevant functions arising in the context of the purely time-dependent Black-Scholes model of option pricing. The subsequent Section 3 formulates the inverse problem in all its facets and characterizes the occurring forward operators. For the nonlinear inverse problem of recovering volatility and interest rate functions simultaneously from two option price functions, we show in Section 4 the identifiability. The proof of the corresponding Proposition 4.1 is rather straightforward, but only based on this uniqueness result, the subsequent case studies make sense, which are of crucial importance to the paper. For the simultaneous identification problem, the local ill-posedness everywhere will be shown in Section 5. This fact motivates regularization strategies for the stable approximate solution of this recovery process. We mention a classical variant of variational regularization in Section 6 and suggest two-parameter Tikhonov regularization approaches more adapted to the practice in Section 7 motivated by initial case studies.

Section 8 is devoted to more comprehensive numerical tests based on synthetic data concerning heuristic rules for choosing the two regularization parameters. We present case studies for two-dimensional L -curve or L -hypersurface methods, quasi-optimality methods, and a combination of both approaches. We are aware of Bakushinsky's veto (cf. [5]) against heuristic rules, because there are worst case situations that prevent convergence of the regularized solutions of the inverse problem to the exact one. However, in [26] (see also [6]) convergence for the quasi-optimality approach was shown under certain conditions imposed on the noise distribution in the data. On the other hand, as for example emphasized in [19], heuristic rules like the L -curve method lead often to better results than some sophisticated noise-level dependent rules which possess formally order-optimal convergence properties.

The appendix outlines the background of the Black-Scholes formula with integrated time-dependent interest rates and time-dependent volatility.

2. Preliminaries. Throughout this paper we consider a variant of the Black-Scholes model derived from the stochastic differential equation

$$(2.1) \quad \frac{dX(t)}{X(t)} = \mu dt + \sigma(t) dW(t) \quad (0 \leq t \leq T)$$

for the time-dependent price $X(t)$ of an asset under consideration with constant drift μ and standard Wiener process W . At the initial time $t = 0$, let there exist two idealized families of

European vanilla call and put options at the financial market with observable prices written on the asset with current asset price $X := X(0) > 0$ and fixed strike $K > 0$. For simplicity, we use $T := 1$ for the maturity interval $0 \leq t \leq T$, which means that all real functions under consideration are defined on the unit interval.

For the implied volatility term structure $\sigma(t)$ ($0 \leq t \leq 1$), we introduce the auxiliary functions $a(t) := \sigma^2(t)$ ($0 \leq t \leq 1$) and

$$(2.2) \quad A(t) := \int_0^t a(\tau) d\tau \quad (0 \leq t \leq 1).$$

Moreover, we consider a *maturity-dependent interest rate* $r(t)$ ($0 \leq t \leq 1$) and the corresponding auxiliary function

$$(2.3) \quad R(t) := \int_0^t r(\tau) d\tau \quad (0 \leq t \leq 1).$$

The *Black-Scholes function* with variables $X > 0$, $K > 0$, $\varrho \geq 0$, $s \geq 0$, and

$$d_{\pm}(X, K, \varrho, s) := \frac{\varrho \pm \frac{s}{2} + \ln\left(\frac{X}{K}\right)}{\sqrt{s}} \quad (\text{i.e., } d_- = d_+ - \sqrt{s})$$

is

$$(2.4) \quad U_{BS}(X, K, \varrho, s) := \begin{cases} X\Phi(d_+(X, K, \varrho, s)) - Ke^{-\varrho}\Phi(d_-(X, K, \varrho, s)) & \text{if } s > 0, \\ \max(X - Ke^{-\varrho}, 0) & \text{if } s = 0, \end{cases}$$

where we denote the cumulative distribution function of the standard normal distribution by

$$\Phi(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{x^2}{2}} dx.$$

REMARK 2.1. It is essential to note that formula (2.4) is stated in terms of ρ and s ; the usual Black-Scholes formula for constant r and σ is obtained by the setting $\rho = r \cdot t$ and $s = \sigma^2 \cdot t$.

The following lemma (see also [22, Lemma 2.1]), which can be easily verified from (2.4), makes assertions on continuity and monotonicity properties of the Black-Scholes function with respect to both the variables ρ and s .

LEMMA 2.2. *Let the parameters $X > 0$ and $K > 0$ be fixed. Then the nonnegative function $U_{BS}(X, K, \rho, s) \leq X$ is continuous for $(\rho, s) \in [0, \infty) \times [0, \infty)$. Moreover, for $(\rho, s) \in [0, \infty) \times (0, \infty)$, this function is continuously differentiable with respect to ρ , where we have with $\nu := \ln\left(\frac{X}{K}\right)$*

$$\frac{\partial U_{BS}(X, K, \rho, s)}{\partial \rho} = \exp(-\rho) K \Phi\left(\frac{s - 2\rho - 2\nu}{2\sqrt{s}}\right) < K,$$

and it is continuously differentiable with respect to s , where we have

$$(2.5) \quad \frac{\partial U_{BS}(X, K, \rho, s)}{\partial s} = \frac{X}{\sqrt{8\pi s}} \exp\left(-\frac{[\nu + \rho]^2}{2s} - \frac{[\nu + \rho]}{2} - \frac{s}{8}\right) > 0.$$

3. Modelling of the inverse problem and characterization of the forward operators.

The following lemma outlines the fair call price function $V_c(t)$ ($0 \leq t \leq 1$) and the price function $V_p(t)$ ($0 \leq t \leq 1$) of the associated put by taking into account the put-call parity for European options; cf., e.g., [25, Chap. 9.4]. An economic motivation of the formula and a sketch of a proof of this lemma can be found in the appendix.

LEMMA 3.1. *By using the Black-Scholes-function U_{BS} from (2.4) we have*

$$(3.1) \quad V_c(t) = U_{BS}(X, K, R(t), A(t)) \quad (0 \leq t \leq 1)$$

and

$$(3.2) \quad V_p(t) = V_c(t) - X + Ke^{-R(t)} \quad (0 \leq t \leq 1).$$

In what follows we use infinite-dimensional Hilbert spaces to model our problem, in particular, $L^2(0, 1)$ with the usual norm and $H := L^2(0, 1) \times L^2(0, 1)$ with the norm $\|(u, v)\|_H := \left(\|u\|_{L^2(0,1)}^2 + \|v\|_{L^2(0,1)}^2 \right)^{1/2}$, although the functions a and r to be recovered and the option price functions are at least continuous in practice. The reasons for doing so are twofold:

1. Instead of the continuous fair option price functions $V_c(t)$ and $V_p(t)$ defined on the maturity interval $[0, 1]$, there are only noisy observations $V_c^\delta(t)$ and $V_p^\delta(t)$ available, for example by bid-ask price gaps, which satisfy the deterministic *noise model*

$$(3.3) \quad \|(V_c, V_p) - (V_c^\delta, V_p^\delta)\|_H = \left(\|V_c - V_c^\delta\|_{L^2(0,1)}^2 + \|V_p - V_p^\delta\|_{L^2(0,1)}^2 \right)^{1/2} \leq \delta$$

with some noise level $\delta \geq 0$. The observed noisy option prices tend to be ‘unsmooth’, incorporating some random effects; the class of quadratically integrable functions is able to model such effects.

2. The simultaneous recovery of volatility and interest rate functions can be associated with the nonlinear forward operator

$$(3.4) \quad F: (a, r) \in H \mapsto (V_c, V_p) \in H,$$

possessing the domain

$$(3.5) \quad \mathcal{D}(F) := \{(a, r) \in H : a(t) \geq \underline{a} > 0, r(t) \geq 0 \text{ a.e.}\},$$

such that solving the inverse problem corresponds to the solution of the nonlinear operator equation

$$(3.6) \quad F((a, r)) = (V_c, V_p), \quad (a, r) \in \mathcal{D}(F) \subset H, (V_c, V_p) \in H.$$

As we will see, the general theory of nonlinear variational regularization in the sense of [17, 32, 33] applies immediately to that operator equation (3.6).

Following the ideas in [22] we can decompose the *total inverse problem (TOP)* modeled by the nonlinear operator equation (3.6) in the Hilbert space H into an

Outer inverse problem (OIP): find the pair $(A, R) \in H$ of antiderivatives to $(a, r) \in H$ from noisy data $(V_c^\delta, V_p^\delta) \in H$ satisfying the noise model (3.3)

and an

Inner inverse problem (IIP): find the pair $(a, r) \in H$ of derivatives to A and R , respectively, from noisy data of A and R .

The forward operator of (OIP)

$$N: (A, R) \mapsto (V_c, V_p)$$

is nonlinear and of Nemytskii-type, which means here that $V_c(t)$ as well as $V_p(t)$ only depend on the two numbers $A(t)$ and $R(t)$ for the same value $t \in [0, 1]$. Based on the compact linear integration operator $J: L^2(0, 1) \rightarrow L^2(0, 1)$, defined as

$$[Ju](t) := \int_0^t u(\tau) d\tau \quad (0 \leq t \leq 1),$$

the forward operator L of the inverse problem (IIP) mapping as

$$L: (a, r) \mapsto (Ja, Jr)$$

is also linear, where the mildly ill-posed differentiation problem, i.e., the inversion of J , is well investigated from analytical and numerical points of view; cf., e.g., [20]. Note that we have to consider the constraints $L: \mathcal{D}(F) \subset H \rightarrow H$ imposed by the domain of F . The studies in [27] (see also [28]) concerning Nemytskii operators indicate that outer problems like (OIP) are not ill-posed but extremely ill-conditioned, in particular with focus on small maturities t . Taking into account the composition structure

$$F = N \circ L,$$

however, the ill-posedness carries over from the inner problem to the total inverse problem; see Section 4 below. Hence, regularization is required for the stable approximate simultaneous recovery of a and r . On the other hand, it is not a priori clear whether it is really advantageous to solve (TIP) not directly by a regularization approach but successively by finding regularized solutions to (OIP) in a first step and then to solve (IIP) in a second step.

PROPOSITION 3.2. *The Nemytskii-type forward operator N of (OIP) is continuous in H , and hence the forward operator F of (TIP) from (3.4) with domain (3.5) maps weak-to-weak continuously in H .*

Proof. To show the continuity of N it is enough to show continuity of the mapping $(A, R) \in \mathcal{D}(F) \subset H \mapsto V_c \in L^2(0, 1)$ based on properties of the Black-Scholes function U_{BS} characterized by Lemma 2.2, since the mapping $(A, R) \in \mathcal{D}(F) \subset H \mapsto V_p \in L^2(0, 1)$ has analog properties. Along the lines of [18] we therefore have to ensure that the Carathéodory condition and the growth condition for the generator function $g(t, \rho, s) = U_{BS}(X, K, \rho, s)$ of the Nemytskii operator $(A(t), R(t)) \mapsto [V_c(A, R)](t) = g(t, R(t), A(t))$ ($0 \leq t \leq 1$) are valid. On the one hand, the smoothness requirements on g for the Carathéodory condition are satisfied due to the assertions of Lemma 2.2 when taking into account that only nonnegative ρ and s are under consideration due to the specific domain (3.5) of F . On the other hand, the growth condition is trivially satisfied due to $U_{BS}(X, K, \rho, s) \leq X$. This proves the continuity of N . Then the weak-to-weak continuity of $F = N \circ L$ is an immediate consequence of the continuity of N , and the weak-to-norm continuity of L is an implication of the compactness of the J -mapping in $L^2(0, 1)$. This completes the proof. \square

4. Simultaneous identifiability of volatility and interest rate. In this section, we show that the nonlinear operator F is *injective* on $\mathcal{D}(F)$, which means that the pair $(a, r) \in \mathcal{D}(F)$ of term structures of volatility and interest rate is *simultaneously identifiable* in a unique manner from fair prices of the term structure (V_c, V_p) of associated call and put options.

PROPOSITION 4.1. *If there is a pair $(a, r) \in \mathcal{D}(F)$ with associated auxiliary functions (A, R) according to (2.2) and (2.3) such that $V_c(t)$ ($0 \leq t \leq 1$) and $V_p(t)$ ($0 \leq t \leq 1$) satisfy the formulas (3.1) and (3.2), then this pair (a, r) is uniquely determined in $\mathcal{D}(F)$.*

Proof. For the pair of fair option price functions (V_c, V_p) under consideration we can conclude as follows: we have from $r(t) \geq 0$ almost everywhere on $[0, 1]$ that, with (2.3), $R(t) \geq 0$ and consequently from formula (3.2) that $X + V_p(t) - V_c(t) \geq 0$ for all $0 < t \leq 1$. This implies the uniqueness of $R(t)$ in formula (3.2) for all $(0 \leq t \leq 1)$ and hence due to (2.3) the uniqueness of $r(t)$ for all $0 < t \leq 1$. Since $R(t)$ ($0 \leq t \leq 1$) is well determined now, we can prove the uniqueness of $A(t)$ ($0 < t \leq 1$) from formula (3.1) taking into account that $a(t) \geq \underline{a} > 0$ almost everywhere on $[0, 1]$ and hence $A(t) > 0$ for all $0 < t \leq 1$. By formula (2.5) from Lemma 2.2 we get $\frac{\partial}{\partial s} U_{BS}(X, K, \rho, s) > 0$ for all $0 < s \leq 1$. This, however, implies for all $0 < t \leq 1$ the strict growth of the function $U_{BS}(X, K, R(t), s)$ with respect to $s > 0$ for fixed parameters X, S , and $R(t)$. Consequently, the function $A(t)$ ($0 \leq t \leq 1$) and hence by (2.2) also the function $a(t)$ ($0 < t \leq 1$) are uniquely determined. This completes the proof of the proposition. \square

5. Local ill-posedness of the total inverse problem (TIP). We refer to [24, Definition 3] for the concept of local ill-posedness for a nonlinear operator equation with forward operator F .

PROPOSITION 5.1. *The total inverse problem (TIP) is locally ill-posed everywhere on $\mathcal{D}(F)$, which means that in any ball in H around $(a, r) \in \mathcal{D}(F)$ with positive radius there is a sequence $\{(a_n, r_n)\}_{n \in \mathbb{N}} \subset \mathcal{D}(F)$ such that $\liminf_{n \rightarrow \infty} \|(a_n, r_n) - (a, r)\|_H > 0$ but $\lim_{n \rightarrow \infty} \|F(a_n, r_n) - F(a, r)\|_H = 0$.*

Proof. Evidently, for all points $(a, r) \in \mathcal{D}(F)$ and balls around with arbitrarily small radii, we have a sequence $\{(a_n, r_n)\}_{n \in \mathbb{N}} \subset \mathcal{D}(F)$ in such a ball that converges weakly to (a, r) but not in the norm of H . Then the sequence $\{(Ja_n, Jr_n)\}_{n \in \mathbb{N}}$ is norm-convergent in H to the element (Ja, Jr) as a consequence of the compactness of J . Moreover, by Proposition 3.2, we have that $F(a_n, r_n)$ is norm-convergent in H to $F(a, r)$ because N is a continuous nonlinear operator mapping in H . This shows the local ill-posedness everywhere. \square

Due to the ill-posedness of the inverse problem, some regularization approach is required for obtaining stable approximate solutions to the uniquely determined pair $(a, r) \in \mathcal{D}(F)$ from noisy data $(V_c^\delta, V_p^\delta) \in H$.

6. A theoretical variant of variational regularization. Following the classical approach of variational regularization we can search for regularized solutions $(a_\alpha^\delta, r_\alpha^\delta) \in \mathcal{D}(F)$ as minimizers of the extremal problem

$$(6.1) \quad T_\alpha^\delta((\tilde{a}, \tilde{r})) \rightarrow \min, \quad \text{subject to } (\tilde{a}, \tilde{r}) \in \mathcal{D}(F) \subset H,$$

for the Tikhonov functional

$$T_\alpha^\delta((\tilde{a}, \tilde{r})) := \|F((\tilde{a}, \tilde{r})) - (V_c^\delta, V_p^\delta)\|_H^2 + \alpha \|(\tilde{a}, \tilde{r}) - (a^*, r^*)\|_H^2,$$

where $\alpha > 0$ is the regularization parameter and a^* and r^* are prescribed initial guesses of the functions a and r to be recovered. Due to the *weak-to-weak continuity* of F (cf. Proposition 3.2) in combination with the closedness and convexity of $\mathcal{D}(F)$, we thus obtain convergence of $(a_\alpha^\delta, r_\alpha^\delta)$ to the uniquely determined exact solution (a, r) of problem (TIP) in the norm of H whenever the regularization parameter $\alpha > 0$ is chosen such that $\alpha \rightarrow 0$ and $\delta^2/\alpha \rightarrow 0$ as $\delta \rightarrow 0$; cf., e.g., [17, Chap. 10]. This condition can be executed by an a priori parameter choice or by using a variant of the discrepancy principle.

7. Two-parameter regularization approaches more adapted to the practice. Since we have to identify two independent functions a and r it seems to be more adapted to practical requirements to employ a two-parameter regularization approach with two convex and stabilizing penalties $\Omega_1(a)$ and $\Omega_2(r)$, which express subjective a priori information about

a and r , separately. Then the regularized solutions $(a_{\alpha,\beta}^\delta, r_{\alpha,\beta}^\delta) \in \mathcal{D}(F)$ are minimizers of the extremal problem (6.1) but with the amended Tikhonov functional

$$(7.1) \quad T_\alpha^\delta((\tilde{a}, \tilde{r})) := \|F((\tilde{a}, \tilde{r})) - (V_c^\delta, V_p^\delta)\|_H^2 + \alpha \Omega_1(\tilde{a}) + \beta \Omega_2(\tilde{r})$$

and two positive regularization parameters α and β .

Using a discretization with an equidistant grid over the unit interval $[0, 1]$ with 50 grid points we investigate three test examples for an asset with present price $X = 100$ at $t = 0$ and a family of options with fixed strike price $K = 95$:

EXAMPLE 7.1. For this example we consider a very smooth implied volatility function

$$a(t) := (t - 0.5)^2 + 0.01 \quad (0 \leq t \leq 1),$$

which is decreasing on the subinterval $[0, 1/2]$ and increasing on $[1/2, 1]$, and the slowly growing interest rate function

$$r(t) := \frac{1}{70} \cdot \ln(50t + 1) + 0.02 \quad (0 \leq t \leq 1).$$

EXAMPLE 7.2. In order to study the impact of a lower degree of smoothness in the volatility function, we choose for this example the term structure

$$a(t) := 0.4 \left(0.5 + \frac{0.9}{1 + 100(2.05t - 0.2)^2} \right) \quad (0 \leq t \leq 1)$$

with a peak near $t = 0.1$ and the same interest rate

$$r(t) := \frac{1}{70} \cdot \ln(50t + 1) + 0.02 \quad (0 \leq t \leq 1)$$

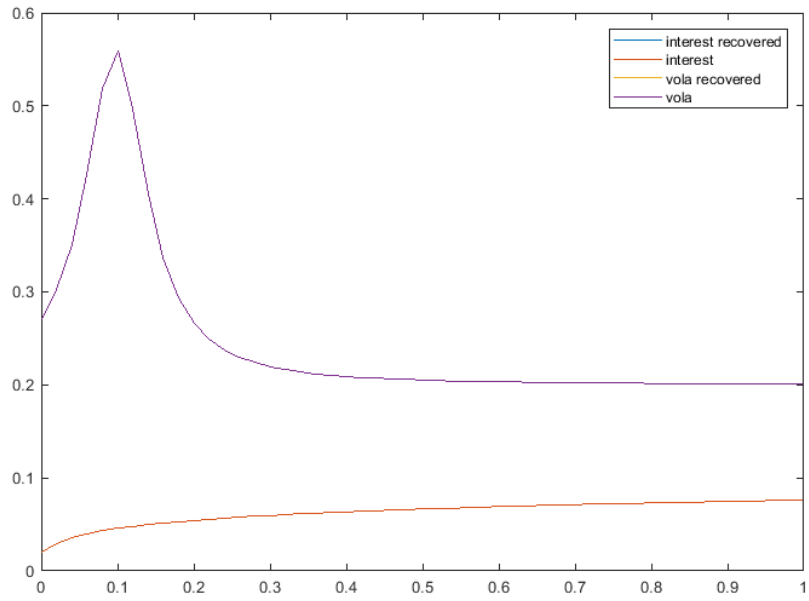
as in Example 7.1.

EXAMPLE 7.3. In a third test case, we repeat the peak structure

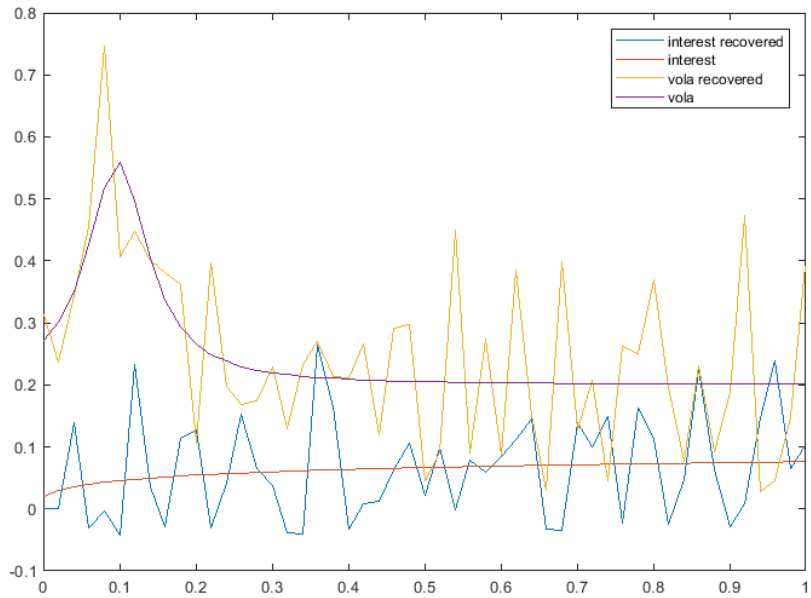
$$a(t) := 0.4 \left(0.5 + \frac{0.9}{1 + 100(2.05t - 0.2)^2} \right) \quad (0 \leq t \leq 1)$$

of Example 7.2 for the volatility function but slightly amend the interest rate to

$$r(t) := \begin{cases} 0.03 + 0.02 \sin(4\pi t) & \text{if } 0 \leq t \leq 0.5 \\ 0.03 + \frac{1}{70} \ln(50(t - 0.05) + 1) & \text{if } 0.5 < t \leq 1 \end{cases} \quad (0 \leq t \leq 1).$$

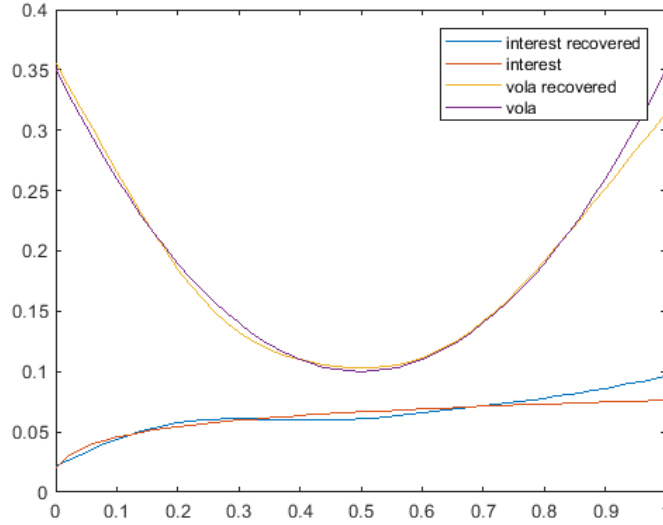


(a) $\delta = 0$.

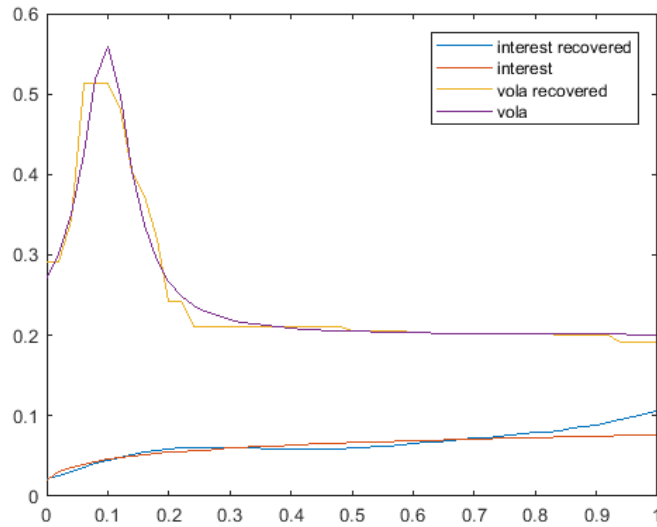


(b) $\delta = 0.01$.

FIG. 7.1. Exact functions a and r and least-squares solutions (unregularized) for Example 7.2.



(a) Example 7.1 with $\alpha = 2, \beta = 3$.



(b) Example 7.2 with $\alpha = 1.5, \beta = 2$.

FIG. 7.2. Regularized solutions $a_{\alpha,\beta}^\delta$ and $r_{\alpha,\beta}^\delta$ versus exact a and r for both examples and appropriately chosen regularization parameters $\alpha > 0, \beta > 0, \delta = 0.01$ and appropriately chosen $\Omega_1(a)$ and $\Omega_2(r)$.

In the numerical case studies for these three examples, with a noise level δ obeying (3.3), a Gaussian additive noise model is used and implemented based on the formula

$$[\underline{V}_c^\delta \mid \underline{V}_p^\delta] = [\underline{V}_c \mid \underline{V}_p] + \frac{\mathcal{E}}{\|\mathcal{E}\|_F} \cdot \|\underline{V}_c \mid \underline{V}_p\|_F \cdot \delta,$$

where the underlining denotes that discretized versions (n -dimensional column vectors) are used instead of the original continuous functions. The symbol $\underline{\mathcal{E}} \in \mathbb{R}^{n \times 2}$ denotes the noise matrix with Gaussian random i.i.d. entries $\epsilon_{ij} \sim \mathcal{N}(0, 1)$, and $\|\cdot\|_F$ designates the Frobenius norm.

To illustrate the ill-posedness (instability) phenomenon of the inverse problem (TIP) we first compare the different behavior of approximate solutions between the recovery with noiseless data ($\delta = 0$) and noisy data ($\delta = 0.01$). This is shown in Figure 7.1 for Example 7.2. No regularization in the least-squares approach was employed, i.e., $\alpha = \beta = 0$. For both functions a and r , highly oscillating unregularized solutions are observed if the data are contaminated with noise.

This motivates in turn to employ variational regularization (6.1) with the functional (7.1) to be minimized. Figure 7.2 shows that for both Examples 7.1 and 7.2 there are two positive parameters α and β for which the regularized solutions $a_{\alpha,\beta}^\delta$ and $r_{\alpha,\beta}^\delta$ are good approximations to the exact volatilities a and interest rates r if appropriate penalty terms are chosen, even if the data are contaminated with noise. In this case discretized versions of the second derivative norm square $\Omega_2(r) = \|r''\|_{L^2(0,1)}^2$ as well as $\Omega_1(a) = \|a''\|_{L^2(0,1)}^2$ in Example 7.1 and the bounded variation norm $\Omega_1(a) = \|a\|_{BV[0,1]}$ in Example 7.2 are employed. The noise level is set to $\delta = 0.01$ in both cases. For the recovery of very smooth functions it makes sense to use penalties based on the norm square of the second derivative, whereas total variation penalties proved to be appropriate if some variants of non-smoothness like discontinuities or peaks occur. Although these two figures appear very promising, the question how to choose the pair (α, β) of positive regularization parameters remains open. As the error level δ is unknown in practice, we rely on heuristic parameter choice rules. In the following section, two such rules are suggested and numerically tested based on the Examples 7.1, 7.2, and 7.3 introduced above.

8. Heuristic rules for two-parameter regularization.

8.1. Parameter choice based on L-hypersurfaces. The L-hypersurface or generalized L-curve framework is an extension of the conventional L-curve. The concept was introduced by Belge et. al. ([8, 7]), but their approach and applications are restricted to linear operators. In the following, we apply this approach to the given nonlinear problem. The conventional L-curve method (cf. [21, 31]) displays the trade-off between the fit to the given (noisy) data and the size of the penalty term for various regularization parameters on a log-log scale. This plot often results in the typical L-curve shape, and the ‘optimal’ regularization parameter can be chosen at the point with the highest curvature. The L-hypersurface generalizes this concept—in our case to three dimensions. The penalty terms are plotted against the residue on a log-log-log scale for various regularization parameters. In the next step the point with the highest Gaussian curvature is determined, and the corresponding regularization parameters are chosen. The Gaussian curvature can be visualized in the corresponding curvature surface, which displays the curvature depending on the regularization parameters on a log-log-log scale. The respective regularization parameters can then be easily determined from this plot.

In what follows, we test the performance of such an approach for Example 7.1 with discretized versions of $\Omega_1(a) = \|a''\|_{L^2(0,1)}^2$, $\Omega_2(r) = \|r''\|_{L^2(0,1)}^2$, and for a noise level $\delta = 0.05$. For appropriate regularization parameters we thus obtain the L-hypersurface of Figure 8.1. It visualizes the trade-off between the penalty terms Ω_1 and Ω_2 and the residue or fit to the data. It is expected that the point with the highest Gaussian curvature is a suitable compromise between these variables. Therefore, the point with the highest curvature is selected and the corresponding regularization parameters α and β are chosen. In a first step, this can be done by evaluating the plot and choosing the parameters or points in question. This method is

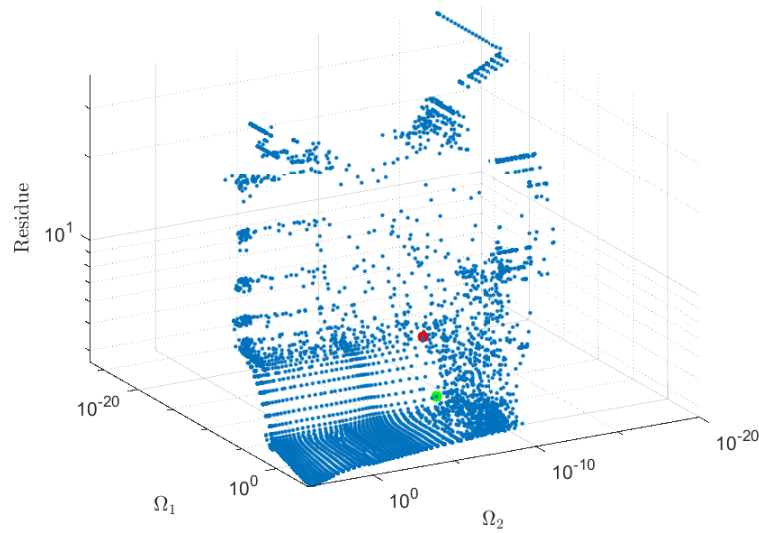


FIG. 8.1. *Generalized L-hypersurface with selected points.*

illustrated by investigating the regularized solutions $a_{\alpha,\beta}^\delta$ and $r_{\alpha,\beta}^\delta$ for selected points from the previously introduced L-hypersurface; see Figure 8.2.

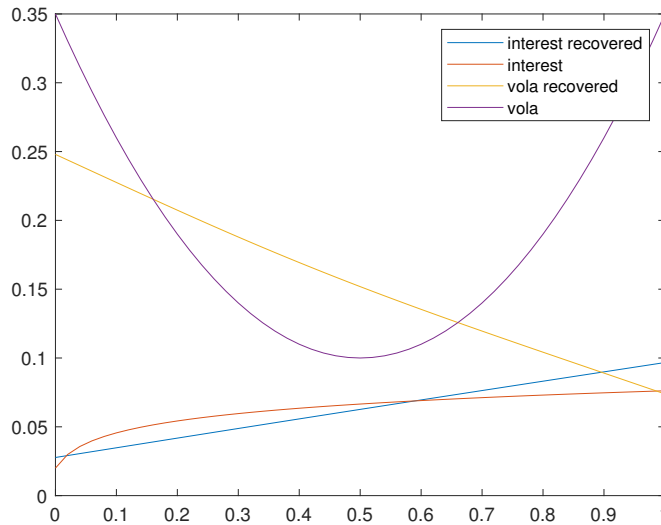
Further numerical experiments show that regularized solutions are very stable with respect to varying β but very sensitive with respect to varying α .

In what follows, instead of ‘guessing’ the curvature from the L-hypersurface, we calculate and plot the Gaussian curvature against the penalty terms or regularization parameters and receive the corresponding curvature surface. Figure 8.3 displays the curvature surfaces for selected $\alpha, \beta \in [10^{-2.75}, 10^{2.25}]$. The plot on the bottom displays the Gaussian curvature against the regularization parameters. Both surfaces display a very pronounced (and unique) maximum. The maximum Gaussian curvature and the appropriate regularization parameters $\alpha = 3.162$ and $\beta = 1778$ can therefore be conveniently determined from this curvature surface.

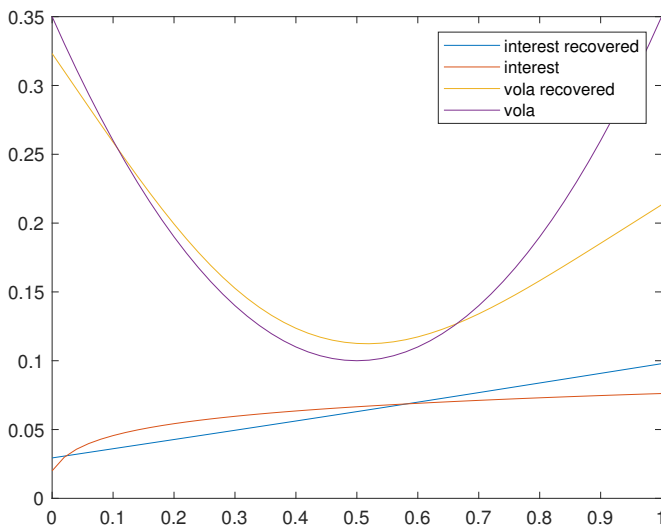
If we look at the regularized solutions $a_{\alpha,\beta}^\delta$ and $r_{\alpha,\beta}^\delta$ for this parameter choice and the exact solutions a and r , the results appear very promising (Figure 8.4), especially considering the high error level of 5%.

Please note that extensive numerical experiments have been conducted in which the L-hypersurface retains always its distinctive shape for various error levels, penalty terms, interest, and volatility functions as well as combinations of strike and underlying.

8.2. Adaption of quasi-optimality and L-curve criteria. The following approach is motivated by the method of Naumova and Pereverzyev [30]. These authors consider a linear operator between Hilbert spaces and deal with the minimization of a Tikhonov functional with two penalty terms and corresponding regularization parameters. The regularization parameters are chosen individually, one after the other, using the quasi-optimality criterion. From the numerical experiments in the previous section we also concluded that it is ‘easier’ to chose the parameter α than β , as α is more stable in a certain sense than β . Extensive numerical experiments with various heuristic parameter choice rules have also shown that a combination



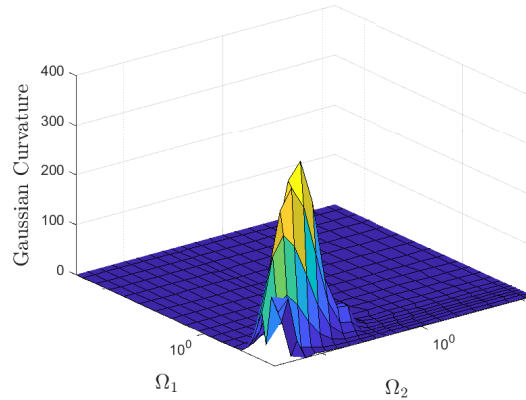
(a) Green point, $a_{\alpha,\beta}^\delta$ and $r_{\alpha,\beta}^\delta$, $\alpha = 56$, $\beta = 1778$



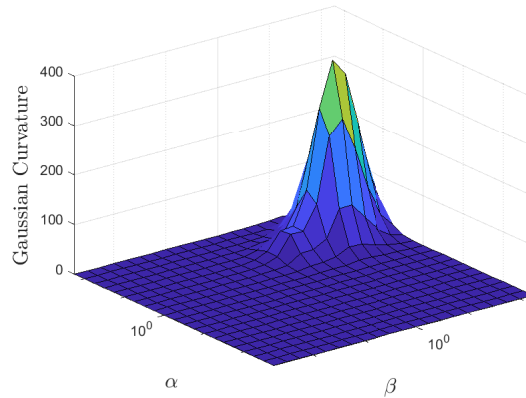
(b) Red point, $a_{\alpha,\beta}^\delta$ and $r_{\alpha,\beta}^\delta$, $\alpha = 31623$, $\beta = 316$

FIG. 8.2. Regularized solutions $a_{\alpha,\beta}^\delta$, $r_{\alpha,\beta}^\delta$ and exact solutions a and r for selected α , β ; regularization parameters chosen from the L-hypersurface; 5% additive Gaussian noise.

of Hansen's L-curve technique in the first stage and quasi-optimality in the second has so far been the most successful. With this idea in mind, the following parameter choice strategy can be employed to find suitable regularization parameters for the Tikhonov functional (7.1). Initially, choose accessible (finite) sets of regularization parameters $P_M^\beta = \{\beta_1, \beta_2, \dots, \beta_N\}$ and $Q_N^\alpha = \{\alpha = \alpha_i = \alpha_0 p^j : j = 0, 1, \dots, M\}$ for $p > 0$ and some β_0 . Suitable $\alpha = \alpha(\beta_j)$ are chosen for all $\beta_j \in P_M^\beta$ using the L-curve rule.



(a) Curvature surface depending on penalty terms.



(b) Curvature surface depending on regularization parameters.

FIG. 8.3. Curvature surface depending on penalty terms and regularization parameters for $\alpha, \beta \in [10^{-2.75}, 10^{2.25}]$.

In the second step, the quasi-optimality criterion is applied to determine $\beta = \beta(\alpha_k)$. Note that in the following all norms are discretized L^2 -norms. When employing the quasi-optimality rule one needs to decide whether to apply the criterion to the series $\{x_{\beta(\alpha_j), \alpha_j}^\delta\}$ or $\{r_{\beta(\alpha_j), \alpha_j}^\delta\}$ with $x = \begin{pmatrix} a \\ r \end{pmatrix}$, which means either determining

$$\|x_{\alpha_j, \beta(\alpha_k)}^\delta - x_{\alpha_j, \beta(\alpha_{k-1})}^\delta\| = \min\{\|x_{\alpha_j, \beta(\alpha_j)}^\delta - x_{\alpha_j, \beta(\alpha_{j-1})}^\delta\| : j = 1, 2, \dots, N\}.$$

or

$$\|r_{\alpha_j, \beta(\alpha_k)}^\delta - r_{\alpha_j, \beta(\alpha_{k-1})}^\delta\| = \min\{\|r_{\alpha_j, \beta(\alpha_j)}^\delta - r_{\alpha_j, \beta(\alpha_{j-1})}^\delta\| : j = 1, 2, \dots, N\}.$$

Subsequently we evaluate this method for Example 7.1. In the first step we apply the L-curve criterion in order to obtain some $\alpha(\beta_j)$ for suitable $\beta_j \in P_M^\beta$. Due to numerical limitations we rely on precalculated values and $P_M^\beta = \{10^k : k = -2, -2.25, -2.5, \dots, 2.75, 3\}$. When plotting the sum of the penalty terms $\Omega = \Omega_1(a) + \Omega_2(r)$ against the residue on a

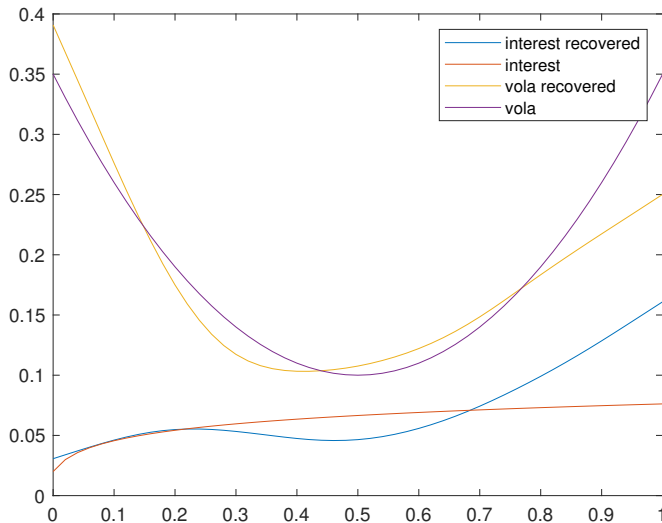


FIG. 8.4. Regularized solutions $a_{\alpha,\beta}^\delta$, $r_{\alpha,\beta}^\delta$ and exact solutions a and r for $\alpha = 10$, $\beta = 17.78$; regularization parameters chosen from the curvature surface in Figure 8.3; 5% additive Gaussian noise.

log-log scale, the L-curve we obtain has the distinct shape for all β_j as in Figure 8.5. It is therefore sensible to evaluate the curvature and chose $\alpha(\beta_j)$ accordingly, i.e., at the point with the highest curvature. In this example we therefore receive $\alpha_j = \alpha(\beta_j) = 17.78$ for all $\beta_j \in P_M^\beta$. In the next step we estimate $\beta(\alpha_k)$. The quasi-optimality criterion is applied to $\{x_{\alpha_j,\beta(\alpha_j)}^\delta\}$ and $\{r_{\alpha_j,\beta(\alpha_j)}^\delta\}$. Figure 8.6 displays $\|x_{\alpha_j,\beta(\alpha_k)}^\delta - x_{\alpha_j,\beta(\alpha_{k-1})}^\delta\|$ (left) and $\|r_{\alpha_j,\beta(\alpha_k)}^\delta - r_{\alpha_j,\beta(\alpha_{k-1})}^\delta\|$ (right) plotted against P_M^β . Recall that $Q_N^\alpha = \{\alpha = \alpha_i = \alpha_0 q^i : i = 0, 1, \dots, N\}$ for $q > 0$ and some α_0 and q need to be chosen accordingly. In this example, $\alpha_0 = 100$ and $q = 0.8$.

Both functions have a (local) minimum at $\beta = 10.74$. Therefore it makes sense to determine the regularized solutions for these parameters and compare it to the exact solution. Figure 8.7 illustrates the regularized solutions $a_{\alpha,\beta}^\delta$ and the exact solutions r and a for $\alpha = 17.78$ and $\beta = 10.74$. Although it appears that $r_{\alpha,\beta}^\delta$ is severely underregularized, this method is promising considering the high error level of $\delta = 0.05$.

Further studies have been conducted in order to confirm these results. In the subsequent example, an error level of $\delta = 0.01$ is used, and we remain in the setting of Example 7.3. Similar to the previous experiment, we again apply the L-curve criterion for $\beta_j \in P_M^\beta$ to obtain $\alpha(\beta_j)$. In this case $P_M^\beta = \{10^k : k = -5, -4.5, -5, \dots, 1.5, 2\}$ was used. The L-curve has again the very distinctive shape observed in the previous example, and we determine $\alpha(\beta_j) = 0.0562$ for all $\beta_k \in P_M^\beta$. Similar to the previous example, Figure 8.8 displays $\|x_{\alpha_j,\beta(\alpha_k)}^\delta - x_{\alpha_j,\beta(\alpha_{k-1})}^\delta\|$ (left) and $\|r_{\alpha_j,\beta(\alpha_k)}^\delta - r_{\alpha_j,\beta(\alpha_{k-1})}^\delta\|$ (right) plotted against P_M^β . Note that the minima of $\|x_{\alpha_j,\beta(\alpha_k)}^\delta - x_{\alpha_j,\beta(\alpha_{k-1})}^\delta\|$ and $\|r_{\alpha_j,\beta(\alpha_k)}^\delta - r_{\alpha_j,\beta(\alpha_{k-1})}^\delta\|$ are again in the same location, and therefore we receive $\beta = 0.859$ in either case.

Figure 8.9 illustrates the exact and regularized solutions for these parameter choices. We note here that the quasi-optimality criterion is very robust for different q in this situation.

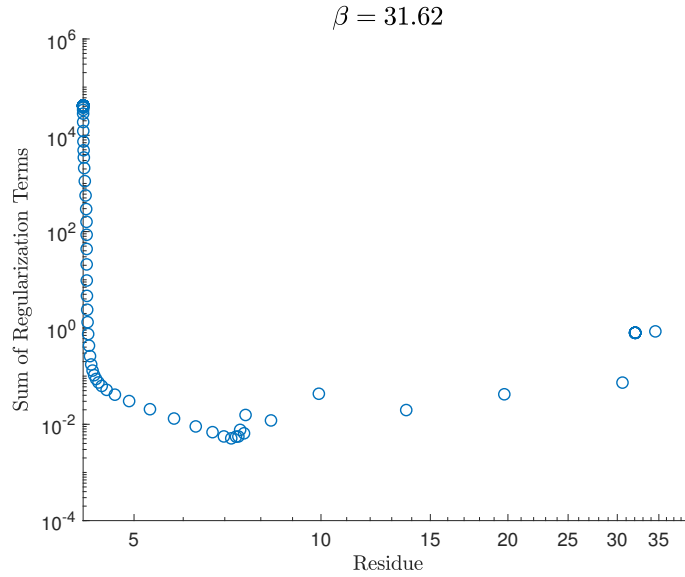


FIG. 8.5. *L*-curve for some fixed $\beta_j=31.62$.

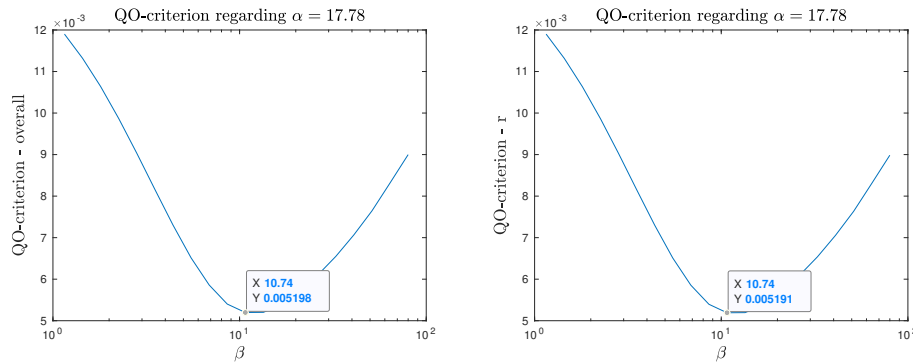


FIG. 8.6. *Quasi-optimality criterion for x and r, for suitable beta.*

8.3. Conclusions. The goal of the present article is to close a certain gap in the treatment of the inverse option pricing problem, namely to identify time-dependent volatility and interest rate functions simultaneously. We investigate ill-posedness and regularization approaches for that specific nonlinear inverse problem with financial background. The injectivity of the forward operator and therefore the identifiability of the function pair (a, r) from data of an associated option price function pair is proven. Variants of Tikhonov regularization using two separate penalty terms are rather successfully executed. In this context, existing heuristic parameter choice rules are adapted and generalized for this particular problem with two regularization parameters.

Numerical case studies for synthetic data indicate that it is possible to recover simultaneously time-dependent interest rate and volatility functions in a stable manner from noisy data. We are also very confident that modifications of these approaches can be applied to real market data and that American options can be incorporated, too.

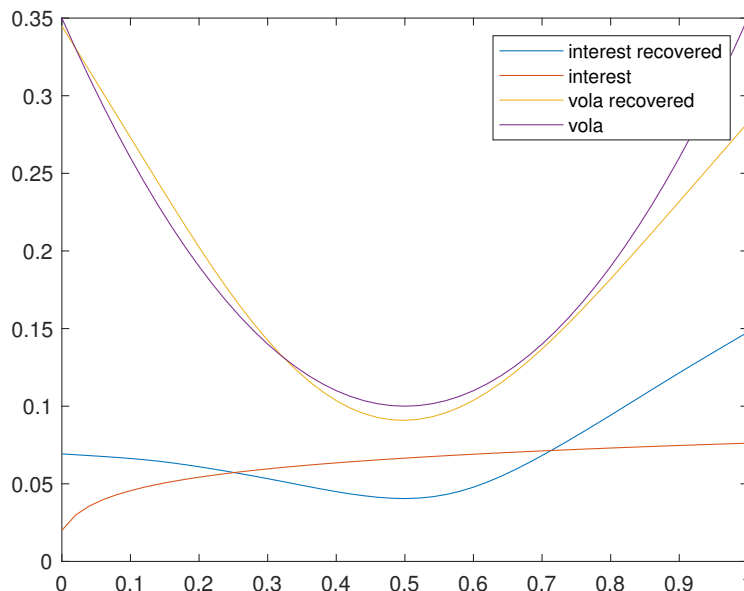


FIG. 8.7. Regularized solutions $r_{\alpha,\beta}^\delta$ and $a_{\alpha,\beta}^\delta$ with exact solutions a and r for $\alpha = 17.78$ and $\beta = 10.74$; regularization parameters chosen from a combination of L-curve and quasi-optimality; 5% additive Gaussian noise.

Acknowledgment. The research of the first two authors is supported by Deutsche Forschungsgemeinschaft (DFG-grant HO1454/12-1) under the auspices of the joint Austrian-German project “Novel Error Measures and Source Conditions of Regularization Methods for Inverse Problems (SCIP)” with the University of Vienna (PI: Prof. Dr. Otmar Scherzer) according to D-A-CH Lead Agency Agreement.

Appendix A. To accept Lemma 3.1 for *time-varying* functions $\sigma^2(\cdot)$ and $r(\cdot)$ with antiderivatives A and R , respectively, we consider a smooth function $v(t, X)$ and the geometric Brownian motion $X(t)$ given by (2.1). Employing Itô’s rule gives

$$\begin{aligned}
 dv(t, X) &= v_t dt + v_X dX + \frac{1}{2} v_{XX} (dX)^2 \\
 \text{(A.1)} \quad &= v_t dt + v_X \cdot (\mu X(t) dt + \sigma X(t) dW(t)) + \frac{1}{2} v_{XX} \cdot (dX(t))^2 \\
 &= \left(v_t + \mu X(t) v_X + \frac{1}{2} \sigma^2 X(t)^2 v_{XX} \right) dt + \sigma X(t) v_X dW(t).
 \end{aligned}$$

Define $\Pi := -v + v_X \cdot X$.

REMARK A.1. The function Π accounts for a portfolio consisting of

- (i) $\Delta := \frac{\partial}{\partial X} v = v_X$ stocks and
- (ii) -1 derivatives.

This is what is called a Δ -hedge in mathematical finance.

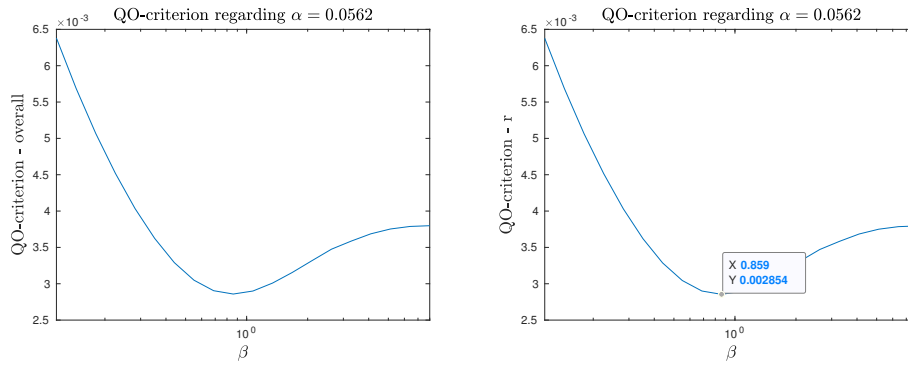


FIG. 8.8. *Quasi-optimality criterion for x and r , for suitable β .*

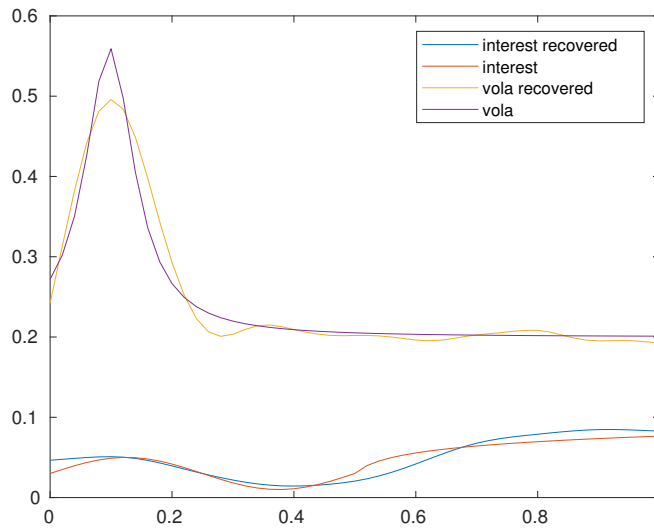


FIG. 8.9. *Regularized solutions $\alpha_{\alpha,\beta}^\delta$ and $r_{\alpha,\beta}^\delta$ and exact solutions a and r for $\alpha = 0.0562$ and $\beta = 0.859$; regularization parameters chosen from a combination of L-curve and quasi-optimality; 1% additive Gaussian noise.*

Using (A.1) we find that

$$\begin{aligned}
 d\Pi(t) &= -dv_t + v_X \cdot dX(t) \\
 &= - \left(v_t + \mu X(t)v_X + \frac{1}{2}\sigma^2 X(t)^2 v_{XX} \right) dt \\
 &\quad - \sigma v_X X(t) dW(t) + v_X (\mu X(t) dt + \sigma X(t) dW(t)) \\
 \text{(A.2)} \quad &= \left(-v_t - \frac{1}{2}\sigma^2 X(t)^2 v_{XX} \right) dt.
 \end{aligned}$$

Equation (A.2) reveals that the price of the portfolio $\Pi(t)$ is not random any longer, precisely, it does not depend on the random quantity $W(t)$, and further the drift μ appearing in the initial equation (2.1) is gone. This means that the portfolio $\Pi(t)$ is deterministic, i.e., not random.

However, there is only one risk-free asset on the market. It has interest r , and thus

$$(A.3) \quad d\Pi(t) = r\Pi(t)dt = r(-v + X(t) \cdot v_X) dt.$$

By comparing (A.2) and (A.3) we find the Black-Scholes differential equation

$$(A.4) \quad v_t + rXv_X + \frac{1}{2}\sigma^2X^2v_{XX} = rv.$$

Note in particular that this equation is free of the drift μ but involves the risk-free interest rate r instead. More importantly, throughout the derivation and particularly in (A.2), the interest rate r and volatility σ can be considered as *time-dependent* functions, which is essential for our setting. It is seen by direct evaluation that

$$v(t, X) := U_{BS}(X, K, \int_t^T r(u)du, \int_t^T a(u)du)$$

solves (A.4). This is (2.4) for time reversed, and thus Lemma 3.1 holds true.

REFERENCES

- [1] Y. ACHDOU AND O. PIRONNEAU, *Computational Methods for Option Pricing*, SIAM, Philadelphia, 2005.
- [2] V. ALBANI, U. M. ASCHER, X. YANG, AND J. P. ZUBELLI, *Data driven recovery of local volatility surfaces*, *Inverse Probl. Imaging*, 11 (2017), pp. 799–823.
- [3] V. ALBANI, A. DE CEZARO, AND J. P. ZUBELLI, *Convex regularization of local volatility estimation*, *Int. J. Theor. Appl. Finance*, 20 (2017), Art. 1750006, 37 pages.
- [4] V. ALBANI AND J. P. ZUBELLI, *Online local volatility calibration by convex regularization*, *Appl. Anal. Discrete Math.*, 8 (2014), pp. 243–268.
- [5] A. B. BAKUSHINSKY, *Remarks on choosing a regularization parameter using the quasioptimality and ratio criterion*, *U.S.S.R Comp. Math. Math. Phys.*, 24 (1984), pp. 181–182.
- [6] F. BAUER AND S. KINDERMANN, *The quasi-optimality criterion for classical inverse problems*, *Inverse Problems*, 24 (2008), Art. 035002, 20 pages.
- [7] M. BELGE, M. KILMER, AND E. MILLER, *Efficient determination of multiple regularization parameters in a generalized L-curve framework*, *Inverse Problems*, 18 (2002), pp. 1161–1183.
- [8] ———, *Simultaneous multiple regularization parameter selection by means of the L-hypersurface with applications to linear inverse problems posed in the wavelet transform domain*, in *Proc. SPIE'98—Bayesian Inference for Inverse Problems*, A. Mohammad-Djafari, ed., SPIE Proceedings Vol. 3459, SPIE, Bellingham, 1998.
- [9] F. BLACK AND M. SCHOLES, *The pricing of options and corporate liabilities*, *J. Polit. Econ.*, 81 (1973), pp. 637–654.
- [10] I. BOUCHOUVEV AND V. ISAKOV, *Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets*, *Inverse Problems*, 15 (1999), pp. R95–R116.
- [11] S. CRÉPEY, *Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization*, *SIAM J. Math. Anal.*, 34 (2003), pp. 1183–1206.
- [12] A. DE CEZARO, O. SCHERZER, AND J. P. ZUBELLI, *Convex regularization of local volatility models from option prices: convergence analysis and rates*, *Nonlinear Anal.*, 75 (2012), pp. 2398–2415.
- [13] A. DE CEZARO AND J. P. ZUBELLI, *The tangential cone condition for the iterative calibration of local volatility surfaces*, *IMA J. Appl. Math.*, 80 (2015), pp. 212–232.
- [14] B. DUPIRE, *Pricing with a smile*, *RISK*, 7 (1994), pp. 18–20.
- [15] H. EGGER AND H. W. ENGL, *Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates*, *Inverse Problems*, 21 (2005), pp. 1027–1045.
- [16] H. EGGER, T. HEIN, AND B. HOFMANN, *On decoupling of volatility smile and term structure in inverse option pricing*, *Inverse Problems*, 22 (2006), pp. 1247–1259.
- [17] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
- [18] H. GOLDBERG, W. KAMPOWSKY, AND F. TRÖLTZSCH, *On Nemytskij operators in L_p -spaces of abstract functions*, *Math. Nachr.*, 155 (1992), pp. 127–140.
- [19] M. HANKE AND P. C. HANSEN, *Regularization methods for large-scale problems*, *Surveys Math. Indust.*, 3 (1993), pp. 253–315.
- [20] M. HANKE AND O. SCHERZER, *Inverse problems light: numerical differentiation*, *Amer. Math. Monthly*, 108 (2001), pp. 512–521.

- [21] P. C. HANSEN AND D. P. O'LEARY, *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM J. Sci. Comput., 14 (1993), pp. 1487–1503.
- [22] T. HEIN AND B. HOFMANN, *On the nature of ill-posedness of an inverse problem arising in option pricing*, Inverse Problems, 19 (2003), pp. 1319–1338.
- [23] B. HOFMANN AND R. KRÄMER, *On maximum entropy regularization for a specific inverse problem of option pricing*, J. Inverse Ill-Posed Probl., 13 (2005), pp. 41–63.
- [24] B. HOFMANN AND R. PLATO, *On ill-posedness concepts, stable solvability and saturation*, J. Inverse Ill-Posed Probl., 26 (2018), pp. 287–297.
- [25] J. C. HULL, *Options, Futures, and Other Derivative Securities*, Prentice-Hall, London, 1997.
- [26] S. KINDERMANN AND A. NEUBAUER, *On the convergence of the quasioptimality criterion for (iterated) Tikhonov regularization*, Inverse Probl. Imaging, 2 (2008), pp. 291–299.
- [27] R. KRÄMER AND P. MATHÉ, *Modulus of continuity of Nemytskiĭ operators with application to the problem of option pricing*, J. Inverse Ill-Posed Probl., 16 (2008), pp. 435–461.
- [28] R. KRÄMER AND M. RICHTER, *Ill-posedness versus ill-conditioning – an example from inverse option pricing*, Appl. Anal., 87 (2008), pp. 465–477.
- [29] S. LU AND S. V. PEREVERZEV, *Regularization Theory for Ill-Posed Problems*, De Gruyter, Berlin, 2013.
- [30] V. NAUMOVA AND S. V. PEREVERZYEV, *Multi-penalty regularization with a component-wise penalization*, Inverse Problems, 29 (2013), Art. 075002, 15 pages.
- [31] T. REGIŃSKA, *A regularization parameter in discrete ill-posed problems*, SIAM J. Sci. Comput., 17 (1996), pp. 740–749.
- [32] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, AND F. LENZEN, *Variational Methods in Imaging*, Springer, New York, 2009.
- [33] T. SCHUSTER, B. KALTENBACHER, B. HOFMANN, AND K. S. KAZIMIERSKI, *Regularization Methods in Banach Spaces*, De Gruyter, Berlin, 2012.
- [34] J. P. ZUBELLI, *Inverse problems in finance: A short survey of calibration techniques*, in Proc. 2nd Brazilian Conference on Statistical Modelling in Insurance and Finance, Maresias, Brazil, 2005, N. Kolev and P. Morettin, eds., Institute of Mathematics and Statistics, University of São Paulo, 2005, pp. 64–75.