On maximum entropy regularization for a specific inverse problem of option pricing

Bernd Hofmann * and Romy Krämer †

Faculty of Mathematics, Technical University Chemnitz
D–09107 Chemnitz, Germany

MSC2000 classification scheme numbers: 65J20, 47H30, 49N45, 47B07, 91B28

Abstract

We investigate the applicability of the method of maximum entropy regularization (MER) to a specific nonlinear ill-posed inverse problem (SIP) in a purely time-dependent model of option pricing, introduced and analyzed for an L^2-setting in [9]. In order to include the identification of volatility functions with a weak pole, we extend the results of [12] and [13], concerning convergence and convergence rates of regularized solutions in L^1, in some details. Numerical case studies illustrate the chances and limitations of (MER) versus Tikhonov regularization (TR) for smooth solutions and solutions with a sharp peak. A particular paragraph is devoted to the singular case of at-the-money options, where derivatives of the forward operator degenerate.

1 Introduction

In this paper, we are dealt with a specific *ill-posed nonlinear inverse problem* that arises in financial markets (for an overview of such problems see [3]). The problem consists in finding (*calibrating*) a time-dependent volatility function defined on a finite time interval $I := [0, T]$ from the term structure on I of observed prices of vanilla call options with a fixed strike $K > 0$. This problem was introduced and discussed in an $L^2(I)$-setting in [9] with a convergence rate analysis of Tikhonov regularization based on the seminal paper [6]. Here, we consider solutions in $L^1(I)$ and show the theoretical and practical applicability of the method of *maximum entropy regularization* including convergence and convergence rates of regularized solutions as well as numerical case studies. In this context, we use the results of [12] and [13] and extend them in order to incorporate the case of reference functions with a weak pole.

*Email: hofmannb@mathematik.tu-chemnitz.de
†Email: rokra@mathematik.tu-chemnitz.de

The paper is organized as follows: In the remaining part of the introduction we define the specific inverse problem (SIP) under consideration in this paper and outline the problem structure by characterizing the forward operator as a composition of a linear integral operator and a nonlinear Nemytskii operator. Properties of the forward operator, which imply the local ill-posedness of the inverse problem, are given in §2. Then in §3 we apply the maximum entropy regularization to the problem (SIP) and discuss sufficient conditions for obtaining convergence rates. There occurs a singular case treated in §4, where strike price of the option and current asset price coincide. A comprehensive case study with synthetic data, presented in §5, completes the paper.

The restricted model under consideration uses a generalized *geometric Brownian motion* as stochastic process for an asset on which options are written. We denote by $X(\tau)$ the positive asset price at time τ. With a constant drift μ, time-dependent volatilities $\sigma(\tau)$ and a standard Wiener process $W(\tau)$ the stochastic differential equation

$$\frac{dX(\tau)}{X(\tau)} = \mu\, d\tau + \sigma(\tau)\, dW(\tau) \qquad (\tau \in I)$$

is assumed to hold. For an asset with current asset price $X := X(0) > 0$ at time $\tau = 0$ we consider a family of European vanilla call options with a fixed strike $K > 0$, a fixed risk-free interest rate $r \geq 0$ and maturities t varying through the whole interval I. We set $a(t) := \sigma^2(t)$ $(t \in I)$ and call this not directly observable function a, which expresses the volatility term structure, *volatility function*. Then it follows from stochastic considerations (for details see, e.g., [15, p.71/72]) that the associated *fair prices* $u(t)$ $(0 < t \leq T)$ of these options satisfy on an *arbitrage-free* market the equation

$$u(t) \;=\; X\,\Phi\left(\frac{\ln\left(\frac{X}{K}\right)+rt+\frac{1}{2}\int_0^t a(\tau)\,d\tau}{\sqrt{\int_0^t a(\tau)\,d\tau}}\right) - K\,e^{-rt}\Phi\left(\frac{\ln\left(\frac{X}{K}\right)+rt-\frac{1}{2}\int_0^t a(\tau)\,d\tau}{\sqrt{\int_0^t a(\tau)\,d\tau}}\right) \tag{1}$$

with the cumulative density function of the standard normal distribution

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

Moreover, the payoff of a European call at expiry provides

$$u(0) \;=\; \max(X - K, 0). \tag{3}$$

The Black-Scholes-type formula (1) – (3) is originally derived for positive continuous volatility functions, but it also yields well-defined values $u(t) \geq 0$ $(t \in I)$ if the function a is Lebesgue-integrable and almost everywhere finite and positive.

For parameters $X > 0$, $K > 0$, $r \geq 0$, $\tau \geq 0$ and $s \geq 0$ we introduce the *Black-Scholes function*

$$U_{BS}(X, K, r, \tau, s) := \begin{cases} X\Phi(d_1) - Ke^{-r\tau}\Phi(d_2) & (s > 0) \\ \max(X - Ke^{-r\tau}, 0) & (s = 0) \end{cases} \tag{4}$$

with

$$d_1 := \frac{\ln\left(\frac{X}{K}\right) + r\tau + \frac{s}{2}}{\sqrt{s}}, \quad d_2 := d_1 - \sqrt{s} \tag{5}$$

and $\Phi(\cdot)$ from formula (2). In terms of the auxiliary function

$$S(t) := \int_0^t a(\tau)\, d\tau \qquad (t \in I) \tag{6}$$

the option prices can be written concisely as

$$u(t) \;=\; U_{BS}(X, K, r, t, S(t)) \qquad (t \in I).$$

Now let a^* denote the *exact volatility function* of the underlying asset and S^* denote the corresponding auxiliary function obtained from a^* via formula (6). Instead of the fair price function

$$u^*(t) \;=\; U_{BS}(X, K, r, t, S^*(t)) \qquad (t \in I). \tag{7}$$

we observe a square-integrable noisy data function $u^\delta(t)$ $(t \in I)$, where u^* and u^δ belong to the *set D_+ of nonnegative functions over the interval I*. Then the specific inverse problem of identifying (calibrating) the volatility term structure a^* from noisy data u^δ can be expressed as follows:

Definition 1.1 (Specific inverse problem – SIP) *From a square-integrable noisy data function $u^\delta(t)$ $(t \in I)$ with noise level $\delta > 0$ and*

$$\|u^\delta - u^*\|_{L^2(I)} = \left(\int_I \left(u^\delta(t) - u^*(t)\right)^2 \, dt \right)^{\frac{1}{2}} \;\leq\; \delta$$

find appropriate approximations a^δ of the function a^*, where both a^δ and a^* are integrable and almost everywhere nonnegative functions over I and we measure the accuracy of a^δ by

$$\|a^\delta - a^*\|_{L^1(I)} = \int_I \left|a^\delta(\tau) - a^*(\tau)\right| \, d\tau.$$

The inverse problem (SIP) corresponds with the solution of the nonlinear operator equation

$$F(a) \;=\; u \qquad (a \in D(F) \subset B_1, \quad u \in D_+ \cap L^2(I) \subset B_2) \tag{8}$$

in the pair of Banach spaces

$$B_1 := L^1(I) \quad \text{and} \quad B_2 := L^2(I)$$

with a given noisy right-hand side, where the *nonlinear forward operator*

$$F = N \circ J : \;\; D(F) \subset B_1 \longrightarrow B_2$$

with domain

$$D(F) := \left\{ a \in L^1(I) : a(\tau) \geq 0 \text{ a.e. in } I \right\} \tag{9}$$

is decomposed into the inner *linear Volterra integral operator* $J : B_1 \longrightarrow B_2$ with

$$[J(h)](t) := \int_0^t h(\tau)\, d\tau \qquad (t \in I) \tag{10}$$

and the outer *nonlinear Nemytskii operator* $N : D(N) := D_+ \cap L^2(I) \subset B_2 \longrightarrow B_2$ defined as

$$[N(S)](t) := U_{BS}(X, K, r, t, S(t)) \qquad (t \in I) \tag{11}$$

by using the Black-Scholes function (4) – (5).

2 Properties of the forward operator and ill-posedness of the inverse problem

In order to characterize the forward operator F in the pair of Banach spaces B_1 and B_2, we first study the components J and N of its decomposition. For example as a consequence of Theorem 4.3.3 in [8] we obtain the *continuity* and *compactness* of the *injective* operator $J : L^1(I) \to L^2(I)$ defined by formula (10).

For studying the operator N, we summarize the main properties of the Black-Scholes function U_{BS} by the following lemma, which can be proven straightforward by elementary calculations.

Lemma 2.1 *Let the parameters $X > 0$, $K > 0$ and $r \geq 0$ be fixed. Then the nonnegative function $U_{BS}(X, K, r, \tau, s)$ is continuous for $(\tau, s) \in [0, \infty) \times [0, \infty)$. Moreover, for $(\tau, s) \in [0, \infty) \times (0, \infty)$, this function is continuously differentiable with respect to τ, where we have*

$$\frac{\partial U_{BS}(X, K, r, \tau, s)}{\partial \tau} = r\, K\, e^{-r\tau}\, \Phi(d_2) \geq 0,$$

and twice continuously differentiable with respect to s, where we have with $\nu := \ln\left(\frac{X}{K}\right)$

$$\begin{aligned} \frac{\partial U_{BS}(X,K,r,\tau,s)}{\partial s} &= \Phi'(d_1)\, X\, \frac{1}{2\sqrt{s}} \\ &= \frac{X}{2\sqrt{2\,\pi\, s}} \exp\left(-\frac{[\nu + r\,\tau]^2}{2\,s} - \frac{[\nu + r\,\tau]}{2} - \frac{s}{8}\right) > 0 \end{aligned} \tag{12}$$

and

$$\begin{aligned} \frac{\partial^2 U_{BS}(X,K,r,\tau,s)}{\partial s^2} &= -\Phi'(d_1)\, X\, \frac{1}{4\sqrt{s}} \left(-\frac{[\nu + r\tau]^2}{s^2} + \frac{1}{4} + \frac{1}{s}\right) \\ &= -\frac{X}{4\sqrt{2\,\pi\, s}} \left(-\frac{[\nu + r\tau]^2}{s^2} + \frac{1}{4} + \frac{1}{s}\right) \exp\left(-\frac{[\nu + r\,\tau]^2}{2\,s} - \frac{[\nu + r\,\tau]}{2} - \frac{s}{8}\right). \end{aligned} \tag{13}$$

Furthermore, we find the limit conditions

$$\lim_{s \to \infty} U_{BS}(X, K, r, \tau, s) = X \tag{14}$$

and

$$\lim_{s \to 0} \frac{\partial U_{BS}(X, K, r, \tau, s)}{\partial s} = \begin{cases} \infty & (X = K\, e^{-r\tau}) \\ 0 & (X \neq K\, e^{-r\tau}) \end{cases} \tag{15}$$

The Nemytskii operator N defined by formula (11) *maps continuously* in $L^2(I)$, since the function $k(\tau, s) := U_{BS}(X, K, r, \tau, s)$ satisfies the Caratheodory condition and the growth condition (see, e.g., [14, p.52]). Namely, due to the formulae (4), (12) and (14) $k(\tau, s)$ is continuous and uniformly bounded with $0 \leq k(\tau, s) \leq X$ for all $(\tau, s) \in I \times [0, \infty)$. From formula (12) of Lemma 2.1 we moreover obtain $\frac{\partial k(\tau,s)}{\partial s} > 0$ for all $(\tau, s) \in I \times (0, \infty)$ and hence the *injectivity* of the operator N on its domain $D(N)$.

Lemma 2.2 *The nonlinear operator $F : D(F) \subset L^1(I) \to L^2(I)$ is compact, continuous, weakly continuous and injective. Thus, the inverse operator F^{-1} defined on the range $F(D(F))$ of F exists.*

Proof: Since the linear operator J maps injective and compact from $L^1(I)$ to $L^2(I)$ and the nonlinear operator N maps injective and continuous from $D(N) \subset L^2(I)$ to $L^2(I)$, the composite operator $F = N \circ J$ is injective, continuous and compact. Consequently, F transforms weakly convergent sequences $\{a_n\}_{n=1}^{\infty} \subset D(F)$ with $a_n \rightharpoonup a_0$ in $L^1(I)$ into strongly convergent sequences $F(a_n) \to F(a_0)$ in $L^2(I)$ and is weakly continuous. Due to the injectivity of F, the inverse $F^{-1} : F(D(F)) \subset L^2(I) \to D(F) \subset L^1(I)$ is well-defined. Note that the range $F(D(F)) \subset D_+$ contains only continuous functions over I ∎

Now we are in search of solution points $a^* \in D(F)$, for which the inverse problem (SIP) written as an operator equation (8) is *locally ill-posed* or *locally well-posed*, respectively (cf. [10, Definition 2]).

Definition 2.3 *We call the operator equation (8) between the Banach spaces B_1 and B_2 locally ill-posed at the point $a^* \in D(F)$ if, for all balls $B_r(a^*)$ with radius $r > 0$ and center a^*, there exist sequences $\{a_n\}_{n=1}^{\infty} \subset B_r(a^*) \cap D(F)$ satisfying the condition*

$$F(a_n) \to F(a^*) \quad \text{in} \ B_2, \quad \text{but} \quad a_n \not\to a^* \quad \text{in} \ B_1 \quad \text{as} \quad n \to \infty .$$

Otherwise, we call the equation (8) locally well-posed at $a^* \in D(F)$ if there exists a radius $r > 0$ such that

$$F(a_n) \to F(a^*) \quad \text{in} \ B_2 \quad \Longrightarrow \quad a_n \to a^* \quad \text{in} \ B_1 \quad \text{as} \quad n \to \infty,$$

for all sequences $\{a_n\}_{n=1}^{\infty} \subset B_r(a^*) \cap D(F)$.

Note that, for the injective operator F under consideration, local ill-posedness at a^* implies discontinuity of F^{-1} at the point $u^* = F(a^*)$ and vice versa continuity of F^{-1} at u^* implies local well-posedness at a^*.

Theorem 2.4 *The operator equation (8) associated with the inverse problem (SIP) is locally ill-posed at the solution point $a^* \in D(F)$ whenever we have*

$$\operatorname*{ess\,inf}_{\tau \in I} a^*(\tau) \ > \ 0. \tag{16}$$

Proof: For a function $a^* \in D(F)$ satisfying (16) let $\varepsilon > 0$ in $a_n(\tau) := a^*(\tau) + \varepsilon \sin(n\tau)$ $(\tau \in I)$ be small enough in order to ensure $a_n \in D(F) \cap B_r(a^*)$. Due to the Riemann-Lebesgue lemma we have $\lim_{n\to\infty} \int_I \sin(n\tau)\varphi(\tau)d\tau = 0$ for $\varphi \in L^\infty(I)$ and thus $a_n \rightharpoonup a^*$. On the other hand, $\int_I |\sin(n\tau)|d\tau \not\to 0$ provides $a_n \not\to a^*$ in $L^1(I)$ as $n \to \infty$. Then the compactness of F implies $F(a_n) \to F(a^*)$ in $L^2(I)$ and hence the local ill-posedness at the point. a^* ∎

By the arguments of the above proof it cannot be excluded that (8) is locally well-posed if $\operatorname{ess\,inf}_{\tau\in I} a^*(\tau) = 0$. For example in the case $a^* = \mathbf{0}$ (zero function) the weak convergence $a_n \rightharpoonup \mathbf{0}$ in $L^1(I)$ implies strong convergence $a_n \to \mathbf{0}$ in $L^1(I)$ if all functions a_n are nonnegative a.e. However, if there is no reason to assume a purely deterministic behavior of the asset price $X(t)$ for some time interval, the ill-posed situation caused by (16) seems to be realistic.

Thus, a regularization approach is required for the stable solution of the nonlinear inverse problem (SIP). The standard *Tikhonov regularization* (TR) approach in the sense of [6] with penalty functional $\|a - \overline{a}\|^2_{L^2(I)}$ for that problem including considerations on the strength of ill-posedness is studied in the paper [9]. Here, we will apply the *maximum entropy regularization* (MER) to (SIP) and extend in this context the results of the papers [12] and [13] concerning convergence and convergence rates to the case of reference functions, which are not necessarily in $L^\infty(I)$.

3 Applicability of maximum entropy regularization

Since the solution space of the inverse problem (SIP) is $L^1(I)$ and we have a stochastic background, the use of maximum entropy regularization (MER) as an appropriate regularization approach (cf. [5, Chapters 5.3 and 10.6] and the references therein) is motivated. We use the penalty functional

$$E(a, \overline{a}) := \int_I \left\{ a(\tau) \ln\left(\frac{a(\tau)}{\overline{a}(\tau)}\right) + \overline{a}(\tau) - a(\tau) \right\} d\tau \tag{17}$$

called in [4] *cross entropy* relative to a fixed reference function $\overline{a} \in L^1(I)$ satisfying the condition

$$0 < \underline{c} \le \overline{a}(\tau) \quad \text{a.e. on } I. \tag{18}$$

For the unique solution $a^* \in D(F)$ (see (9)) of equation $F(a) = u^*$ with the exact right-hand side u^* (see (7)) we assume in the sequel $a^* \in D(E)$ with

$$D(E) := \left\{ a \in D(F) \subset L^1(I) : E(a, \overline{a}) < \infty \right\}. \tag{19}$$

Since the weak continuity of F implies the weak closedness of the operator, we obtain from Lemma 2.2:

Corollary 3.1 *The nonlinear operator $F : D(F) \subset L^1(I) \to L^2(I)$ possessing a convex and in $L^1(I)$ weakly closed domain $D(F)$ is continuous, weakly closed and injective.*

For the operator equation (8) with an operator F as characterized by Corollary 3.1 it is useful to consider regularized solutions a_α^δ solving the extremal problem

$$\|F(a) - u^\delta\|^2_{L^2(I)} + \alpha\, E(a, \overline{a}) \longrightarrow \min, \qquad \text{subject to} \qquad a \in D(E). \tag{20}$$

If the unknown solution a^* is normalized by specifying $\|a^*\|_{L^1(I)}$, frequently the penalty functional

$$\widetilde{E}(a, \overline{a}) := \int_I a(\tau)\, \ln\left(\frac{a(\tau)}{\overline{a}(\tau)}\right)\, d\tau \tag{21}$$

is considered instead of (17). This situation has been studied in [12] and [13] combined with a reference $\overline{a} \in L^\infty(I)$ satisfying the condition

$$0 < \underline{c} \le \overline{a}(\tau) \le \overline{c} < \infty \quad \text{a.e. on } I. \tag{22}$$

Note that we have

$$E(a, \overline{a}) = \widetilde{E}(a, e \cdot \overline{a}) + \int_I \overline{a}(\tau)\, d\tau$$

and the functional (17) attains its minimum for $a = \overline{a}$, whereas (21) attains its minimum for $a = \frac{\overline{a}}{e}$. Moreover, the domains $\widetilde{D}(E) := \left\{a \in D(F) \subset L^1(I) : \widetilde{E}(a, \overline{a}) < \infty\right\}$ and $D(E)$ (see (19)) coincide. Whenever $\overline{a}$ satisfies (22), the extremal problem (20) is equivalent to

$$\|F(a) - u^\delta\|^2_{L^2(I)} + \alpha \widetilde{E}(a, e \cdot \overline{a}) \longrightarrow \min, \qquad \text{subject to} \qquad a \in \widetilde{D}(E) \tag{23}$$

and the theoretical results of [12] and [13] apply to regularized solutions a_α^δ obtained by solving (20).

Using [4, Lemmas 2.1 - 2.3] and the following Lemma 3.2, which is as an extension of [12, Lemma 3], one can show along the lines of the proofs of [12, Theorems 1 and 2] that for all regularization parameters $\alpha > 0$ the minimization problem (20) is solvable (not necessarily unique) and the regularized solutions a_α^δ stably depend on the data u^δ. In contrast to (22) we include in (18) functions $\overline{a}$ with a weak pole. This seems to be reasonable, since the admissible solutions a^* of (SIP) may also have such a pole.

Lemma 3.2 *For $\{a_n\}_{n=1}^\infty \subset D(E)$, $a_0 \in D(E)$ and $\overline{a}$ satisfying (18) we have*

$$a_n \rightharpoonup a_0 \text{ in } L^1(I) \ \wedge\ E(a_n, \overline{a}) \to E(a_0, \overline{a}) \implies a_n \to a_0 \text{ in } L^1(I) \text{ as } n \to \infty,$$

i.e., weak convergence and convergence in entropy imply strong convergence in $L^1(I)$.

Proof: Let $\mathcal{B}$ denote the Borel subsets of I. We define on $(I, \mathcal{B})$ a measure

$$\theta(A) := \int_A e \cdot \overline{a}(\tau) d\tau \qquad (A \in \mathcal{B})$$

and consider the finite measure space $L^1(I, \mathcal{B}, \theta)$. Weak convergence of a sequence $\{b_n\}_{n=1}^\infty \in L^1(I, \mathcal{B}, \theta)$ to an element $b_0 \in L_1(I, \mathcal{B}, \theta)$ means, for all $f \in L^\infty(I, \mathcal{B}, \theta)$,

$$\int_I b_n f d\theta \to \int_I b_0 f d\theta \qquad \text{as } n \to \infty.$$

Note that $f \in L^\infty(I, \mathcal{B}, \theta)$ is equivalent to $f \in L^\infty(I)$. Moreover, for $b_n := \frac{a_n}{e \cdot \overline{a}}$ and $b_0 := \frac{a_0}{e \cdot \overline{a}}$ the equivalences

$$a_n \to a_0 \quad \text{in } L^1(I) \iff b_n \to b_0 \quad \text{in } L^1(I, \mathcal{B}, \theta),$$

$$a_n \rightharpoonup a_0 \quad \text{in } L^1(I) \iff b_n \rightharpoonup b_0 \quad \text{in } L^1(I, \mathcal{B}, \theta)$$

and

$$\int_I \left\{a_n \ln \tfrac{a_n}{\overline{a}} + \overline{a} - a_n\right\} d\tau \to \int_I \left\{a_o \ln \tfrac{a_o}{\overline{a}} + \overline{a} - a_o\right\} d\tau \iff \int_I \{b_n \ln b_n\} d\theta \to \int_I \{b_o \ln b_o\} d\theta$$

hold. Theorem 2.7 in [2] asserts that

$$b_n \rightharpoonup b_0 \quad \text{in } L^1(I, \mathcal{B}, \theta) \qquad \text{and} \qquad \int_I \{b_n \ln b_n\}\, d\theta \to \int_I \{b_0 \ln b_0\}\, d\theta$$

imply $b_n \to b_0$ in $L^1(I, \mathcal{B}, \theta)$. This assertion combined with the above implications proves the lemma. ■

Now we can prove a convergence theorem for maximum entropy method. Note that for the entropy functional (21) and a reference function $\overline{a}$ satisfying (22) this result directly follows from [12, Theorem 3].

Theorem 3.3 *Let $a^* \in D(E)$ be the exact solution of problem (SIP) associated with noiseless data u^*. Then, for a sequence $\{u^{\delta_n}\}_{n=1}^\infty$ of noisy data with $\|u^{\delta_n} - u^*\|_{L^2(I)} \le \delta_n$ and $\delta_n \to 0$ as $n \to \infty$ and for a sequence $\{\alpha_n = \alpha(\delta_n)\}_{n=1}^\infty$ of positive regularization parameters with*

$$\alpha_n \to 0 \quad \text{and} \quad \frac{\delta_n^2}{\alpha_n} \to 0 \quad \text{as} \quad n \to \infty,$$

any sequence $\{a_{\alpha_n}^{\delta_n}\}_{n=1}^\infty$ of corresponding regularized solutions converges to the exact solution in $L^1(I)$, i.e., we have

$$\lim_{n\to\infty} \|a_{\alpha_n}^{\delta_n} - a^*\|_{L^1(I)} \;=\; 0.$$

Proof: By definition of $a_{\alpha_n}^{\delta_n}$ we have

$$\|F\left(a_{\alpha_n}^{\delta_n}\right) - u^{\delta_n}\|_{L^2(I)}^2 + \alpha_n E\left(a_{\alpha_n}^{\delta_n}, \overline{a}\right) \le \|F(a^*) - u^{\delta_n}\|_{L^2(I)}^2 + \alpha_n E\left(a^*, \overline{a}\right). \tag{24}$$

By the choice of $\alpha(\delta)$ there exists a constant $C_0 > 0$ such that $E\left(a_{\alpha_n}^{\delta_n}, \overline{a}\right) \le C_0$ for sufficiently large n. Since the level sets $\mathcal{E}_C := \{a \in D(E) : E(a, \overline{a}) \le C\}$ are weakly compact in $L^1(I)$ (see, e.g., [4, Lemma 2.3]), there are a subsequence $\left\{a_{\alpha_{n_k}}^{\delta_{n_k}}\right\}_{k=1}^\infty$ of $\{a_{\alpha_n}^{\delta_n}\}_{n=1}^\infty$ and an element $\widetilde{a} \in L^1(I)$ with $a_{\alpha_{n_k}}^{\delta_{n_k}} \rightharpoonup \widetilde{a}$ in $L^1(I)$ as $k \to \infty$. Moreover, (24) implies $\lim_{k\to\infty} \|F\left(a_{\alpha_{n_k}}^{\delta_{n_k}}\right) - u^{\delta_{n_k}}\|_{L^2(I)}^2 = 0$ and therefore $F\left(a_{\alpha_{n_k}}^{\delta_{n_k}}\right) \to u^*$ in $L^2(I)$. Thus, by the weak closedness of F we have $\widetilde{a} \in D(F)$ and $F(\widetilde{a}) = u^*$. Since F is injective, this implies $\widetilde{a} = a^*$. It remains to show that $a_{\alpha_{n_k}}^{\delta_{n_k}} \to a^*$ in $L^1(I)$ as $k \to \infty$, since then any subsequence of $\{a_{\alpha_n}^{\delta_n}\}_{n=1}^\infty$ has the same convergence property. Applying Lemma 3.2 we only have to show $E\left(a_{\alpha_{n_k}}^{\delta_{n_k}}, \overline{a}\right) \to E(a^*, \overline{a})$ as $k \to \infty$. From the weak lower semicontinuity of

$E(\cdot, \overline{a})$ together with (24) and the choice of $\alpha(\delta)$ we immediately derive this convergence in entropy by the inequalities

$$E(a^*, \overline{a}) \leq \liminf_{k\to\infty} E\left(a_{\alpha_{n_k}}^{\delta_{n_k}}, \overline{a}\right) \leq \limsup_{k\to\infty} E\left(a_{\alpha_n}^{\delta_n}, \overline{a}\right) \leq E(a^*, \overline{a})\,. \blacksquare$$

The convergence formulated in Theorem 3.3 may be arbitrarily slow. To obtain a convergence rate of regularized solutions,

$$\|a_\alpha^\delta - a^*\|_{L^1(I)} = \mathcal{O}\left(\sqrt{\delta}\right) \quad \text{as } \delta \to 0, \tag{25}$$

in the proof of the next theorem we follow the ideas of Theorem 1 from [13] (originally formulated and proven in Chinese in [13]) and extend those results to our situation. Here we use the Landau symbol $\mathcal{O}$ in (25) in the sense of the existence of a constant $C > 0$ satisfying for sufficiently small δ the estimate above $\|a_\alpha^\delta - a^*\|_{L^1(I)} \leq C\sqrt{\delta}$. Note that convergence rates results for the maximum entropy regularization avoiding the assumption of weak closedness of the nonlinear forward operator are given in [7].

Theorem 3.4 *Let the operator $H : D(H) \subset L^1(I) \longrightarrow L^2(I)$ be continuous and weakly closed and let the domain $D(H)$ of H, containing functions which are nonnegative a.e. on I, be convex. Moreover, let the operator equation $H(a) = u^*$, for the given right-hand side $u^* \in L^2(I)$, have a unique solution $a^* \in D(H)$ with $E(a^*, \overline{a}) < \infty$. Then we have, for data $u^\delta \in L^2(I)$ with $\|u^\delta - u^*\|_{L^2(I)} \leq \delta$ and regularization parameters chosen as $\alpha \sim \delta$, regularized solutions a_α^δ solving the extremal problems*

$$\|H(a) - u^\delta\|^2_{L^2(I)} + \alpha\, E(a, \overline{a}) \longrightarrow \min, \quad \textit{subject to } a \in D(H) \textit{ with } E(a, \overline{a}) < \infty,$$

which provide a convergence rate (25) whenever there exists a continuous linear operator $G : L^1(I) \longrightarrow L^2(I)$ with a (Banach space) adjoint $G^* : L^2(I) \longrightarrow L^\infty(I)$ and a positive constant L such that the following three conditions are satisfied:

(i) $\quad \|H(a) - H(a^*) - G(a - a^*)\|_{L^2(I)} \leq \dfrac{L}{2}\, \|a - a^*\|^2_{L^1(I)}$ *for all* $a \in D(H)$,

(ii) *there exists a function* $w \in L^2(I)$ *satisfying* $\ln\left(\frac{a^*}{\overline{a}}\right) = G^* w$

and

(iii) $\quad L\, \|w\|_{L^2(I)}\, \|a^*\|_{L^1(I)} < 1.$

Proof: By definition of a_α^δ we have

$$\|H\left(a_\alpha^\delta\right) - u^\delta\|^2_{L^2(I)} + \alpha E\left(a_\alpha^\delta, \overline{a}\right) \leq \|H(a^*) - u^\delta\|^2_{L^2(I)} + \alpha E(a^*, \overline{a})$$

and hence

$$\|H\left(a_\alpha^\delta\right) - u^\delta\|^2_{L^2(I)} + \alpha E\left(a_\alpha^\delta, a^*\right) \leq \|H(a^*) - u^\delta\|^2_{L^2(I)} + \alpha \int_I \left(a^* - a_\alpha^\delta\right) \ln\left(\frac{a^*}{\overline{a}}\right) d\tau.$$

With condition (ii) we then obtain

$$\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}^2+\alpha E\left(a_\alpha^\delta,a^*\right)\leq\|H(a^*)-u^\delta\|_{L^2(I)}^2+\alpha\int\limits_I\left(a^*-a_\alpha^\delta\right)G^*\,w\,d\tau.$$

The integral in the last formula can be understood as a duality product: $G^*\,w\in L^\infty(I)$ is applied to $a^*-a_\alpha^\delta\in L^1(I)$. Using the definition of the adjoint operator G^* we obtain

$$\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}^2+\alpha E\left(a_\alpha^\delta,a^*\right)\leq\|H(a^*)-u^\delta\|_{L^2(I)}^2+\alpha\int\limits_I G\left(a^*-a_\alpha^\delta\right)\,w\,d\tau. \qquad (26)$$

For $r\left(a_\alpha^\delta,a^*\right):=H\left(a_\alpha^\delta\right)-H(a^*)-G\left(a_\alpha^\delta-a^*\right)$ we have by condition (i)

$$\|r\left(a_\alpha^\delta,a^*\right)\|_{L^2(I)}\leq\frac{L}{2}\|a_\alpha^\delta-a^*\|_{L^1(I)}^2$$

Substituting

$$G\left(a^*-a_\alpha^\delta\right)=H(a^*)-H\left(a_\alpha^\delta\right)+r\left(a_\alpha^\delta,a^*\right)$$

in (26) we get

$$\begin{aligned}&\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}^2+\alpha E\left(a_\alpha^\delta,a^*\right)\\&\quad\leq\|H(a^*)-u^\delta\|_{L^2(I)}^2+\alpha\left\langle w,H(a^*)-H\left(a_\alpha^\delta\right)+r\left(a_\alpha^\delta,a^*\right)\right\rangle_{L^2(I)},\end{aligned}$$

where $\langle\cdot,\cdot\rangle_{L^2(I)}$ denotes the common inner product in $L^2(I)$. Because of formula (1.7) in [4, p.1558] we have

$$\left(\frac{2}{3}\|a_\alpha^\delta\|_{L^1(I)}+\frac{4}{3}\|a^*\|_{L^1(I)}\right)^{-1}\|a_\alpha^\delta-a^*\|_{L^1(I)}^2\leq E\left(a_\alpha^\delta,a^*\right).$$

Therefore

$$\begin{aligned}&\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}^2+\alpha\left(\frac{2}{3}\|a_\alpha^\delta\|_{L^1(I)}+\frac{4}{3}\|a^*\|_{L^1(I)}\right)^{-1}\|a_\alpha^\delta-a^*\|_{L^1(I)}^2 \qquad (27)\\&\quad\leq\delta^2+\alpha\delta\|w\|_{L^2(I)}+\alpha\|w\|_{L^2(I)}\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}+\frac{1}{2}\alpha L\|w\|_{L^2(I)}\|a_\alpha^\delta-a^*\|_{L^1(I)}^2.\end{aligned}$$

Following the proof of Theorem 3.3 here we also have $\lim\limits_{\delta\to0}\|a_\alpha^\delta-a^*\|_{L^1(I)}=0$ for the parameter choice $\alpha\sim\delta$. For sufficiently small values δ condition (iii) implies

$$\frac{1}{2}L\|w\|_{L^2(I)}\left(\frac{2}{3}\|a_\alpha^\delta\|_{L^1(I)}+\frac{4}{3}\|a^*\|_{L^1(I)}\right)<1. \qquad (28)$$

Thus we have

$$\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}^2\leq\delta^2+\alpha\delta\|w\|_{L^2(I)}+\alpha\|w\|_{L^2(I)}\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}.$$

With the implication $a,b,c\geq0\ \wedge\ a^2\leq b^2+ac\ \Rightarrow\ a\leq b+c$ we obtain

$$\|H\left(a_\alpha^\delta\right)-u^\delta\|_{L^2(I)}\leq\sqrt{\delta^2+\alpha\delta\|w\|_{L^2(I)}}+\alpha\|w\|_{L^2(I)}$$

and by substituting this relation into (27)

$$\alpha \left[\left(\frac{2}{3} \|a_\alpha^\delta\|_{L^1(I)} + \frac{4}{3} \|a^*\|_{L^1(I)} \right)^{-1} - \frac{1}{2} L \|w\|_{L^2(I)} \right] \|a_\alpha^\delta - a^*\|^2_{L^1(I)} \tag{29}$$
$$\leq \delta^2 + \alpha\delta \|w\|_{L^2(I)} + \alpha \|w\|_{L^2(I)} \left(\sqrt{\delta^2 + \alpha\delta \|w\|_{L^2(I)}} + \alpha \|w\|_{L^2(I)} \right).$$

For the parameter choice $\alpha \sim \delta$ we finally obtain $\|a_\alpha^\delta - a^*\|_{L^1(I)} = \mathcal{O}(\sqrt{\delta})$ for $\delta \to 0$ from formula (29) by considering the inequality (28). ∎

In order to apply Theorem 3.4 with $H = F$ to our problem (SIP), we restrict the domain of F in the form

$$\widetilde{D}(F) := \left\{ a \in L^1(I) : a(\tau) \geq \underline{c} > 0 \text{ a.e. in } I \right\}$$

and assume $a^* \in \widetilde{D}(F) \cap D(E)$ for a given lower positive bound $\underline{c}$ also occurring in condition (18). Since $\widetilde{D}(F)$ is also convex and closed, $F : \widetilde{D}(F) \subset L^1(I) \to L^2(I)$ remains weakly closed. Note that the functions $S = J(a)$ according to (10) with $a \in \widetilde{D}(F)$ satisfy the condition

$$S(t) \geq \underline{c}\, t \quad (t \in I). \tag{30}$$

Setting $H := F$ and $D(H) := \widetilde{D}(F)$ the operator G in Theorem 3.4 can be considered as the Fréchet derivative $F'(a^*)$ of F at the point a^* neglecting the fact that $\widetilde{D}(F)$ has an empty interior in $L^1(I)$ (cf. [5, Remark 10.30]). If there exists a linear operator G satisfying the condition (i) in Theorem 3.4, then the structure of G can be verified as a (formal) Gâteaux derivative by a limiting process outlined in [9, §5] in the form $G = M \circ J$ or

$$[G(h)](t) = m(t)\,[J(h)](t) \quad \left(t \in I, \quad h \in L^1(I)\right). \tag{31}$$

with a linear multiplication operator M defined by the multiplier function

$$m(0) = 0, \quad m(t) = \frac{\partial U_{BS}(X, K, r, t, S^*(t))}{\partial s} > 0 \quad (0 < t \leq T), \tag{32}$$

where $S^* := J(a^*)$ and we can prove:

Theorem 3.5 *In the case $X \neq K$ we have $m \in L^\infty(I)$ and the operator G defined by the formulae (31) and (32) maps continuously from $L^1(I)$ to $L^2(I)$. Then the condition*

$$\|F(a) - F(a^*) - G(a - a^*)\|_{L^2(I)} \leq \frac{L}{2} \|a - a^*\|^2_{L^1(I)} \quad \textit{for all} \quad a \in \widetilde{D}(F) \tag{33}$$

is satisfied with a constant

$$L = \sqrt{T}\, C_2, \quad \textit{where} \quad C_2 := \sup_{(\tau,s)\in\mathcal{M}_{\underline{c}}} \left| \frac{\partial^2 U_{BS}(X, K, r, \tau, s)}{\partial s^2} \right| < \infty$$

is determined from the set

$$\mathcal{M}_{\underline{c}} := \{(\tau, s) \in I\!R^2 : s \geq \underline{c}\,\tau, \;\; 0 < \tau \leq T\}.$$

Proof: We make use of the fact that a multiplication operator M is continuous in $L^2(I)$ if the multiplier function m belongs to $L^\infty(I)$. From formula (12) we obtain for $(\tau, s) \in I \times (0,\infty)$ the estimate

$$\left|\frac{\partial U_{BS}(X,K,r,\tau,s)}{\partial s}\right| = \frac{X}{\sqrt{8\,\pi\,s}} \exp\left(-\frac{\left[\ln\left(\frac{X}{K}\right)+r\,\tau\right]^2}{2\,s} - \frac{\ln\left(\frac{X}{K}\right)+r\,\tau}{2} - \frac{s}{8}\right)$$
$$\leq \sqrt{\frac{X\,K}{8\,\pi}}\,\frac{1}{\sqrt{s}}\,\exp\left(-\frac{\left[\ln\left(\frac{X}{K}\right)+r\,\tau\right]^2}{2\,s}\right).$$

This implies for $(\tau,s) \in \mathcal{M}_{\underline{c}}$

$$\left|\frac{\partial U_{BS}(X,K,r,\tau,s)}{\partial s}\right| \leq \sqrt{\frac{X\,K}{8\,\pi}}\,\left(\frac{K}{X}\right)^{\frac{r}{\underline{c}}}\,\frac{1}{\sqrt{s}}\,\exp\left(-\frac{\left[\ln\left(\frac{X}{K}\right)\right]^2}{2\,s}\right). \tag{34}$$

For $X \neq K$ the right expression in the inequality (34) is continuous with respect to $s \in (0,\infty)$ and tends to zero as $s \to 0$ and as $s \to \infty$. With a finite constant $C_1 := \sup\limits_{(\tau,s)\in\mathcal{M}_{\underline{c}}} \left|\frac{\partial U_{BS}(X,K,r,\tau,s)}{\partial s}\right| < \infty$ we then have $m \in L^\infty(I)$ in that case. In order to prove the condition (33) for $X \neq K$, we use the structure of the second derivative $\frac{\partial^2 U_{BS}(X,K,r,\tau,s)}{\partial s^2}$ as expressed by formula (13). Similar considerations as in the case of the first derivative also show the existence of a constant $C_2 := \sup\limits_{(\tau,s)\in\mathcal{M}_{\underline{c}}} \left|\frac{\partial^2 U_{BS}(X,K,r,\tau,s)}{\partial s^2}\right| < \infty$.

Then we can estimate for $S = J(a)$, $S^* = J(a^*)$, $a, a^* \in \widetilde{D}(F)$ and $T \in I$

$$|[F(a) - F(a^*) - G(a-a^*)](t)| =$$
$$= \left|U_{BS}(X,K,r,t,S(t)) - U_{BS}(X,K,r,t,S^*(t)) - \tfrac{\partial U_{BS}(X,K,r,t,S^*(t))}{\partial s}(S(t)-S^*(t))\right|$$
$$= \tfrac{1}{2}\left|\tfrac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}(S(t)-S^*(t))^2\right| \leq \frac{C_2}{2}\,\|a-a^*\|^2_{L^1(I)},$$

where S_{im} with $\min(S(t),S^*(t)) \leq S_{im}(t) \leq \max(S(t),S^*(t))$ for $0 < t \leq T$ is an intermediate function such that the pairs of real numbers $(t,S(t))$, $(t,S^*(t))$ and $(t,S_{im}(t))$ all belong to the set $\mathcal{M}_{\underline{c}}$. This provides

$$\|F(a) - F(a^*) - G(a-a^*)\|_{L^2(I)} \leq \frac{\sqrt{T}\,C_2}{2}\,\|a-a^*\|^2_{L^1(I)}$$

and hence the condition (33), which proves the theorem. ∎

In order to interpret the conditions (ii) and (iii) in Theorem 3.4 for problem (SIP) with $H = F$ in the case $X \neq K$, we write (ii) as

$$\ln\left(\frac{a^*(t)}{\overline{a}(t)}\right) = \int_t^T m(\tau)\,w(\tau)d\tau \quad \left(t \in I, \quad w \in L^2(I)\right). \tag{35}$$

If (ii) is satisfied, then the function $\ln\left(\frac{a^*}{\overline{a}}\right)$ belongs to the Sobolev space $H^1(I)$, which is embedded in the space of continuous functions $C(I)$. Moreover, we get

$$a^*(T) = \overline{a}(T) \quad \text{and} \qquad w = -\frac{\left[\ln\left(\frac{a^*}{\overline{a}}\right)\right]'}{m} \in L^2(I), \tag{36}$$

where the weight function $\frac{1}{m}$ does not belong to $L^\infty(I)$, since we have $\lim_{t\to 0} m(t) = 0$. As a consequence of condition (iii) the norm $\|w\|_{L^2(I)}$ has to be small enough with respect to L and $\|a^*\|_{L^1(I)}$ in order to ensure the convergence rate (25).

Note that the strength of the conditions (ii) and (iii) in the case $X \neq K$ also depends on the rate of growth of $\frac{1}{m(t)} \to \infty$ as $t \to 0$. If we restrict the reference functions $\overline{a}$ by condition (22), then we have an exponential growth rate indicating an extremely ill-posed behavior of (SIP), but localized in a neighborhood of $t = 0$. Namely, from formula (12) we derive

$$\frac{1}{m(t)} = C\,\sqrt{S^*(t)}\,\exp(\psi(t)) \quad (0 < t < T)$$

with a constant $C > 0$ and

$$\psi(t) = \frac{\nu^2}{2S^*(t)} + \frac{r^2\,t^2}{2S^*(t)} + \frac{\nu r\,t}{S^*(t)} + \frac{\nu}{2} + \frac{r\,t}{2} + \frac{S^*(t)}{8}, \quad \nu := \ln\left(\frac{X}{K}\right) \neq 0.$$

With $a^* \in \widetilde{D}(F)$ and formula (35) we obtain a constant $\overline{\overline{c}}$ depending on w and $\overline{c}$ with $a^*(\tau) \leq \overline{\overline{c}}$ a.e. in I. Hence we derive $\underline{c}\,t \leq S^*(t) \leq \overline{\overline{c}}\,t$ $(t \in I)$. This implies the estimates

$$\widetilde{C}\,\sqrt{t}\,\exp\left(\frac{\nu^2}{2\,\overline{\overline{c}}\,t}\right) \leq \frac{1}{m(t)} \leq \hat{C}\,\sqrt{t}\,\exp\left(\frac{\nu^2}{2\,\underline{c}\,t}\right) \qquad (0 < t \leq T) \tag{37}$$

below and above with positive constants $\widetilde{C}$ and $\hat{C}$.

4 The singular case

Analyzing the structure of the second partial derivative $\frac{\partial^2 U_{BS}(X,K,r,\tau,s)}{\partial s^2}$ in formula (13) for $X \neq K$ we find that the upper bound C_2 defined in Theorem 3.5 and hence $L = \sqrt{T}\,C_2$ grow to infinity as $X - K \to 0$. The limit case $X = K$ of *at-the-money options*, however, represents a singular situation (cf. also [9, §3]), since then M fails to be a bounded linear operator in $L^2(I)$ due to $\lim_{t\to 0} m(t) = \infty$ (see formula (37) for $\nu = \ln\left(\frac{X}{K}\right) = 0$). Consequently, G defined by the formulae (31) – (32) is not necessarily a continuous operator from $L^1(I)$ to $L^2(I)$ and the properties of this operator G may vary when the point $a^* \in \widetilde{D}(F) \subset L^1(I)$ changes. We formulate this variation in the following more in detail:

Theorem 4.1 *In the case $X = K$ the operator G defined by the formulae (31) – (32) is a bounded linear operator from $L^1(I)$ to $L^2(I)$ if $a^* \in \widetilde{D}(F)$ satisfies the condition*

$$a^*(\tau) \;\geq\; c\,\tau^{-\beta} \qquad (\tau \in I) \tag{38}$$

for a fixed constant $0 < \beta < 1$. On the other hand, that linear operator G is unbounded for

$$a^* \in \widetilde{D}(F) \cap L^\infty(I). \tag{39}$$

Proof: For $X = K$ the formula (12) attains the form

$$\frac{\partial U_{BS}(X,K,r,\tau,s)}{\partial s} \;=\; \frac{X}{2\sqrt{2\,\pi\,s}}\,\exp\left(-\frac{r^2\,\tau^2}{2\,s} - \frac{r\,\tau}{2} - \frac{s}{8}\right). \tag{40}$$

The following considerations are based on this formula. If $a^* \in \widetilde{D}(F)$ satisfies the condition (38), then we can estimate with $S^* = J(a^*)$ and some positive constants $C, \widetilde{C}$ and $\widetilde{\widetilde{C}}$:

$$\|G(h)\|^2_{L^2(I)} \leq C \int_I \frac{1}{S^*(t)} \left(\int_0^t h(\tau)\, d\tau \right)^2 dt \leq \widetilde{C} \left(\int_I \frac{dt}{t^{1-\beta}} \right) \|h\|^2_{L^1(I)} \leq \widetilde{\widetilde{C}} \|h\|^2_{L^1(I)}. \quad (41)$$

This proves the continuity of G in that case. On the other hand, for a^* satisfying (39) and consequently $\underline{c}\, t \leq S^*(t) \leq \overline{c}\, t$ we consider the sequences of functions

$$h_n(\tau) = \begin{cases} 2n(1 - n\tau) & (0 \leq \tau \leq 1/n) \\ 0 & (1/n < \tau \leq T) \end{cases} \qquad \text{with} \quad \|h_n\|_{L^1(I)} = 1$$

and

$$[J(h_n)](t) = \begin{cases} 2nt - n^2 t^2 & (0 \leq t \leq 1/n) \\ 1 & (1/n < t \leq T). \end{cases}$$

Then we derive from the structure of $J(h_n)$

$$\lim_{n\to\infty} \|G(h_n)\|^2_{L^2(I)} = \lim_{n\to\infty} \int_I m^2(t)\, [J(h_n)]^2(t)\, dt = \infty \quad \text{if} \quad m \notin L^2(I).$$

Now we have in that case with positive constants $C, \widetilde{C}$ and K

$$\int_I m^2(t)\, dt = C \int_I \frac{1}{S^*(t)} \exp\left(\frac{r^2 t^2}{S^*(t)} - rt - \frac{S^*(t)}{4} \right) dt$$

$$\geq \widetilde{C} \int_I \frac{\exp(-Kt)}{t}\, dt \;\geq\; \widetilde{C} \exp(-KT) \int_I \frac{dt}{t} = \infty.$$

Thus the operator G is unbounded in that situation. ∎

As Theorem 4.1 indicates, the operator F fails to be Gâteaux-differentiable at any point $a^* \in \widetilde{D}(F) \cap L^\infty(I)$. Then a condition

$$\|F(a) - F(a^*) - G(a - a^*)\|_{L^2(I)} \leq \frac{L}{2} \|a - a^*\|^2_{L^1(I)} \quad (42)$$

for all $a \in \widetilde{D}(F)$ requiring even the existence of a Fréchet derivative G cannot hold at such a point. Although the Gâteaux derivative G exists as a bounded linear operator from $L^1(I)$ to $L^2(I)$ in the special case of elements a^* with a weak pole satisfying (38) we disbelieve the existence of a constant $0 < L < \infty$ such that (42) is valid for all $a^* \in \widetilde{D}(F)$. However, we have no stringent proof of that fact.

After all we note that, for $X = K$, the operator G according to (31) – (32) mapping from $L^2(I)$ to $L^2(I)$ as considered in [9, §5] is continuous for all $a^* \in L^2(I) \cap \widetilde{D}(F)$. Namely, for $h \in L^2(I)$ with $\left(\int_0^t h(\tau)\, d\tau \right)^2 \leq t\, \|h\|^2_{L^2(I)}$ as a consequence of the Schwarz inequality we estimate with positive constants $\hat{C}$ and $\hat{\hat{C}}$ and $S^*(t) \ge \underline{c}\, t$ $(t \in I)$ in analogy to (41):

$$\|G(h)\|^2_{L^2(I)} \le \hat{C} \left(\int_I \frac{t}{S^*(t)}\, dt \right) \|h\|^2_{L^2(I)} \le \hat{\hat{C}} \|h\|^2_{L^2(I)}.$$

But we also conjecture that an inequality

$$\|F(a) - F(a^*) - G(a - a^*)\|_{L^2(I)} \le \frac{L}{2} \|a - a^*\|^2_{L^2(I)} \quad \text{for all} \quad a \in L^2(I) \cap \widetilde{D}(F),$$

cannot hold.

5 Numerical case studies

In [9] we find a case study concerning the solution of (SIP) using a discrete version of a second order Tikhonov regularization (TR) approach with solutions of the extremal problem

$$\|F(a) - u^\delta\|^2_{L^2(I)} + \alpha \|a''\|^2_{L^2(I)} \longrightarrow \min, \qquad \text{subject to} \qquad a \in D(F) \cap H^2(I). \tag{43}$$

Here, we compare this approach with the results of a discrete version of maximum entropy regularization (MER) with respect to the character and quality of regularized solutions. In particular, we tried to find situations in which maximum entropy regularization performs better than Tikhonov regularization. In the paper [1] the applicability of these two regularization methods for the classical moment problem was studied. The authors concluded that Tikhonov regularization of second order is superior if the solution is smooth, whereas maximum entropy regularization leads to better results if the solution has sharp peaks.

We compared for $T = 1$ the behavior of the maximum entropy regularization (MER) according to (20) and of the second order Tikhonov regularization (TR) according to (43) implemented in a discretized form for the convex and rather smooth volatility function

$$a_1^*(\tau) = ((\tau - 0.5)^2 + 0.1)^2 \qquad (0 \le \tau \le 1)$$

in a first study and for the volatility function

$$a_2^*(\tau) = 0.1 + \frac{0.9}{1 + 100(2\tau - 1)^2} \qquad (0 \le \tau \le 1)$$

with a sharp peak at the point $\tau = 0.5$ in a second study.

For our case studies we approximated the functions a by the vector $\underline{a} = (a_1, \ldots, a_N)^T$ with $a_i = a\left(\frac{2i-1}{2N}\right)$. Using the values $X = 0.6$, $K = 0.5$, $r = 0.05$ we computed the exact option price data $u_j^* = u^*(t_j)$ at time $t_j := \frac{j}{N}$ $(j = 1, \ldots, N)$ according to formula (7). Perturbed by a random noise vector $\underline{\eta} = (\eta_1, \ldots, \eta_N)^T \in \mathbb{R}^N$ with normally distributed components $\eta_i \sim \mathcal{N}(0,1)$, which are i.i.d. for $i = 1, 2, ..., n$, the vector $\underline{u}^* := (u_1^*, \ldots, u_N^*)^T$ gives the noisy data vector $\underline{u}^\delta$ with components

$$u_j^\delta := u_j^* + \delta \frac{\|\underline{u}^*\|_2}{\|\underline{\eta}\|_2} \eta_j \qquad (j = 1, \ldots, N),$$

where $\|\cdot\|_2$ denotes the Euclidean vector norm. The discrete regularized solutions $\underline{a}_\alpha^\delta = \left((a_\alpha^\delta)_1, \ldots, (a_\alpha^\delta)_N\right)^T$ were determined as minimizers of discrete analogs of the extremal problems (23) and (43) and the accuracy of approximate solutions was measured in the Manhattan norm $\|\underline{a}_\alpha^\delta - \underline{a}^*\|_1 = \sum_{j=1}^{N} |(a_\alpha^\delta)_j - a_j^*|$, which is proportional to the discrete L^1-norm.

In the first case study with exact solution a_1^* we used the discretization level $N = 20$ and in both studies and all figures the noise level $\delta = 0.001$. Figure 1 shows the unregularized solution ($\alpha = 0$). Small data errors cause significant perturbations in the least-squares solution. Therefore a regularization seems to be necessary. For the maximum entropy regularization (MER) with reference function

$$\overline{a}_1(\tau) := (0.5 \cdot (\tau - 0.5)^2 + 0.16)^2 \qquad (0 \le \tau \le 1)$$

the error $\|\underline{a}_\alpha^\delta - \underline{a}_1^*\|_1$ as a function of the regularization parameter $\alpha > 0$ is sketched in figure 2. Figure 3 shows for the same reference function the best solution obtained by maximum entropy regularization minimizing the discrete L^1-error norm over all regularization parameters $\alpha > 0$. Figure 4 displays alternatively the best regularized solution computed by second order Tikhonov regularization (TR).

Regularization method	Error $\|\underline{a}_{\alpha_{\text{opt}}}^\delta - \underline{a}_1^*\|_1$
(TR)	0.0560
(MER) with $\overline{a}_1$	0.0896
(MER) with $\overline{a}_2$	0.1196

Table 1: Accuracy of best regularized solutions for smooth solution

Table 1 compares the errors $\|\underline{a}_\alpha^\delta - \underline{a}_1^*\|_1$ for the best regularized solutions $\underline{a}_\alpha^\delta$ obtained by (TR) and by (MER) with the two different reference functions $\overline{a}_1$ as defined above and

$$\overline{a}_2(\tau) := 0.07 \qquad (0 \le \tau \le 1).$$

It can be concluded, that for the smooth function a_1^* to be recovered the best solutions were obtained by Tikhonov regularization of second order. Hence, for the test function a_1^* the smoothness information, which is used by Tikhonov regularization of second order, seems to be more appropriate than the information about the shape of a^*, which is reflected by the reference function $\overline{a}_1$ in the context of maximum entropy regularization. In absence of any information about the shape of a^* one has to use a constant reference function, for example $\overline{a}_2$, which does not provide acceptable regularized solutions here.

In our second case study with exact solution a_2^* we used $N = 50$ and compared (TR) and (MER) with the well-approximating reference function

$$\overline{a}_3(\tau) = 0.12 + 0.5/(1 + 100(2\tau - 1)^2) \qquad (0 \le \tau \le 1)$$

and with the constant reference function

$$\overline{a}_4(\tau) := 0.5 \qquad (0 \le \tau \le 1).$$

The figures 5, 6, 7, 8 show the unregularized solution, the best regularized solution obtained by (TR) and by (MER) with reference function $\overline{a}_3$ and $\overline{a}_4$, respectively. The best possible errors $\|\underline{a}_\alpha^\delta - \underline{a}_2^*\|_1$ are compared in table 2. Here the maximum entropy results with the reference function $\overline{a}_3$ are better than the results obtained with (TR). On the other hand, (MER) with reference function $\overline{a}_4$ leads to a regularized solution, which approximates the exact solution a_2^* quite well for $0 < \tau < 0.95$, but deviates at the right end of the interval, because for $\tau \approx 1$ the data information decreases (cf. formulae (6), (7) and (36)) and the influence of the reference function dominates.

Regularization method	Error $\|\underline{a}_{\alpha_{\mathrm{opt}}}^\delta - \underline{a}_2^*\|_1$
(TR)	0.6429
(MER) with $\overline{a}_3$	0.2694
(MER) with $\overline{a}_4$	0.7169

Table 2: Accuracy of best regularized solutions for peak solution

It should be remarked that the peak shape of the function a_2^* is well-recovered without regularization, whereas the smooth parts of the function away from the peaks require a regularization.

In conclusion, we can say that our investigations have confirmed the results of the article [1]. For determining a smooth exact solution maximum entropy regularization was inferior to Tikhonov regularization of second order, whereas it was superior in the peak case, provided we used an appropriate reference function.

In a final consideration we checked the convergence rates of (MER) solutions for $\delta \to 0$ in the peak case of an exact solution a_2^*. With $N = 50$ and the reference functions $\overline{a}_3$ and $\overline{a}_4$ we compared the $\|.\|_1$-errors for noise levels $\delta_1 = 0.01$, $\delta_2 = 0.001$ and, $\delta_3 = 0.0001$ (see tables 3 and 4). Although the reference functions both do not satisfy the condition (36), the convergence rate seems to be nearly proportional to $\sqrt{\delta}$. However, for the less informative reference function $\overline{a}_4$ the absolute error levels are significantly larger than for $\overline{a}_3$ and smaller optimal regularization parameters α_{opt} occur.

δ	α_{opt}	$\|\underline{a}_{\alpha_{\mathrm{opt}}}^\delta - \underline{a}^*\|_1$
0.01	0.0118	0.7719
0.001	0.00209	0.2694
0.0001	0.000292	0.0770

Table 3: Errors of optimal (MER) solutions for a_2^* with $\overline{a}_3$ depending on δ

Acknowlegdgements

The authors are grateful to Jin Qi-nian (Rutgers University, New Brunswick, USA) for his helpful remarks concerning the cited results on convergence rates.

δ	α_{opt}	$\|\underline{a}^{\delta}_{\alpha_{\text{opt}}} - \underline{a}^*\|_1$
0.01	0.0035	2.396
0.001	0.00034	0.717
0.0001	0.000044	0.1555

Table 4: Errors of optimal (MER) solutions for a_2^* with $\overline{a}_4$ depending on δ

References

[1] Amato, U.; Occorsio, M.R.: Influence of the smoothing functional in the approximation of first kind Fredholm integral equations. *Atti. Semin. Mat. Fis. Univ. Modena* **40** (1992), 577–590.

[2] Borwein, J.M.; Lewis, A.S.: Convergence of best entropy estimates. *SIAM J. Optimization* **2** (1991), 191-205.

[3] Bouchouev, I.; Isakov, V.: Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. Inverse Problems **15** (1999), R95–R116.

[4] Eggermont, P.P.B.: Maximum Entropy Regularization for Fredholm Integral Equations of the first kind. *SIAM J. Math. Anal.* **6** (1993), 1557-1576.

[5] Engl, H.W.; Hanke, M.; Neubauer, A.: *Regularization of Inverse Problems.* Dordrecht: Kluwer 1996.

[6] Engl, H.W.; Kunisch, K; Neubauer, A.: Convergence rates for Tikhonov regularisation of non-linear ill-posed problems. *Inverse Problems* **5** (1989), 523–540.

[7] Engl, H.W.; Landl, G.: Maximum entropy regularization of nonlinear ill-posed problems. In: V. Lakshmikantham (ed.), *Proceedings of the First World Congress on Nonlinear Analysts*, Vol. 1. Berlin: de Gruyter 1996, 513–525.

[8] Gorenflo, R.; Vessella, S.: *Abel Integral Equations.* Berlin: Springer 1991.

[9] Hein, T.; Hofmann, B.: On the nature of ill-posedness of an inverse problem arising in option pricing. *Inverse Problems* **19** (2003), 1319–1338.

[10] Hofmann, B.; Scherzer, O.: Factors influencing the ill-posedness of nonlinear problems. *Inverse Problems* **10** (1994), 1277–1297.

[11] Hull, J.C.: *Options, Futures & Other Derivatives.* Fourth Edition. London: Prentice-Hall International 2000.

[12] Jin Qi-nian; Hou Zong-yi: Maximum entropy method for solving nonlinear ill-posed problems. *Chinese Science Bulletin* **41** (1996), 1589–1593.

[13] Jin Qi-nian; Hou Zong-yi: Maximum entropy method for solving nonlinear ill-posed problems.II (Chinese with English summary). *J. Fudan Univ., Nat. Sci.* **36** (1997), 679–684.

[14] Krasnoselskii, M.A: *Asymptotics of Nonlinearities and Operator Equations.* Basel: Birkhäuser 1995.

[15] Kwok, Y.K.: *Mathematical Models of Financial Derivatives.* Singapore: Springer 1998.

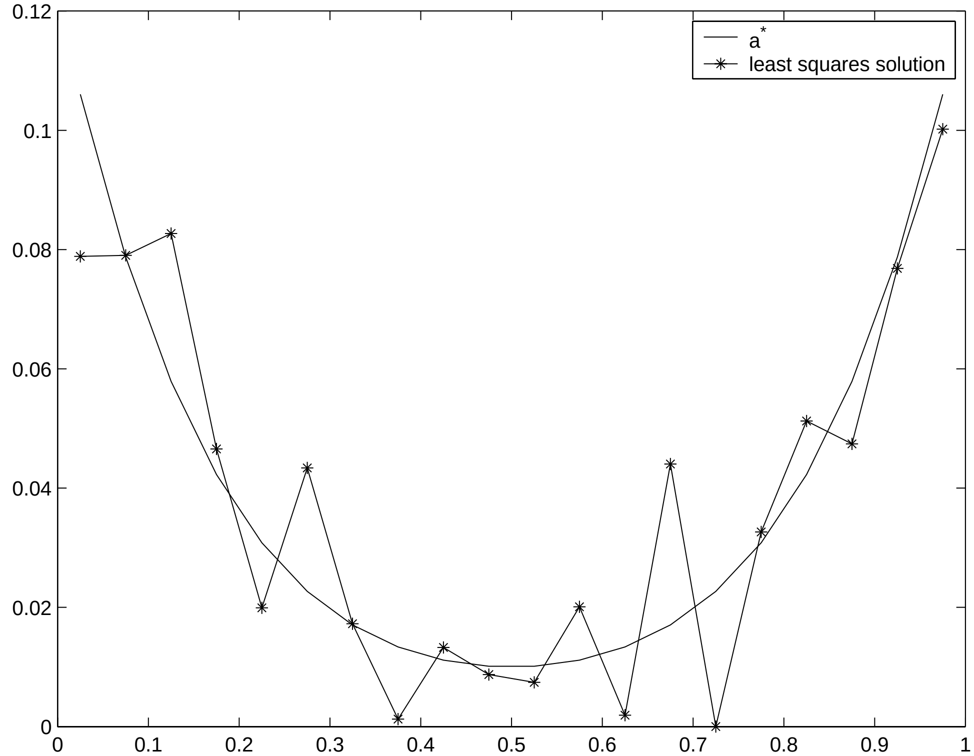

Figure 1: Unregularized least-squares solution for a_1^* and $\delta = 0.001$

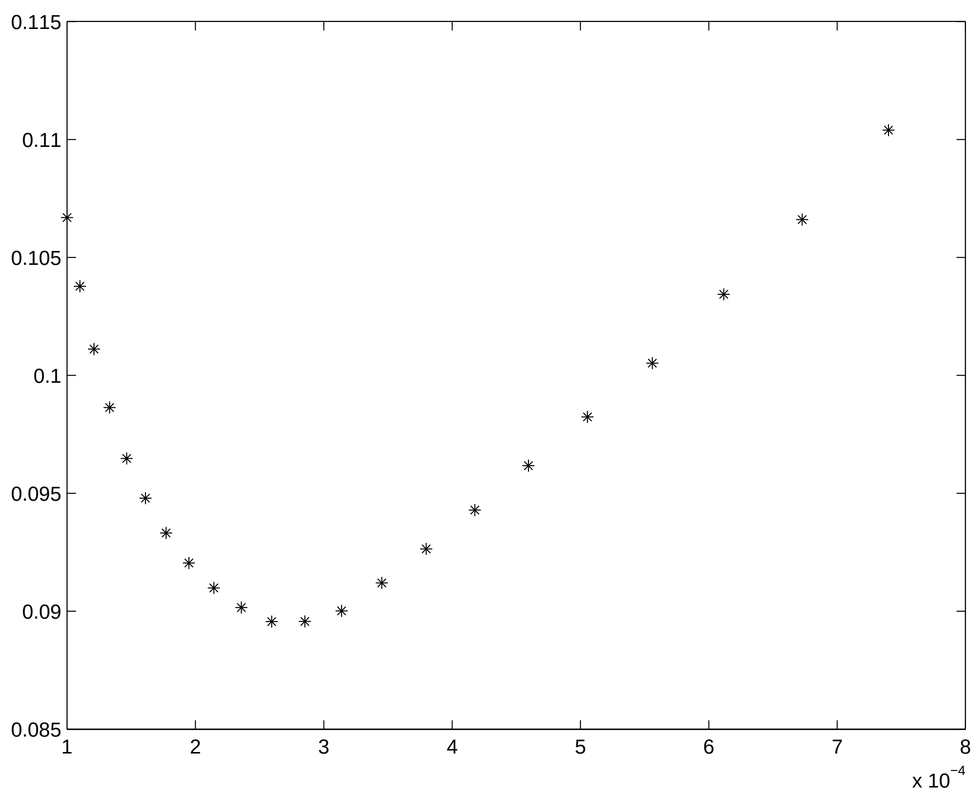

Figure 2: Regularization error $\|\underline{a}_\alpha^\delta - \underline{a}_1^*\|_1$ of (MER) for a_1^* depending on α

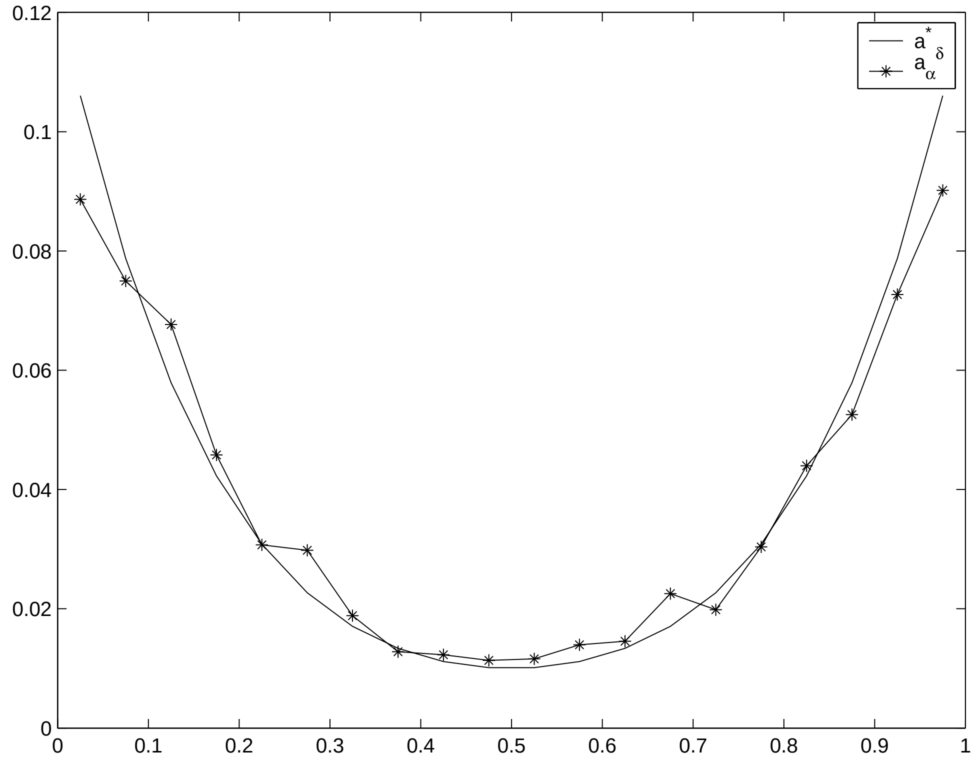

Figure 3: Optimal solution of (MER) for a_1^* with $\overline{a}_1$

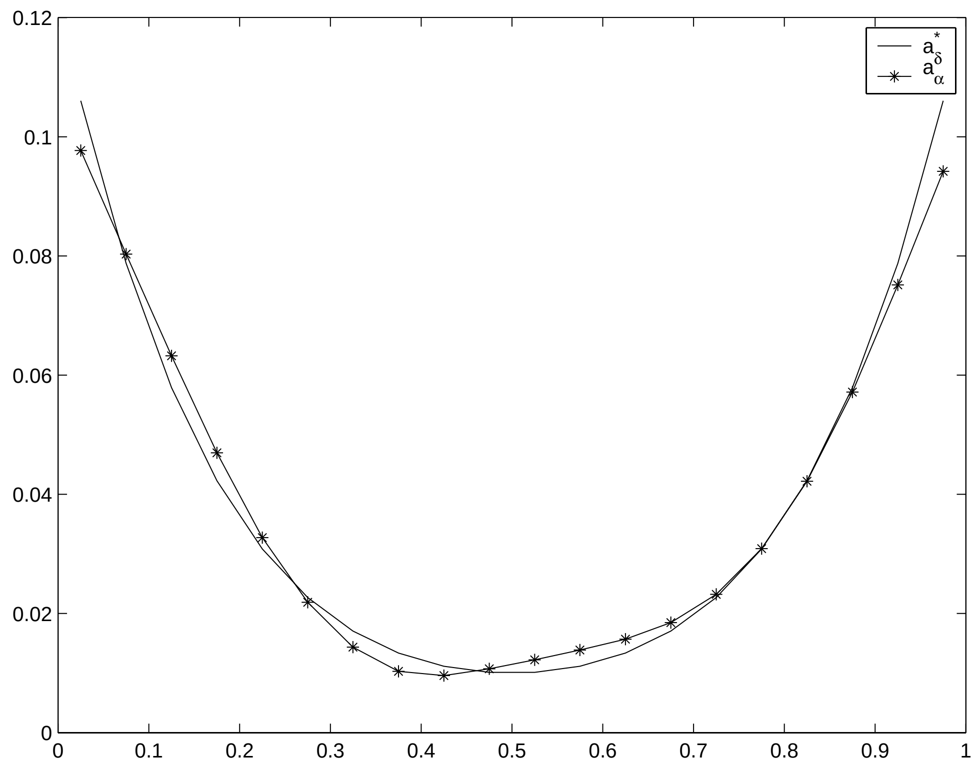

Figure 4: Optimal solution of (TR) for a_1^*

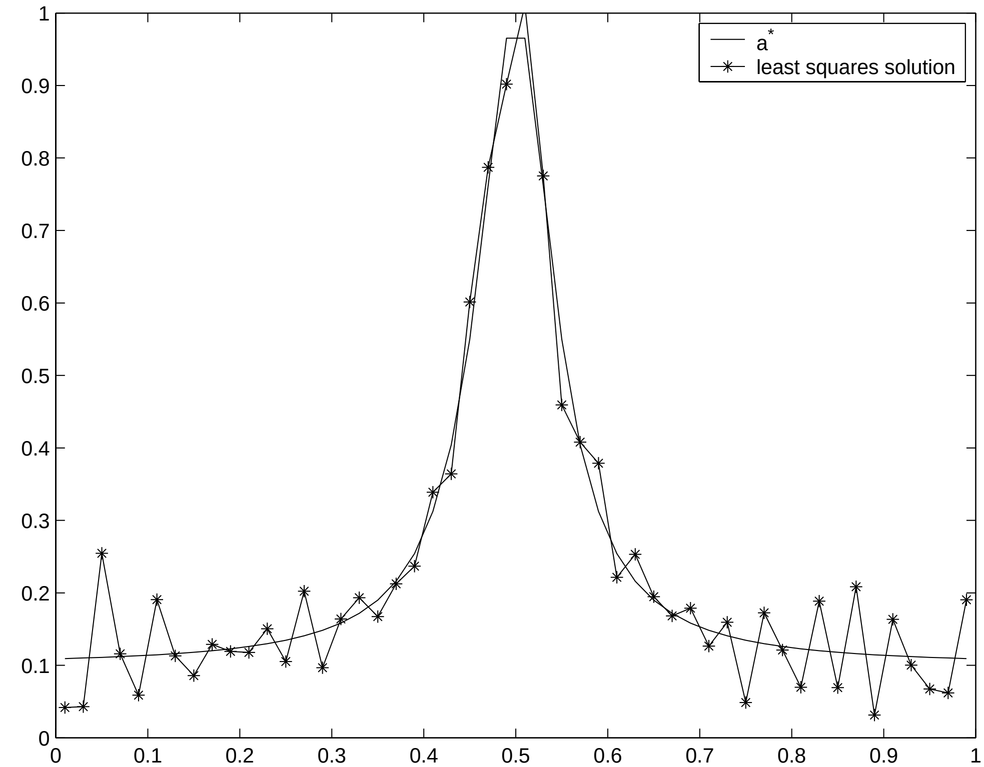

Figure 5: Unregularized least-squares solution for a_2^* and $\delta = 0.001$

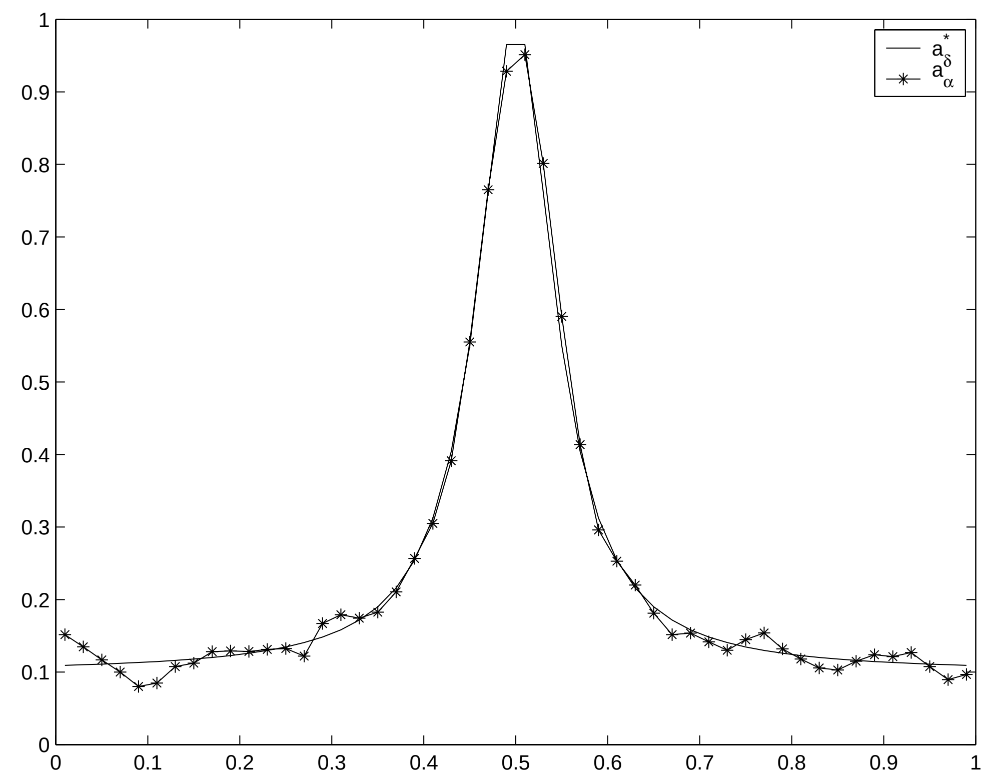

Figure 6: Optimal solution of (TR) for a_2^*

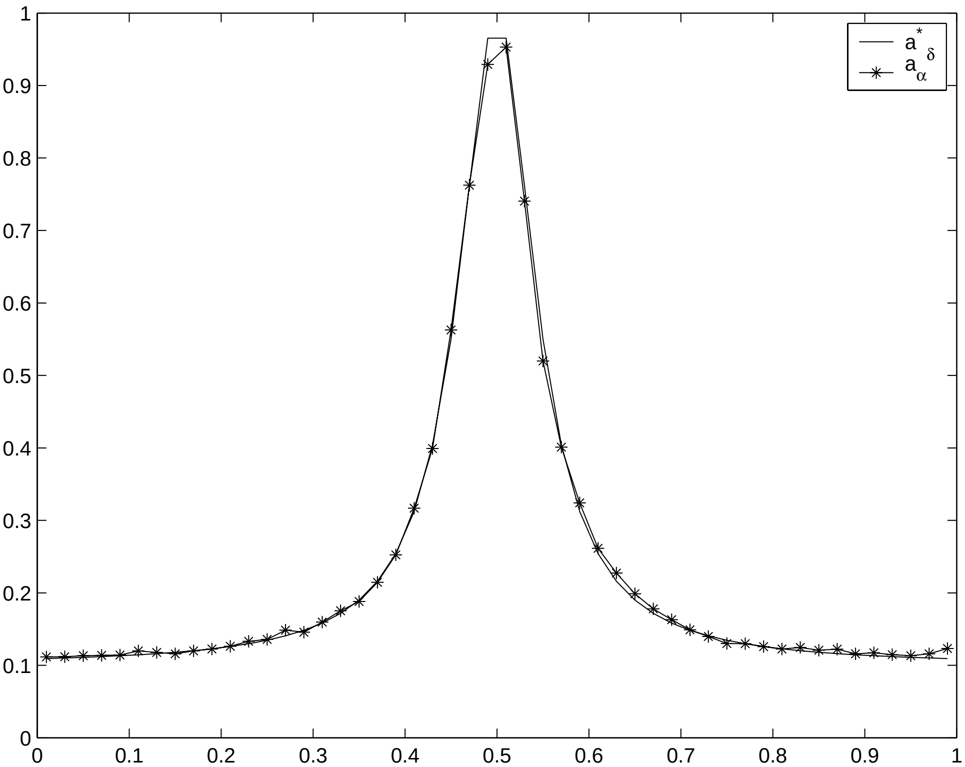

Figure 7: Optimal solution of (MER) for a_2^* with $\overline{a}_3$

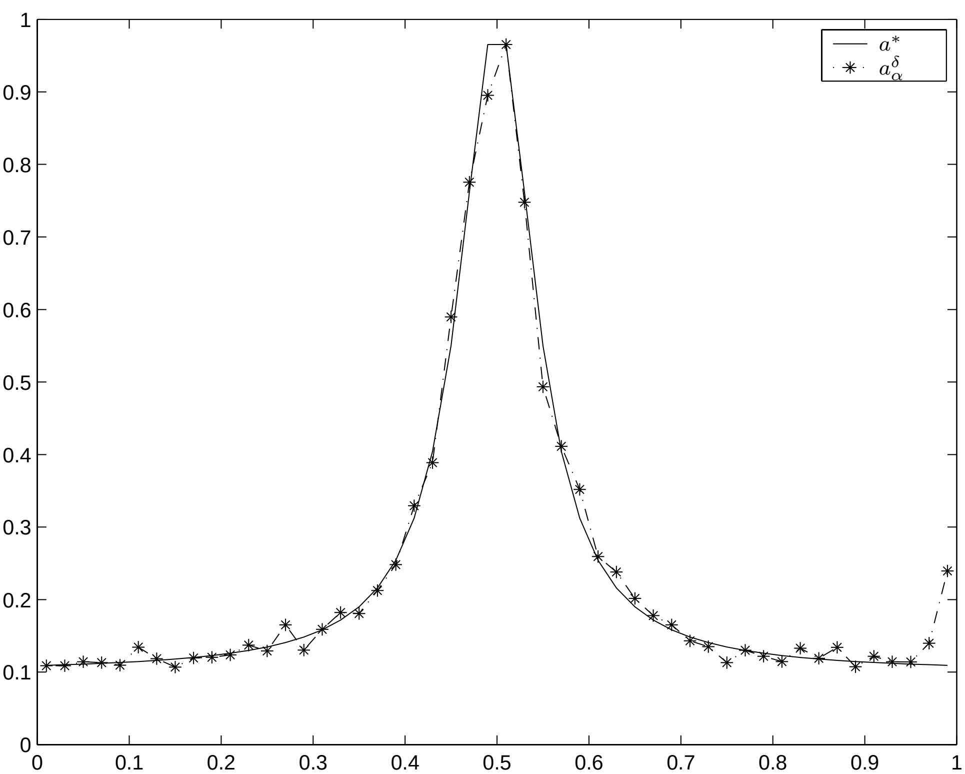

Figure 8: Optimal solution of (MER) for a_2^* with $\overline{a}_4$