A multi-parameter regularization approach for estimating parameters in jump diffusion processes

DANA DÜVELMEYER* AND BERND HOFMANN †

Abstract

In this paper, we consider the inverse problem of estimating simultaneously the five parameters of a jump diffusion process based on return observations of a price trajectory. We show that there occur some ill-posedness phenomena in the parameter estimation problem, because the forward operator fails to be injective and small perturbations in the data may lead to large changes in the solution. We illustrate the instability effect by a numerical case study. To obtain stable approximate solutions of the estimation problem, we use a multi-parameter regularization approach, where a least-squares fitting of empirical densities is superposed by a quadratic penalty term of fitted semi-invariants with weights. A little number of required weights is controlled by a discrepancy principle. For the realization of this control, we propose and justify a fixed point iteration, where an exponent can be chosen arbitrarily positive. A numerical case study completing the paper shows that the approach provides satisfactory results and that the amount of computation can be reduced by an appropriate choice of the free exponent.

MSC2000 subject classification: 65J20, 62F10, 91B84

Keywords: Parameter estimation, jump diffusion processes, ill-posedness, regularization, fixed point iteration

*Department of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, GERMANY. Email: dana.duevelmeyer@mathematik.tu-chemnitz.de.

†Department of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, GERMANY. Email: hofmannb@mathematik.tu-chemnitz.de. Corresponding author.

1 Introduction

For modeling the time-dependent stochastic behavior of prices of stocks or stock indices jump diffusion processes are rather helpful. Such processes are able to close some gaps between the mathematical model and observed market phenomena occurring when a geometric Brownian motion is used as price process (see, e.g., [13, Chapter 9]). However, the number of parameters to be determined grows from two to five if we replace the geometric Brownian motion by a jump diffusion process. The aim of this paper is to analyze the parameter estimation problem and its properties for a jump diffusion model introduced below.

In accordance with [13] we call a stochastic process $(S_t, t \in [0, \infty))$ a jump diffusion process if it is satisfying the stochastic differential equation

$$dS_t = S_t((\mu - \lambda\nu)dt + \sigma dW_t) + S_{t^-} dN_t^c$$

provided that $(W_t, t \in [0, \infty))$ is a standard Wiener process, $(N_t, t \in [0, \infty))$ is a Poisson process with intensity λ and $(N_t^c, t \in [0, \infty))$ is a compound Poisson process associated to $(N_t, t \in [0, \infty))$ with jump amplitude $(Y_j - 1)$ and expectation $\nu = \mathrm{E}\{Y_j - 1\}$. If T_j denotes a jump time we have $N^c_{T_j^+} = N^c_{T_j^-} + (Y_j - 1)$ and hence it yields $S_{T_j^+} = S_{T_j^-} + S_{T_j^-}(Y_j - 1) = S_{T_j^-} Y_j$ for the jump diffusion process. The processes $(W_t, t \in [0, \infty))$, $(N_t, t \in [0, \infty))$ and the jumps $(Y_j)_{j \geq 1}$ are mutually independent. Additionally, we assume $\log Y_j \sim N(\mu_Y, \sigma_Y^2)$ such that $\nu = e^{\mu_Y + \frac{1}{2}\sigma_Y^2} - 1$. Consequently we have two diffusion parameters μ and σ describing drift and volatility of the geometric Brownian motion and three jump parameters. The parameter λ is specifying the number of jumps whereas the parameters μ_Y and σ_Y are determining mean and volatility of them.

Our inverse problem is to estimate from observed process data the five scalar parameters $\mu \in \mathbb{R}, \sigma > 0, \lambda \geq 0, \mu_Y \in \mathbb{R}$ and $\sigma_Y \geq 0$, which we collect in the vector $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)^T \in \mathbb{R}^5$. This vector completely determines the assumed price dynamics. After some considerations concerning the forward operator we will formulate the inverse problem more precisely as an operator equation at the the end of this section. For inverse problems in the context of stochastic considerations see also in general [7], [8, Chapter 5] and [11, Section 4.1.6]. To estimate $\underline{p} \in D$ with an assumed domain

$$D = \{\underline{p} \in \mathbb{R}^5 : \sigma > 0, \lambda \geq 0, \sigma_Y \geq 0\} \tag{1}$$

we observe values $S_{t_0}, S_{t_1}, ..., S_{t_n}$ of the price process under consideration with an appropriate time step $\tau > 0$ and $t_i = t_0 + i\tau$ $(i = 0, 1, 2, ..., n)$. In particular, we use the logarithmic returns $r_{\tau,i} = \log\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right)$ $(i = 1, 2, ..., n)$ as data for fitting $\underline{p}$.

The paper is organized as follows: We begin with a brief discussion of the returns and their stochastic properties and introduce the operator of the forward problem. In section 2 we illustrate some ill-posedness phenomena which occur by solving the inverse problem numerically and discus reasons for the instability. To overcome this instability of the conventional least-squares fitting we propose in section 3 a multi-parameter regularization approach. It is shown that such an approach can help to estimate the unknown parameter vector of jump diffusion in a stable manner provided that realistic error bounds for the semi-invariants can be prescribed. We illustrate its applicability by a numerical case study with synthetic data in section 4. On the other hand, in section 5 we propose a modification of the multi-parameter algorithm based on some exponent variation. A numerical example, which shows that appropriately chosen exponents can reduce the number of iterates and hence the total amount of computation, completes the paper.

Our considerations to estimate a parameter vector $\underline{p}$ of the jump diffusion process are based on the logarithmic returns. By using the generalized Itô-calculus (see [14]) for semi-martingales we obtain after some computations:

Proposition 1 *The natural logarithm of the price process* $\log S_t$ *fulfills the stochastic differential equation*

$$d(\log S_t) = \tilde{\mu} dt + \sigma dW_t + d\widetilde{N}_t^c\,,$$

where $\tilde{\mu} = (\mu - \lambda\nu - \frac{1}{2}\sigma^2)$ *and* $(\widetilde{N}_t^c, t \in [0, \infty))$ *denotes a compound Poisson process associated to* $(N_t, t \in [0, \infty))$ *with jump amplitude* $\log Y_j$.

From proposition 1 we directly derive the structure

$$r_\tau = \log\left(\frac{S_\tau}{S_0}\right) = (\tilde{\mu}\tau + \sigma W_\tau) + \sum_{j=1}^{N_\tau} \log Y_j$$

of logarithmic returns. The stationarity of the returns r_τ expressed by the equality in distribution of r_τ and

$$\begin{aligned} r_\tau^x &:= \log \frac{S_{x+\tau}}{S_x} = \log S_{x+\tau} - \log S_0 + \log S_0 - \log S_x \\ &= r_{x+\tau} - r_x = \tilde{\mu}\tau + \sigma(W_{x+\tau} - W_x) + \sum_{j=N_x+1}^{N_{x+\tau}} \log Y_j \end{aligned}$$

follows directly from the stationarity of the increments of a Wiener process and a Poisson process for a fixed time difference τ, since r_τ^x has the same distribution as $\tilde{\mu}\tau + \sigma(W_{x+\tau} - W_x) + \sum_{j=1}^{N_{x+\tau}-N_x} \log Y_j$. By applying the law of total probability we express the distribution function as

$$\begin{aligned} F(x, \underline{p}) = \mathrm{P}(r_\tau \leq x) &= \sum_{j=0}^{\infty} \mathrm{P}(N_\tau = j)\, \mathrm{P}(r_\tau \leq x \mid N_\tau = j) \\ &= \sum_{j=0}^{\infty} \frac{e^{(-\lambda\tau)}(\lambda\tau)^j}{j!} \Phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right). \end{aligned}$$

Consequently, the density function attains the form

$$f(x, \underline{p}) = \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!\sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right),$$

where $\Phi(x) = \int_{-\infty}^{x} \phi(z)\, dz$ and $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ denote distribution and density function of the standard normal distribution. For all $\underline{p} \in D$ with D from (1) we have

$$f(x, \underline{p}) \leq \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!\sqrt{\sigma^2\tau}} \frac{1}{\sqrt{2\pi}} = \frac{1}{\sigma\sqrt{2\pi\tau}} < \infty$$

and hence $f(\,\cdot\,, \underline{p})$ is in $L^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$.

By applying the Lévy-Khintchine formula (see [16, p. 195]) we get the characteristic function

$$\begin{aligned} \varphi\left(\theta, \underline{p}\right) &= \mathrm{E}\, e^{(i\theta r_\tau)} \\ &= \exp\left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\tau\theta^2 + \lambda\tau\left(\exp\left(-\frac{\sigma_Y^2}{2}\theta^2 + i\mu_Y\theta\right) - 1\right)\right) \end{aligned}$$

of the returns and can calculate the semi-invariants $s_{\tau,k}(\underline{p})$ (see [15, p. 289]) from the k-th derivative

$$s_{\tau,k}(\underline{p}) = \frac{(\log \varphi)^{(k)}(0,\underline{p})}{i^k} .$$

This implies

$$\begin{aligned}
s_{\tau,1}(\underline{p}) &= (\lambda\mu_Y + \tilde{\mu})\tau \\
s_{\tau,2}(\underline{p}) &= ((\sigma_Y^2 + \mu_Y^2)\lambda + \sigma^2)\tau \\
s_{\tau,3}(\underline{p}) &= (3\sigma_Y^2 + \mu_Y^2)\lambda\tau\mu_Y \\
s_{\tau,4}(\underline{p}) &= (3\sigma_Y^4 + 6\sigma_Y^2\mu_Y^2 + \mu_Y^4)\lambda\tau \\
s_{\tau,5}(\underline{p}) &= (15\sigma_Y^4 + 10\sigma_Y^2\mu_Y^2 + \mu_Y^4)\lambda\tau\mu_Y \\
&\vdots \qquad\qquad\qquad \vdots
\end{aligned} \tag{2}$$

We use the following general relation (see [15, p. 290]) to compute the moments from the semi-invariants or vice versa.

Proposition 2 *Let ξ be a random variable with $\mathrm{E}\,|\xi|^n < \infty$. Then for all $k \leq n$ we have for the interplay of the first k-th moments $m_k = \mathrm{E}\xi^k$ and the first k-th semi-invariants s_k of ξ*

$$m_k = \sum_{\lambda^{(1)}+\ldots+\lambda^{(q)}=k} \frac{1}{q!}\frac{k!}{\lambda^{(1)}!\cdots\lambda^{(q)}!}\prod_{\nu=1}^{q} s_{\lambda^{(\nu)}},$$

and

$$s_k = \sum_{\lambda^{(1)}+\ldots+\lambda^{(q)}=k} \frac{(-1)^{q-1}}{q}\frac{k!}{\lambda^{(1)}!\cdots\lambda^{(q)}!}\prod_{\nu=1}^{q} m_{\lambda^{(\nu)}} ,$$

where $\sum\limits_{\lambda^{(1)}+\ldots+\lambda^{(q)}=k}$ *denotes the summation over all ordered sets of natural numbers* $\{\lambda^{(\nu)}\ (\nu = 1, 2, ..., q)\}$ *with* $\sum\limits_{\nu=1}^{q}\lambda^{(\nu)} = k$.

From proposition 2 we can conclude as follows:

$$
\begin{aligned}
s_1 &= m_1 = \mathrm{E}\xi \\
s_2 &= m_2 - m_1^2 = \mathrm{E}(\xi - \mathrm{E}\xi)^2 = \mathrm{D}^2\xi \\
s_3 &= m_3 - 3m_2m_1 + 2m_1^3 = \mathrm{E}(\xi - \mathrm{E}\xi)^3 \\
s_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4 = \ldots \\
&= \mathrm{E}(\xi - \mathrm{E}\xi)^4 - 3s_2^2 \\
s_5 &= m_5 - 5m_4m_1 - 10m_3m_2 + 20m_3m_1^2 + 30m_2^2m_1 - 60m_2m_1^3 + 24m_1^5 \\
&= \ldots = \mathrm{E}(\xi - \mathrm{E}\xi)^5 - 10s_2s_3 \\
&\vdots \qquad\qquad\qquad \vdots
\end{aligned}
$$

Note that the relations above can also be used to calculate empirical semi-invariants from empirical moments or central moments.

It seems to be natural to estimate the parameter vector $\underline{p}$ of the jump diffusion process by conventional statistical techniques like the maximum likelihood method or the moment method. There are papers like [12] which are dealt with the parameter estimation in jump diffusion models and occurring pitfalls in the estimation process (see also [2]). In [12] the author emphasizes that it is invalid to use standard maximum likelihood procedures for estimating jump parameters. There are also serious mathematical and numerical problems if one tries to solve the nonlinear equations (2) with semi-invariants $s_{\tau,k}$ as given data and $\underline{p}$ as vector of unknown parameters. This approach is very similar to the statistical method of moments.

We would like to point out at that our inverse problem of determining $\underline{p}$ is closely related to the well-known Hamburger moment problem. For the ill-posedness of moment problems see, e.g., [1]. By the moment problem an unknown density function is to be determined from given moments, whereas we are only searching for a small number of intrinsic parameters of the density function. Nevertheless, some ill-posedness phenomena also occur in our inverse problem.

We assume $\underline{p}^* \in D$ to be the exact parameter vector to be determined and analyze the estimation problem by using methods of inverse problem theory in order to find approximate solutions $\underline{p}^\delta$ of $\underline{p}^*$, which stably depend on the vector $\underline{S}^\delta = (S^\delta_{t_0}, S^\delta_{t_1}, \ldots, S^\delta_{t_n})^T$ of noisy price data and associated returns $\underline{r}^\delta_\tau = (r^\delta_{\tau,1}, ..., r^\delta_{\tau,n})^T$.

Therefore we consider the empirical density function of the empirical returns belonging to the data $\underline{S}^\delta$ and choose that parameter vector $\underline{p}^\delta$ which minimizes the distance between the density function $f(\,\cdot\,,\underline{p}^\delta)$ and the empirical density function. We use the L^2-norm for measuring the distance between the two densities.

In this context, we define the operator $\mathcal{A}\,:\,\underline{p}\,\mapsto\,f$ of the forward problem which maps the parameter vector $\underline{p}\in D$ to the density function $f(\,\cdot\,,\underline{p})\;\in\;L^1(\mathbb{R})\cap L^\infty(\mathbb{R})$ using the series expansion

$$[\mathcal{A}(\underline{p})](x)=f(x,\underline{p})=\sum_{j=0}^{\infty}e^{-\lambda\tau}\frac{(\lambda\tau)^j}{j!\sqrt{\sigma^2\tau+j\sigma_Y^2}}\phi\left(\frac{x-(\tilde{\mu}\tau+j\mu_Y)}{\sqrt{\sigma^2\tau+j\sigma_Y^2}}\right)\quad(x\in\mathbb{R}). \tag{3}$$

Thus, the inverse problem can be written as the nonlinear operator equation

$$\mathcal{A}(\underline{p})=f\qquad(\underline{p}\in D\subset X:=\mathbb{R}^5,\;f\in\mathcal{A}(D)\subset Y)\,, \tag{4}$$

where $\mathcal{A}(D)$ denotes the range of the operator $\mathcal{A}$ and Y is a Banach space with norm $\|\cdot\|_Y$. We will focus on the situation where the well-defined nonlinear operator $\mathcal{A}$ is a mapping from X to $Y:=L^2(\mathbb{R})$ with domain (1). In [3] and [4] we have shown that this operator $\mathcal{A}$ is continuous for all parameter vectors $\underline{p}\in D$.

For fixed $\underline{p}$ the function $[\mathcal{A}(\underline{p})](x)$ is an infinite mixture of weighted density functions $g_j(x,\underline{p})$ of Gaussian variables with mean $\tilde{\mu}\tau+j\mu_Y$ and variance $\sigma^2\tau+j\sigma_Y^2$, that is

$$g_j(x,\underline{p})=\frac{1}{\sqrt{\sigma^2\tau+j\sigma_Y^2}}\phi\left(\frac{x-(\tilde{\mu}\tau+j\mu_Y)}{\sqrt{\sigma^2\tau+j\sigma_Y^2}}\right).$$

Furthermore, the weights $w_j=e^{-\lambda\tau}\frac{(\lambda\tau)^j}{j!}$ correspond with the probability that j jumps occur. Such a mixture of bell-shaped curves also appears at considering Fredholm integral equations in L^2 or C-spaces with continuous and bell-shaped kernels (see for example [11, example 2.3]). Even if our forward operator $\mathcal{A}$ maps from a finite dimensional space only and although we have an infinite sum instead of an integral, the mixture of bell-shaped curves (3) leads to smoothing properties which are comparable to those of integral operators mapping between two infinite dimensional spaces. In particular, they cause ill-conditioning phenomena which will be illustrated and discussed in section 2.

2 Ill-posedness phenomena of the inverse problem

The practical inverse problem of parameter estimation consists in finding appropriate approximations $\underline{p}^\delta$ of the true vector $\underline{p}^\star$ with $\mathcal{A}(\underline{p}^\star) = f^\star$ using the noisy data function $f^\delta \in Y$ with $\left\| f^\delta - f^\star \right\|_Y \le \delta$ and noise level $\delta > 0$. The noisy data are obtained from empirical density functions of the empirical returns $\underline{r}_\tau^\delta$.

Uniqueness and stability of solutions play an important role in the solution process. Therefore we will briefly discuss some properties of $\mathcal{A}$. The operator $\mathcal{A}$ is obviously not injective on D. To see this, we consider the parameter vectors $\underline{p}_1 = (\mu, \sigma, \lambda, 0, 0)^T$ and $\underline{p}_2 = (\mu, \sigma, 0, \mu_Y, \sigma_Y)^T$. Both vectors map to the same density function

$$[\mathcal{A}(\underline{p}_1)](x) = [\mathcal{A}(\underline{p}_2)](x) = \frac{1}{\sqrt{\sigma^2 \tau}} \phi \left(\frac{x - \tilde{\mu}\tau}{\sqrt{\sigma^2 \tau}} \right) \quad (x \in \mathbb{R}),$$

which is a normal density function, because the jump part is eliminated. In the case of $\underline{p}_1$ the jump size is always zero and in the second case of $\underline{p}_2$ jumps do not occur. However, this trivial case is the only example for non-injectivity. The following proposition is proven in [17].

Proposition 3 *The operator $\mathcal{A}$ is injective on the restricted domain*

$$\hat{D} = \left\{ \underline{p} \in \ D : \ \lambda \left(\sigma_Y^2 + \mu_Y^2 \right) \ \neq \ 0 \right\} .$$

Consequently the operator equation (4) is uniquely solvable if and only if a solution $\underline{p} \in \hat{D}$ exists. Moreover, the equation

$$\mathcal{A}(\underline{p}_1) = \mathcal{A}(\underline{p}_2) \tag{5}$$

can only hold for some $\underline{p}_1 \neq \underline{p}_2$ whenever $\underline{p}_1$ and $\underline{p}_2$ from (1) both belong to the set $D \setminus \hat{D}$, that means $\lambda_1 \, ({\mu_Y}_1^2 + {\sigma_Y}_1^2) = \lambda_2 \, ({\mu_Y}_2^2 + {\sigma_Y}_2^2) = 0$. Furthermore, it is easy to prove that the diffusion parameters coincide in every case $\mathcal{A}(\underline{p}_1) = \mathcal{A}(\underline{p}_2)$, even the operator $\mathcal{A}$ fails to be injective.

Proposition 4 *For a pair of vectors $\underline{p}_1 = (\mu_1, \sigma_1, \lambda_1, {\mu_Y}_1, {\sigma_Y}_1)^T \in D$ and $\underline{p}_2 = (\mu_2, \sigma_2, \lambda_2, {\mu_Y}_2, {\sigma_Y}_2)^T \in D$ with $\underline{p}_1 \neq \underline{p}_2$ satisfying (5) we have $\mu_1 = \mu_2$, $\sigma_1 = \sigma_2$ and $\underline{p}_1, \underline{p}_2 \in D \setminus \hat{D}$.*

The parameter vector $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)^T \in D \setminus \hat{D}$ is a solution of (4) if and only if $(\mu, \sigma, 0, \hat{\mu}_Y, \hat{\sigma}_Y) \in D \setminus \hat{D}$ and $(\mu, \sigma, \hat{\lambda}, 0, 0) \in D \setminus \hat{D}$ are solutions of (4) for arbitrary $\hat{\mu}_Y \in \mathbb{R}$, $\hat{\sigma}_Y \geq 0$ and $\hat{\lambda} \geq 0$. As a consequence of proposition 4 we can extend the domain of injectivity as follows.

Corollary 5 *The operator $\mathcal{A}$ is injective even on the domain*

$$\tilde{D} = \left\{\underline{p} \in \ D : \ \lambda \neq \ 0\right\} .$$

Since the operator equation (4) is not uniquely solvable for all density functions f from the range $\mathcal{A}(D)$, the equation (4) is ill-posed. We can easily find a sequence $\{f_n\}$ where $f_n = \mathcal{A}(\underline{p}_n)$ converges to $f_0 = \mathcal{A}(\underline{p}_0)$ in the norm of Y, but the sequence $\left\{\underline{p}_n\right\}$ with

$$\underline{p}_n \in U(f_n) := \left\{\underline{p} \in D : \ \mathcal{A}(\underline{p}) = f_n\right\}$$

does not converge to $\underline{p}_0 \in U(f_0)$ in X. However, we can show that under some conditions the diffusion parameters μ and σ of parameter vectors $\underline{p}_n \in U(f_n)$ converge to the diffusion parameters of parameter vectors $\underline{p}_0 \in U(f_0)$, i.e.,

$$\lim_{n\to\infty} \mu_n = \mu_0 \quad \text{and} \quad \lim_{n\to\infty} \sigma_n = \sigma_0 \, . \tag{6}$$

In general we cannot ensure the convergence of the jump parameters λ, μ_Y and σ_Y, but we obtain the limit condition

$$\lim_{n\to\infty} \lambda_n({\mu_Y}_n^2 + {\sigma_Y}_n^2) = \lambda_0({\mu_Y}_0^2 + {\sigma_Y}_0^2) \, . \tag{7}$$

Note that the limit case $\lambda \to \infty$ may lead to an asymptotical non-injectivity of the operator $\mathcal{A}$. Then also the equations (6) and (7) are not necessarily fulfilled asymptotically. In order to prevent this case, we restrict this parameter by an upper bound $\lambda_{\max}$. Besides restricting the jump intensity we also restrict the jump heights and consider for sufficiently large constants $\lambda_{\max}$, ${\mu_Y}_{\max}$ and ${\sigma_Y}_{\max}$ parameter vectors in the restricted domain

$$D_{\max} := \left\{\underline{p} \in D : \ \lambda \leq \lambda_{\max} < \infty \, , |\mu_Y| \leq {\mu_Y}_{\max} < \infty \, , \sigma_Y \leq {\sigma_Y}_{\max} < \infty\right\} .$$

The limitation of μ_Y and σ_Y is not essential, because the jumps and absolute returns $|r_\tau|$ increase arbitrarily as $|\mu_Y| \to \infty$ or $\sigma_Y \to \infty$. In [3] and [4] we have proven the following proposition.

Proposition 6 *Let $\{f_n\} \subset \mathcal{A}(D_{max})$ be a sequence which converges in Y to the density function $f_0 \in \mathcal{A}(D_{max})$. Then the inverse images $\underline{p}_n = (\mu_n, \sigma_n, \lambda_n, \mu_{Yn}, \sigma_{Yn})^T \in U(f_n) \cap D_{max}$ and $\underline{p}_0 = (\mu_0, \sigma_0, \lambda_0, \mu_{Y0}, \sigma_{Y0})^T \in U(f_0) \cap D_{max}$ fulfill (6) and (7). Every infinite subsequence $\left\{\underline{p}_{n_k}\right\} \subset U(f_n) \cap D_{max}$ has an accumulation point $\underline{\hat{p}} \in U(f_0) \cap D_{max}$. Moreover, if additionally $f_0 \in \mathcal{A}(D_{max} \cap \hat{D})$, then for a sufficiently large n the sets $U(f_n) \cap D_{max}$ and $U(f_0)$ are both a singleton and the sequence $\left\{\underline{p}_n\right\}$ converges to $\underline{p}_0$.*

If we accept solutions only in $D_{\max}$, the operator equation (4) is stably solvable in terms of properties formulated in proposition 6. Therefore we consider the least-squares problem

$$\Psi(\underline{p}) := \left\|\mathcal{A}(\underline{p}) - f^\delta\right\|^2_{L^2(\mathbb{R})} \longrightarrow \min, \quad \text{subject to} \quad \underline{p} \in D_{\max}. \tag{8}$$

Since $\sigma = 0$ is excluded, the set $D_{\max}$ of feasible solutions in (8) is not closed. Nevertheless this extremal problem is solvable in all practical cases where the empirical density function f^δ is uniformly bounded by a constant $B > 0$. Namely, by restricting the intensity parameter λ we particularly get the existence of a positive lower bound

$$\sigma \geq \frac{e^{-\lambda_{\max}\tau}}{\left(\delta + \sqrt{B}\right)^2 \sqrt{2\pi\tau}} > 0$$

for all parameter vectors $\underline{p}$ from the set $M_\eta = \left\{\underline{p} \in D_{\max} : \left\|\mathcal{A}(\underline{p}) - f^\delta\right\|_{L^2(\mathbb{R})} \leq \eta\right\}$ and sufficiently small $\eta \leq \delta$.

Unfortunately, there occur some instability effects by solving the least squares problem (8) numerically even if the noise level δ is very small (see also [5]). These ill-posedness phenomena are caused by an ill-conditioned extremal problem

$$\left\|\mathrm{A}(\underline{p}) - \underline{z}^\delta\right\|_2^2 \to \min, \quad \text{subject to} \quad \underline{p} \in D_{\max} \tag{9}$$

with Euclidean norm $\|\cdot\|_2$ and minimizer $\underline{p}^\delta$, which we obtain after discretization. In this context, we search for least-squares solutions of the discretized version

$$\mathrm{A}(\underline{p}) = \underline{z} \tag{10}$$

of equation (4). Since the empirical density function has the form of a histogram, we discretize the density function $f(\,\cdot\,,\underline{p})$ in the same way and denote the discretized function as $\tilde{f}(\,\cdot\,,\underline{p})$. Hence we consider for the grid $x_0 < x_1 < \ldots < x_n$ the discretized operator $\mathrm{A} : D \subset \mathbb{R}^5 \to \mathbb{R}^n$ mapping $\underline{p} \in D$ to the vector $\underline{z} = (\tilde{f}(x_0,\underline{p}), \tilde{f}(x_1,\underline{p}), \ldots, \tilde{f}(x_{n-1},\underline{p}))^T \in \mathbb{R}^n$. For studying the stability of least-squares solutions we perturb the vector $\underline{z}^* = \mathrm{A}(\underline{p}^*)$ component-wise by normally distributed and uncorrelated errors $\epsilon_i \sim N(0,\delta^2)$ by setting $z_i^\delta = z_i^*(1+\epsilon_i)$ such that we have

$$\mathrm{E}\|\underline{z}^* - \underline{z}^\delta\|_2^2 = \sum_{i=0}^{n-1} (z_i^*)^2\, \mathrm{E}\left(\epsilon_i^2\right) = \delta^2\, \|\underline{z}^*\|_2^2$$

and

$$\mathrm{E}\left(\frac{\|\underline{z}^* - \underline{z}^\delta\|_2^2}{\|\underline{z}^*\|_2^2}\right) = \delta^2\,.$$

Figure 1 illustrates the used functions and data. The data vector $\underline{z}^\delta$ can be considered as a skeleton of the empirical density function obtained from noisy returns $\underline{r}_\tau^\delta$. Since we do not compute the data $\underline{z}^\delta$ from the returns, we must calculate the empirical moments $m_{\tau,k}^\delta$ through

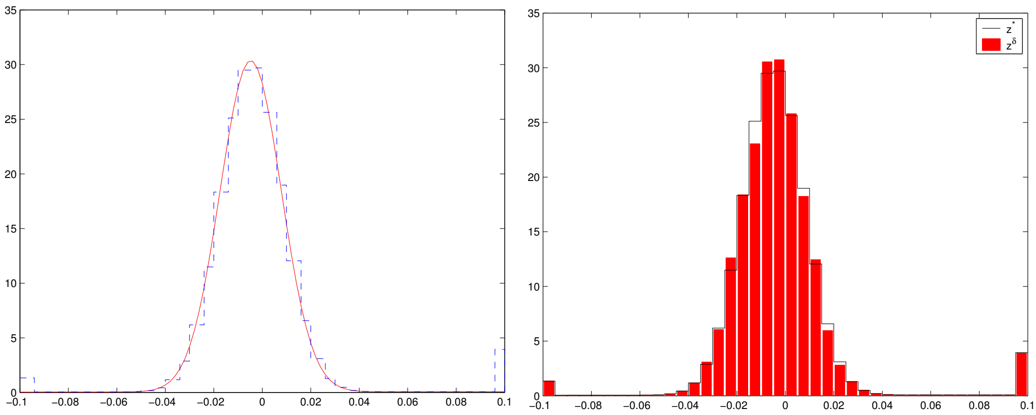
$$m_{\tau,k}^\delta \approx \sum_{i=0}^{n-1} \frac{1}{k+1}\left(x_{i+1}^{k+1} - x_i^{k+1}\right) z_i^\delta\,,$$

where we use the approximation

$$m_{\tau,k}(\underline{p}) = \int_{\mathbb{R}} x^k f(x,\underline{p})\,dx \approx \int_{\mathbb{R}} x^k \tilde{f}(x,\underline{p})\,dx = \sum_{i=0}^{n-1} z_i \frac{1}{k+1}(x_{i+1}^{k+1} - x_i^{k+1})\,.$$

Using proposition 2 we thus obtain also empirical semi-invariants $s_{\tau,k}^\delta$.

The following numerical examples will show, that non-injectivity of the forward operator $\mathcal{A}$ is not the only ill-posedness phenomenon occurring in the inverse problem of determining $\underline{p}$. In order to illustrate the instability as the second ingredient of ill-posedness, here we generated for the parameter vector $\underline{p}^* = (0.1, 0.2, 10.0, 0.1, 0.2)^T$ the exact right hand side $\underline{z}^*$ concerning daily returns ($\tau = \frac{1}{250}$). For studying the case of noisy data we actually computed the weakly perturbed vector $\underline{p}^\delta$ with $\delta = 0.01$. A simulated price trajectory



$f(\,\cdot\,,\underline{p}^*)$ and $\tilde{f}(\,\cdot\,,\underline{p}^*)$ $\qquad$ $\underline{z}^*$ and $\underline{z}^\delta$

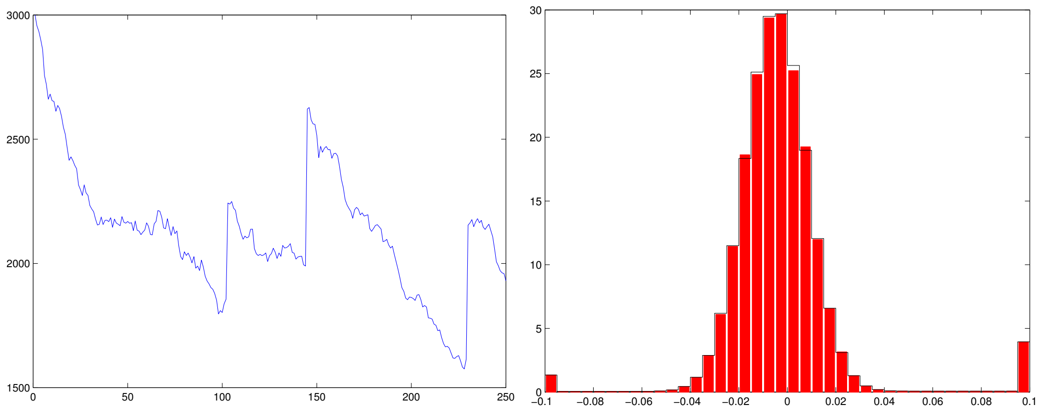
Figure 1: Discretization and perturbed data

associated with those settings is presented in the left-hand picture of figure 2, whereas the right-hand picture shows the histograms corresponding to the exact data $\underline{z}^*$ and the noisy data $\underline{z}^\delta$. The semi-invariants $s^\delta_{\tau,k}$ of the noisy data are displayed in table 2. The deviation between the exact semi-invariants and $s^\delta_{\tau,k}$ has the same order of magnitude like the noise level δ.

k	$s_{\tau,k}(\underline{p}^*)$	$s^\delta_{\tau,k}$	deviation
1	-0.000779874	-0.000766772	1.68%
2	$+0.002160000$	$+0.002162937$	0.14%
3	$+0.000520000$	$+0.000519441$	0.11%
4	$+0.000292000$	$+0.000291870$	0.04%
5	$+0.000112400$	$+0.000112369$	0.03%
6	$+0.000069640$	$+0.000069666$	0.04%
7	$+0.000033940$	$+0.000033959$	0.06%
8	$+0.000022893$	$+0.000022913$	0.09%

Table 1: Semi-invariants of the perturbed data ($\tau = \frac{1}{250}$)

We begin our study with the case of unperturbed data $\underline{z}^*$. Solving the least-squares problem (9) iteratively with an appropriate initial guess leads to the very good approximation $\underline{p} = (0.1022, 0.2000, 9.9920, 0.1002, 0.2003)^T$ of the exact solution $\underline{p}^\star = (0.1, 0.2, 10.0, 0.1, 0.2)^T$. Since the estimated parameters are close to the exact ones, the density functions $\tilde{f}(\,\cdot\,,\underline{p})$ and $\tilde{f}(\,\cdot\,,\underline{p}^*)$



Simulated prices

Histograms to $\underline{z}^*$ and $\underline{z}^\delta$

Figure 2: Sample trajectory and data

nearly coincide and the deviations between semi-invariants are also rather small (see table 2). For the sake of simplicity we convert daily semi-invariants using the scaling property $s_{c\tau,k}(\underline{p}) = cs_{\tau,k}(\underline{p})$ with $c = \frac{1}{\tau} = 250$ into annualized semi-invariants with $\tau = 1$ and write in this case for the $k-$th semi-invariant s_k instead of $s_{1,k}$.

k	$s_k(\underline{p})$	$s_k(\underline{p}^*)$	deviation
1	-0.1935712	-0.1949685	0.72%
2	$+0.5414362$	$+0.5400000$	0.27%
3	$+0.1306442$	$+0.1300000$	0.50%
4	$+0.0734753$	$+0.0730000$	0.65%
5	$+0.0283397$	$+0.0281000$	0.85%
6	$+0.0175866$	$+0.0174100$	1.01%
7	$+0.0085878$	$+0.0084850$	1.21%
8	$+0.0058021$	$+0.0057233$	1.38%

Table 2: Annualized semi-invariants of the estimated parameter vector $\underline{p}$

The computation of a least-squares solution for a first realization of the weakly perturbed data vector $\underline{z}^\delta$, however, provides us with rather large deviations between exact and estimated parameters, although the noise level $\delta = 0.01$ is not too high (see table 3). In particular, the parameter μ responds very sensitively to data changes. However, the density function of the estimated parameters fits the data very well. A very interesting effect is that the semi-invariants and moments of the estimated parameters clearly deviate from the exact ones (see table 4). This observation will be used in section 3 in order to construct a specific approach of stabilization.

parameters	$\underline{p}^\delta$ (estimated)	$\underline{p}^*$	deviation
μ	-0.0424	$+0.1$	142.35%
σ	$+0.2003$	$+0.2$	0.16%
λ	$+10.6748$	$+10.0$	6.75%
μ_Y	$+0.0865$	$+0.1$	13.54%
σ_Y	$+0.1784$	$+0.2$	10.82%

Table 3: Estimated parameters for a first realization of $\underline{z}^\delta$ with $\delta = 0.01$

k	$s_k(\underline{p}^\delta)$	$s_k(\underline{p}^*)$	deviation
1	-0.2900873	-0.1949685	48.79%
2	$+0.4594911$	$+0.5400000$	14.91%
3	$+0.0949759$	$+0.1300000$	26.94%
4	$+0.0482313$	$+0.0730000$	33.93%
5	$+0.0162548$	$+0.0281000$	42.15%
6	$+0.0090766$	$+0.0174100$	47.87%
7	$+0.0038871$	$+0.0084850$	54.19%
8	$+0.0023572$	$+0.0057233$	58.81%

Table 4: Annualized semi-invariants of the estimated parameter vector $\underline{p}^\delta$

We repeated the computations with six additional different realizations of the data vector $\underline{z}^\delta$. The results are given in table 5.
They show the instability of the least-squares problem (8) expressed by a wide range of possible parameter vectors $\underline{p}^\delta$ obtained by varying data perturbations with only one per cent noise. This instability requires the use of a regularization method (see, e.g., [6] and [8]) for finding approximate solutions of $\underline{p}^*$ in a stable manner.

3 Multi-parameter regularization approach

The instability of least-squares solutions $\underline{p}^\delta$ with respect to varying data $\underline{z}^\delta$ as observed in section 2 sufficiently motivates the use of a regularization

	μ	σ	λ	μ_Y	σ_Y
$\underline{p}^*$	0.1	0.2	10.0	0.1	0.2
1st run	0.6413	0.1993	8.6430	0.1522	0.2792
2nd run	−0.1225	0.1992	11.4048	0.0759	0.1617
3rd run	−0.0913	0.2003	11.1035	0.0803	0.1685
4th run	0.0354	0.2003	10.3857	0.0924	0.1876
5th run	0.2557	0.1990	9.5495	0.1128	0.2198
6th run	0.5253	0.1995	8.8196	0.1415	0.2637

Table 5: Least-squares solutions $\underline{p}^\delta$ of six further noisy data simulations with $\delta = 0.01$

approach for determining the parameter vector $\underline{p}$. Since there is no other a priori information that prefers a specific parameter vector $\underline{p}$, we exploit the fact pointed out in the case study that the semi-invariants and moments of the estimated parameters sensitively respond to parameter changes. Hence, we use the first l semi-invariants for a regularization, because they are scalable in time, i.e. $s_{c\tau,k}(\underline{p}) = cs_{\tau,k}(\underline{p})$ $(c > 0)$, whereas we have for the moments $m_{c\tau,k}(\underline{p}) \neq cm_{\tau,k}(\underline{p})$ for $k = 1, 2, ..., l$. Consequently, we can use without loss of generality annualized semi-invariants in the following. The order of magnitude varies for different semi-invariants. In this context, we assume that the upper bounds δ_k $(k = 1, ..., l)$ of admissible semi-invariant deviations can be prescribed in a useful manner. For example it is well known from statistics that usually $\left|s_1(\underline{p}^\star) - s_1^\delta\right|$ is larger than deviations of higher semi-invariants. This motivates the application of a multi-parameter regularization approach as introduced and analyzed in [8, §4.2] (see also [9]).

So we are going to compute the optimal vector $\underline{p}^\delta_{opt}$ as a minimizer of the problem

$$\| \mathrm{A}(\underline{p}) - \underline{z}^\delta\|_2^2 \to \min\,, \quad \text{subject to} \quad \underline{p} \in M^\delta \tag{11}$$

with

$$M^\delta = \left\{\underline{p} \in D_{\max} :\ \left|s_k(\underline{p}) - s_k^\delta\right| \le \delta_k, \quad k = 1, \ldots, l\right\}\,,$$

where s_k^δ is the k-th empirical semi-invariant computed from data $\underline{z}^\delta$. The optimal vector $\underline{p}^\delta_{opt}$, however, represents an appropriate trade-off between acceptably small discrepancy values $\| \mathrm{A}(\underline{p}^\delta_{opt}) - \underline{z}^\delta\|_2$ and a required fitting $\underline{p}^\delta_{opt} \in M^\delta$ of semi-invariants. In order to compute $\underline{p}^\delta_{opt}$ in an efficient manner, we search for a saddle point of the Lagrangian functional

$$L(\underline{p},\underline{\alpha}) = \|\,\mathrm{A}(\underline{p}) - \underline{z}^\delta\|^2 + \sum_{k=1}^{l} \alpha_k \left(\left|s_k(\underline{p}) - s_k^\delta\right|^2 - \delta_k^2 \right) \tag{12}$$

of the in general non-convex optimization problem (11) with multiplier vectors $\underline{\alpha} = (\alpha_1, ..., \alpha_l)^T \in \mathbb{R}_+^l$. The numerical computation of such saddle-point can be performed iteratively by computing minimizers $\underline{p}_{\underline{\alpha}}^\delta$ of

$$F(\underline{p},\underline{\alpha}) = \|\,\mathrm{A}(\underline{p}) - \underline{z}^\delta\|_2^2 + \Omega(\underline{p},\underline{\alpha},\underline{z}^\delta) \;\rightarrow\; \min\,, \quad \text{subject to} \quad \underline{p} \in D_{\max} \tag{13}$$

with penalty functional

$$\Omega(\underline{p},\underline{\alpha},\underline{z}^\delta) = \sum_{k=1}^{l} \alpha_k \left|s_k(\underline{p}) - s_k^\delta\right|^2 \tag{14}$$

and a sequence of vectors $\underline{\alpha} = (\alpha_1, ..., \alpha_l)^T$. This a specific multi-parameter approach for finding regularized least-squares solutions to equation (10). The approach is based theoretically on the following lemma 7 and theorem 8.

Lemma 7 *A pair $(\hat{\underline{p}},\hat{\underline{\alpha}}) \in D_{max} \times \mathbb{R}_+^l$ is a saddle-point of the Lagrangian functional (12) to the optimization problem (11) if $\hat{\underline{p}} = \underline{p}_{\hat{\underline{\alpha}}}^\delta$ is a minimizer of (13) for the regularization parameter vector $\underline{\alpha} = \hat{\underline{\alpha}}$ and satisfies simultaneously the l equations*

$$\hat{\alpha}_k \left(|s_k(\hat{\underline{p}}) - s_k^\delta|^2 - \delta_k^2 \right) = 0 \quad (k = 1, \ldots, l) \tag{15}$$

and the additional conditions

$$|s_k(\hat{\underline{p}}) - s_k^\delta|^2 \le \delta_k^2 \quad \text{if} \quad \hat{\alpha}_k = 0 \quad (k = 1, \ldots, l)\,. \tag{16}$$

Proof

The pair $(\hat{\underline{p}},\hat{\underline{\alpha}})$ is a saddle-point of the Lagrangian functional (12) to the optimization problem (11) if it fulfills for all $\underline{p} \in D_{\max}$ and all $\underline{\alpha} \in \mathbb{R}_+^l$ the inequalities

$$L(\hat{\underline{p}},\underline{\alpha}) \le L(\hat{\underline{p}},\hat{\underline{\alpha}}) \le L(\underline{p},\hat{\underline{\alpha}})\,.$$

The right inequality is evidently satisfied if $\underline{\hat{p}} = \underline{p}^{\delta}_{\underline{\hat{\alpha}}}$ is a minimizer of (13) for the regularization parameter vector $\underline{\hat{\alpha}}$. The left inequality, however, is equivalent to

$$\sum_{k=1}^{l} (\hat{\alpha}_k - \alpha_k) \left(\left| s_k(\underline{\hat{p}}) - s_k^{\delta} \right|^2 - \delta_k^2 \right) \geq 0 \quad \text{for all } \underline{\alpha} \in \mathbb{R}^l_+ \,. \tag{17}$$

Certainly, the pair of conditions (15) and (16) implies the inequality (17). This proves the lemma.

□

It is well-known from optimization theory that $\underline{\hat{p}}$ solves the optimization problem (11) whenever $(\underline{\hat{p}}, \underline{\hat{\alpha}}) \in D_{\max} \times \mathbb{R}^l_+$ in a saddle-point in the sense of lemma 7. Adapted from this result we construct an iteration process that approaches this solution. In this context, we reformulate the equations (15) as fixed point equations in the form

$$\hat{\alpha}_k \; = \; \hat{\alpha}_k \, \frac{|s_k(\underline{\hat{p}}) - s_k^{\delta}|^2}{\delta_k^2} \qquad (k = 1, \ldots, l). \tag{18}$$

Moreover, we choose a small positive number $0 < \varepsilon \ll 1$ and an initial guess $\underline{\alpha}^{(0)} \in \mathbb{R}^l_+$ with positive components $\alpha_k^{(0)}$ $(k = 1, \ldots, l)$ for starting the iteration process

$$\underline{p}^{(j)} := \underline{p}^{\delta}_{\underline{\alpha}^{(j)}} \qquad (j = 0, 1, \ldots) \; ;$$

$$\alpha_k^{(j+1)} := \alpha_k^{(j)} \max \left\{ \frac{\left| s_k(\underline{p}^{(j)}) - s_k^{\delta} \right|^2}{\delta_k^2}, \, \varepsilon \right\} \qquad (j = 0, 1, \ldots \; ; \; k = 1, \ldots, l). \tag{19}$$

Theorem 8 *If the iteration (19) converges, i.e.,* $\underline{\alpha}^{(j)} \to \underline{\hat{\alpha}} \in \mathbb{R}^l_+$ *and* $\underline{p}^{(j)} \to \underline{\hat{p}} \in D_{max}$ *as* $j \to \infty$*, then the pair* $(\underline{\hat{p}}, \underline{\hat{\alpha}})$ *is a saddle-point of the Lagrangian functional (12) and hence* $\underline{\hat{p}} = \underline{p}^{\delta}_{opt}$ *is an optimal solution of (11).*

Proof

We apply lemma 7 which proves the theorem if all the hypotheses of this lemma are satisfied. By construction the limit vectors $\underline{\hat{\alpha}}$ and $\underline{\hat{p}}$ of iteration (19) satisfy for all $\underline{p} \in D_{\max}$ the inequality $F(\underline{\hat{p}}, \underline{\hat{\alpha}}) \leq F(\underline{p}, \underline{\hat{\alpha}})$ such that $\underline{\hat{p}}$ is a solution of (13) for $\underline{\alpha} = \underline{\hat{\alpha}}$. Moreover, the pair $(\underline{\hat{\alpha}}, \underline{\hat{p}})$ fulfills for $k = 1, \ldots, l$ the equations

$$\hat{\alpha}_k = \hat{\alpha}_k \max \left\{ \frac{\left| s_k(\underline{\hat{p}}) - s_k^\delta \right|^2}{\delta_k^2}, \varepsilon \right\} . \tag{20}$$

For fixed k equation (20) holds if $\hat{\alpha}_k = 0$ or $\left| s_k(\underline{\hat{p}}) - s_k^\delta \right|^2 = \delta_k^2$. This implies relation (15). In order to show relation (16), we assume that $\left| s_k(\underline{\hat{p}}) - s_k^\delta \right|^2 > \delta_k^2$ and $\hat{\alpha}_k = 0$ would hold simultaneously for some k. Then there is a $j_0 = j_0(k)$ such that $\left| s_k(\underline{p}^{(j)}) - s_k^\delta \right|^2 > \delta_k^2$ is valid for all $j \geq j_0$, since the $k-$th semi-invariant $s_k(\underline{p})$ is continuous with respect to parameter vectors $\underline{p} \in D$. Thus we have

$$\max \left\{ \frac{\left| s_k(\underline{p}^{(j)}) - s_k^\delta \right|^2}{\delta_k^2}, \, \varepsilon \right\} > 1$$

and by (19) the inequality $\alpha_k^{(j+1)} > \alpha_k^{(j)} > 0$ for all $j \geq j_0$. Then the limit $\hat{\alpha}_k = \lim_{j\to\infty} \alpha_k^{(j)}$ cannot be zero as assumed and consequently condition (16) is valid. This proves the theorem. $\square$

Unfortunately, theorem 8 makes only an assertion provided that the iteration process (19) converges. Indeed, it seems to be very difficult to formulate sufficient conditions for getting a contraction mapping in the sense of Banach's fixed point theorem. However, the iteration along the lines of the following algorithm seems to converge rather stable in a wide field of applications (see also [10]).

Algorithm 9

Step 0 *Choose* $j_{max} \in \mathbb{N}$, $\varepsilon, \varepsilon_1 > 0$ *sufficiently small and* $\underline{\alpha}^{(0)}$ *with positive components. Set* $j := 0$.

Step 1 *Compute* $\underline{p}^{(j)} = \underline{p}^\delta_{\underline{\alpha}^{(j)}}$ *as a minimizer of* $F(\underline{p}, \underline{\alpha}^{(j)})$ *by solving (13).*

Step 2 *Compute* $\alpha_k^{(j+1)} = \alpha_k^{(j)} \max \left\{ \frac{\left| s_k(\underline{p}^{(j)}) - s_k^\delta \right|^2}{\delta_k^2}, \, \varepsilon \right\}$ $(k = 1, 2, ..., l)$.

Step 3 *If* $\|\underline{\alpha}^{(j+1)} - \underline{\alpha}^{(j)}\|_2 \leq \varepsilon_1$ *or* $j + 1 \geq j_{max}$, *set* $\underline{\hat{p}} := \underline{p}^{(j)}$ *and stop; otherwise set* $j := j + 1$ *and go to step 1.*

4 Numerical case studies

The multi-parameter regularization approach formulated in algorithm 9 was tested by a case study comparable to that of section 2 with a parameter vector $\underline{p}^* = (0.1, 0.2, 10.0, 0.1, 0.2)^T$, $\tau = \frac{1}{250}$ and $\varepsilon = \varepsilon_1 = 1e{-}05$. However, we have chosen the higher noise level $\delta = 0.1$, because this choice seems to corresponds to a more realistic situation in financial practice. The a priori chosen error bounds δ_k $(k = 1, 2, ..., 5)$ are listed in the fifth column in table 8. The empirical semi-invariants s_k^δ obtained from the data, which provide a basis for the penalty term (14), and the corresponding exact semi-invariants of the parameter vector $\underline{p}^*$ are compared in table 6. For the used data the iteration

k	s_k^δ (empirical semi-inv.)	$s_k(\underline{p}^*)$ (exact semi-inv.)	$\lvert s_k(\underline{p}^*) - s_k^\delta\rvert$	deviation
1	−0.17002795	−0.19496852	0.02494	12.79%
2	+0.54421567	+0.54000000	0.00422	0.78%
3	+0.13101847	+0.13000000	0.00102	0.78%
4	+0.07375075	+0.07300000	0.00075	1.00%
5	+0.02837745	+0.02810000	0.00028	0.99%

Table 6: Comparison of empirical and exact semi-invariants

process was convergent and provided the regularized solution $\underline{\hat{p}} = \underline{p}^\delta_{\underline{\hat{\alpha}}}$, which is compared with $\underline{p}^*$ in table 7. The deviation between the regularized solution and the true parameter vector (in per cent) is given for the five components in the fourth column of table 7. As expected the determination of μ is extremely difficult (about 16 per cent error), whereas the standard deviations σ and σ_Y are estimated rather good with errors less than 2 per cent. The computed

parameter	$\underline{p}^*$	$\underline{\hat{p}} = \underline{p}^\delta_{\underline{\hat{\alpha}}}$	deviation
μ	0.1	0.115857	15.86%
σ	0.2	0.197669	1.17%
λ	10.0	9.491907	5.08%
μ_Y	0.1	0.103871	3.87%
σ_Y	0.2	0.203410	1.71%

Table 7: Some optimal multi-parameter regularized solution and its performance

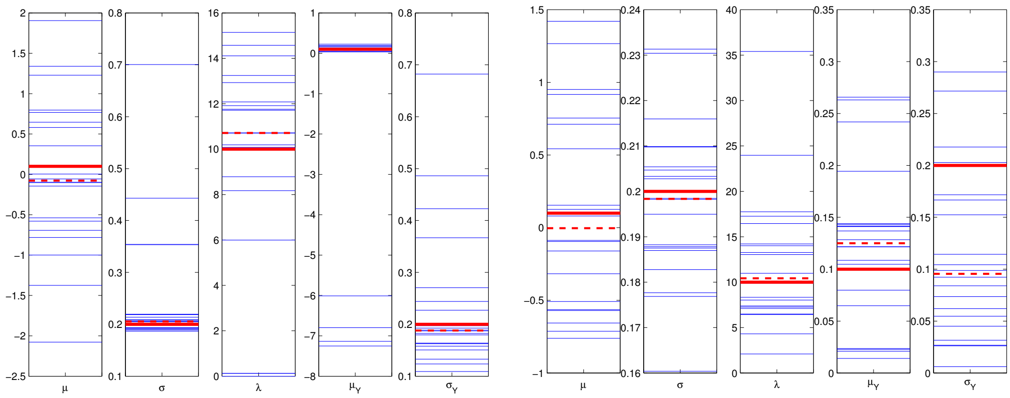
optimal regularization parameter vector $\underline{\hat{\alpha}}$ is given for that data realization in the last column of table 8. Furthermore, table 8 indicates that just for the second and third semi-invariants the restrictions given by δ_2 and δ_3 are active for the optimal solution and hence only α_2 and α_3 are positive.

k	s_k^δ	$s_k(\underline{\hat{p}})$	$\lvert s_k(\underline{\hat{p}}) - s_k^\delta\rvert$	δ_k	$\underline{\hat{\alpha}}$
1	−0.17002795	−0.17002795	0.00680543	0.01	0.00000000
2	+0.54421567	+0.53421568	0.00999999	0.01	0.28657429
3	+0.13101847	+0.13301848	0.00200001	0.002	6.34543394
4	+0.07375075	+0.07527707	0.00152632	0.002	0.00000000
5	+0.02837745	+0.02983393	0.00145648	0.002	0.00000000

Table 8: Prescribed bounds and more details of the iteration result

As a conclusion one can say that the multi-parameter approach that combines the least-squares fitting of the empirical density function obtained from return data with the fitting of empirical semi-invariants works quite well for the case study situation. This kind of regularization is able to surmount part of instability of conventional least-squares fittings and leads to fairly acceptable approximate parameters for the jump diffusion process.

At the end of this section we illustrate the stabilizing effect of the multi-parameter regularization by comparing this method with a conventional maximum likelihood estimation for a practical situation. In this context we simulated 20 times series of 250 prices, which correspond to the case that the parameter vector has to be estimated from daily closing prices of one trading year. The results of the maximum likelihood estimation are displayed in the left hand side and the corresponding results of the multi-parameter regularization on the right hand side of figure 3. In order to analyze the results correctly, it is important to notice the different scales used in both graphics. Each bar contains the results for the corresponding parameter. The thin lines represent the estimated values of a parameter, the dashed line the median of all estimated values and the bold face line the exact parameter value. We should mention that 250 data are not enough to expect very good estimates, since they correspond to a large noise level $\delta \gg 0.1$. There are some significant outliers for σ, λ, μ_Y and σ_Y using the maximum likelihood estimation. They are caused by a misinterpretation of the jump part. If the jumps are not identified, but misinterpreted as a high volatility, then the estimated value of the parameter σ is to large and either jump intensity λ or jump mean μ_Y and jump volatility σ_Y to small. The left-hand schema of figure 3 shows



Maximum likelihood estimation Multi-parameter regularization

Figure 3: Results of 20 simulations all based on 250 daily closing prices

that instability phenomena also occur in the case of maximum likelihood estimation. On the other hand, the stabilizing effect of the multi-parameter regularization approach is illustrated by the right-hand schema of figure 3 in form of reduced variations for the majority of estimated parameters over 20 simulations.

5 A modification based on exponent variation

In section 4 we have illustrated that the multi-parameter regularization introduced in section 3 is suitable to obtain stable approximate solutions for the parameter estimation problem under consideration. The multi-parameter algorithm, however, can be improved by a modification based on exponent variation. The proposed algorithm 9 is based on condition (15) and the associated fix point equations (18). Instead of (15) we can also use the equivalent conditions

$$\underline{\hat{\alpha}}_k \left(\left| s_k(\underline{\hat{p}}) - s_k^\delta \right|^\gamma - \delta_k^\gamma \right) = 0 \quad (k = 1, \ldots, l)$$

for arbitrary exponents $\gamma > 0$ and modify the iteration (19) as

$$\underline{p}^{(j)} = \underline{p}^{\delta}_{\underline{\alpha}^{(j)}} \quad \text{for} \quad (j = 0, 1, \ldots)\,;$$

$$\alpha_k^{(j+1)} = \alpha_k^{(j)} \max\left\{ \frac{\left|s_k(\underline{p}^{(j)}) - s_k^{\delta}\right|^{\gamma}}{\delta_k^{\gamma}}, \varepsilon \right\} \quad (j = 0, 1, \ldots\ ;\ k = 1, \ldots, l)\,. \tag{21}$$

The assertion of theorem 8 remains valid for every $\gamma > 0$ and the appropriate choice of γ can be exploited to accelerate the rate of convergence and hence to improve the efficiency of the algorithm.

Finally, we present the results of two further numerical case studies concerning the effect of appropriately chosen values γ for the iteration process (21). For both studies we used a noise level $\delta = 0.05$, a maximum number of iterations $j_{\max} = 3000$ and prescribed bounds δ_k $(k = 1, ..., 5)$ as given below. All other parameters and settings were taken as in the case study of section 4.

In a first study we used the bounds $\delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = 0.005$ and we compared the required number of iterations depending on the exponent γ which varied in the range $[0.1, 4.5]$. We simulated 5 data vectors $\underline{z}^{\delta}$ and estimated for each γ the parameter vector $\underline{p}^{\delta}$. In most cases the iteration converged, particularly for $\gamma < 2$. It could be seen that the exponent γ did not affect the optimization result provided that the iteration converged. Namely, the essential decimal places of the regularized solutions $\underline{\hat{p}} = \underline{p}^{\delta}_{\underline{\hat{\alpha}}}$ coincide for all γ in case of convergence. However, the exponent γ strongly influences the number of required iterations (see table 9).

Table 9 suggest to choose $\gamma < 2$. This suggestion could be confirmed by a second study, the results of which are displayed in figure 4. This figure shows mean and median of the required iterations from 10 runs, where we took the bounds $\delta_1 = 0.08, \delta_2 = 0.02, \delta_3 = \delta_4 = \delta_5 = 0.004$ and give results also for smaller exponents $\gamma < 0.1$. For very small γ the required number of iterations also grows such that an intermediate interval, for example $\gamma \in [0.2, 1.5]$, can be recommended for use in this iteration process. Comparable suggestions for the choice of the iteration exponent were given in [10], where an inverse eigenvalue problem was solved by a similar multi-parameter regularization approach.

γ	run 1	run 2	run 3	run 4	run 5
0.1	281	325	1115	167	751
0.2	135	241	655	94	484
0.3	106	366	854	69	264
0.4	74	302	719	49	185
0.5	89	309	555	48	189
0.6	63	272	576	54	153
0.7	76	272	490	46	105
0.8	45	488	1013	42	142
0.9	49	292	1468	61	150
1.0	50	179	2533	69	91
1.1	45	551	559	36	137
1.2	49	648	786	52	137
1.3	62	947	1126	34	157
1.4	29	519	3000	66	111
1.5	46	439	808	62	181
1.6	48	564	1506	77	127
1.7	161	862	3000	50	165
1.8	95	433	3000	159	159
1.9	93	997	3000	59	438
2.0	83	1233	3000	72	3000
2.5	112	3000	3000	3000	354
3.5	847	3000	3000	3000	3000
4.5	3000	3000	3000	3000	3000

Table 9: Required iterations depending on γ, first study

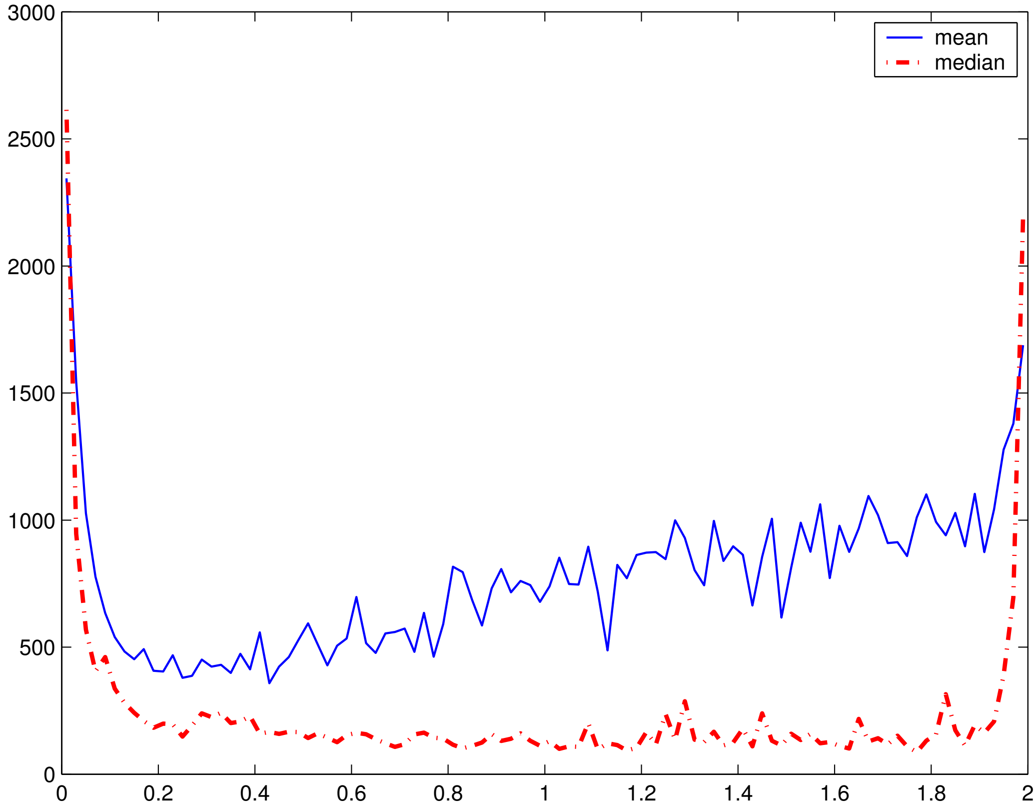


Figure 4: Required iterations depending on γ, second study

References

[1] D. D. Ang, R. Gorenflo, V. K. Le, and D. D. Trong. *Moment Theory and Some Inverse Problems in Potential Theory and Heat Conduction.* Springer, Berlin, 2002.

[2] R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman & Hall/CRC, London, 2004.

[3] D. Düvelmeyer. *Inkorrektheitsphänomene und Regularisierung bei der Parameterschätzung für Jump-Diffusions-Prozesse. PhD thesis.* Chemnitz University of Technology, Faculty of Mathematics, Chemnitz, 2005.

[4] D. Düvelmeyer. Some stability results of parameter identification in a jump diffusion model. In J. vom Scheidt, editor, *Tagungsband zum Workshop Stochastische Analysis Klingenthal 2004 (ISSN 1612-5665)*, pages 27–43, Chemnitz University of Technology, Faculty of Mathematics, 2005.

[5] D. Düvelmeyer and B. Hofmann. Ill-posedness of parameter estimation in jump diffusion processes. In J. vom Scheidt, editor, *Tagungsband zum Workshop Stochastische Analysis Bärenstein 2003 (ISSN 1612-5665)*, pages 5–20, Chemnitz University of Technology, Faculty of Mathematics, 2004.

[6] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer, Dordrecht, 1996.

[7] S. N. Evans and P. B. Stark. Inverse problems as statistics. *Inverse Problems*, 18:R55–97, 2002.

[8] B. Hofmann. *Regularization for Applied Inverse and Ill-Posed Problems.* Teubner, Leipzig, 1986.

[9] B. Hofmann. A control algorithm for a multiparameter regularization process. *U.S.S.R. Comput. Math. Math. Phys*, 28:88–90, 1988. translated from Russian.

[10] B. Hofmann. An improved multi-parameter adjustment algorithm for inverse eigenvalue problems. *Zeitschrift für Analysis und ihre Anwendungen*, 7:511–518, 1988.

[11] B. Hofmann. *Mathematik inverser Probleme.* B. G. Teubner, Stuttgart, 1999.

[12] P. Honoré. *Pitfalls in Estimating Jump Diffusion Models. Working Paper (www.hha.dk/∼phs).* University of Aarhus Graduate School of Business, Aarhus, Denmark, 1998.

[13] R. C. Merton. *Continuous-Time Finance.* Blackwell Publishers Inc., Cambridge, 1992.

[14] P. Protter. *Stochastic Integration and Differential Equations.* Springer Verlag, Berlin, 1992.

[15] A. N. Shiryaev. *Probability.* Springer, New York, 1995.

[16] A. N. Shiryaev. *Essentials of Stochastic Finance - Facts, Models, Theory.* World Scientific Publishing, Singapore, 1999.

[17] H.-J. Starkloff, D. Düvelmeyer, and B. Hofmann. A note on uniqueness of parameter identification in a jump diffusion model. In J. vom Scheidt, editor, *Tagungsband zum Workshop Stochastische Analysis Klingenthal 2004 (ISSN 1612-5665)*, pages 251–256, Chemnitz University of Technology, Faculty of Mathematics, 2005.