Parameter estimation in a generalized bivariate Ornstein-Uhlenbeck model

R. Krämer, M. Richter, B. Hofmann

Technische Universität Chemnitz, Fakultät für Mathematik, 09107 Chemnitz, Germany

Abstract

In this paper, we consider the inverse problem of calibrating a generalization of the bivariate Ornstein-Uhlenbeck model introduced by Lo and Wang. Even though the generalized Black-Scholes option pricing formula still holds, option prices change in comparison to the classical Black-Scholes model. The time-dependent volatility function and the other (real-valued) parameters in the model are calibrated simultaneously from option price data and from some empirical moments of the logarithmic returns. This gives an ill-posed inverse problem, which requires a regularization approach. Applying the theory of Engl, Hanke and Neubauer concerning Tikhonov regularization we show convergence of the regularized solution to the true data and study the form of source conditions which ensure convergence rates.

Keywords: inverse problem, financial analysis, volatility calibration, parameter estimation, regularization

MSC2000 classification scheme numbers: 65J20, 91B24, 47H30
1 Introduction

During the last decades a great diversity of price models of financial assets has been developed. It is well-known that as long as it is only possible to observe asset prices (or the corresponding returns) in a discrete scheme, it is always possible to find a model based on a geometric Brownian motion with constant volatility coefficients and stochastic drift terms having identical distributions as the observed returns (cf. [7]). Due to this fact one must not argue that empirically observed returns which fail to have independent normal distributions require extensions of the classical model in order to price options accurately.

On the other hand it is clear that by introducing random effects into the corresponding models via the drift for a given (fixed) behaviour of the observed data there are changes in option prices, even though the option price formula itself is unaffected by changes in the drift. Consequently, the study of corresponding models is meaningful, where the estimate of volatility has to be reinterpreted in the light of the specific model which is assumed.

We consider the price $P(t)$ of a financial asset during the time interval $t \in [0, T]$. By $p(t)$ the logarithm of the asset price is denoted, i.e. $p(t) = \ln P(t)$. The basis for the model which is analyzed in this paper forms the bivariate Ornstein-Uhlenbeck model of Lo and Wang, introduced in [6]. It is assumed that the logarithm of the asset price $p(t)$ has a linear trend $\mu t$. We consider the process

$$q(t) := p(t) - \mu t,$$

which we call in the following detrended log price process to emphasize that $q(t)$ contains no deterministic trend component. We assume that $q(t)$ satisfies the following pair of stochastic differential equations,

$$dq(t) = - (\gamma q(t) - \lambda X(t)) dt + \sigma(t) dW_q(t)$$
$$dX(t) = -\beta X(t) dt + \sigma_X dW_X(t),$$

(1.1)

where $\gamma \geq 0$, $\lambda \geq 0$, $\beta \geq 0$, $\sigma > 0$ and $\sigma_X > 0$ are real parameters, the initial conditions $q(0) = q_0$, $X(0) = X_0$ hold and $W_q$ and $W_X$ are correlated Wiener processes with correlation coefficient $\kappa$, i.e. $E(W_q(t)W_X(t)) = \kappa t$.

In the following we restrict our considerations to the case of independent Wiener processes $W_q$ and $W_X$, i.e. $\kappa = 0$. On the other hand we are interested in a more general behaviour of derivative prices as the constant volatility coefficient $\sigma$ would admit. Therefore we generalize this model inasmuch as we allow the volatility $\sigma$ to be time-dependent (but still non-random).

Assumption 1.1 We assume that the system of stochastic differential equations

$$dq(t) = - (\gamma q(t) - \lambda X(t)) dt + \sigma(t) dW_q(t)$$
$$dX(t) = -\beta X(t) dt + \sigma_X dW_X(t),$$

(1.2)
with $\gamma \geq 0$, $\lambda \geq 0$, $\beta \geq 0$, $\sigma_X > 0$ and a time-dependent volatility function $\sigma(t)$ with $\sigma(t) > 0$, $0 \leq t \leq T$ holds for the detrended log price process $q(t)$. The initial values $q_0 = q(0)$ and $X_0 = X(0)$ are assumed to be stochastic variables with $E q_0 = E X_0 = 0$.

Furthermore, we assume $q_0$, $X_0$, $W_q$, $W_X$ to be mutually independent. The volatility $\sigma_X$ of the drift process $X(t)$ is still supposed to be constant.

Moreover, the existence of a bond with riskless interest rate $r$ is assumed.

Our aim is to identify the parameters in this model. This identification problem can be regarded as an inverse problem. As data we use observed vanilla call option prices and some empirical moments of the logarithmic returns

$$r_{\tau}(t) := p(t) - p(t - \tau) = \ln \left( \frac{P(t)}{P(t - \tau)} \right).$$

To be precise, we assume that we can observe the price data $P(t_k)$ for $t_k = k\tau$, $k = 0, \ldots, N$ (or equivalently the corresponding log prices $p(t_k)$). From these data we calculate the logarithmic returns $r_{\tau}(t_k)$, $k = 1, \ldots, N$, and some derived quantities, which can be interpreted as empirical moments of $r_{\tau}(t_k)$. Hence, we have to take into account that the assumption of the stationarity of the returns is not fulfilled.

At time zero we observe furthermore the prices of European vanilla call options written on the asset with strike $K > 0$ and maturity $t$ varying in

$$I := [0, T] := [0, N\tau].$$

For $t \in I$ we denote by $u(t)$ the price of such an option with maturity $t$. Using these data $u(t)$ ($t \in I$) we try to identify the time-dependent volatility function $\sigma(s)$ ($s \in I$) and the other parameters of interest in (1.2). Clearly, the logarithmic returns $r_{\tau}(t)$ are also influenced by the trend parameter $\mu$. However, an identification of $\mu$ shall not be a part of this paper.

For the seek of simplicity we will restrict our considerations to the two special cases

(i) $\lambda = 0$, where the detrended log price process $q(t)$ is an Ornstein-Uhlenbeck process,

and

(ii) $\gamma = 0$, where $q(t)$ is a stochastic process with independent Ornstein-Uhlenbeck drift $X(t)$.

The paper is organized as follows. In Section 2 we show that the option price formula of Black-Scholes still holds in the cases (i) and (ii). Furthermore we derive explicit formulas for the theoretical moments of $r_{\tau}(t)$ for each of the cases (i) and (ii). These results enable us to formulate the inverse problem under consideration as nonlinear operator equation.
Section 3 discusses properties of the forward operator as well as uniqueness, solvability and ill-posedness of the inverse problem. Based on these properties we apply in Section 4 the theory of [1] concerning nonlinear Tikhonov regularization to the inverse problem. Particularly, we show convergence of the regularized solution to the true data and study the form of source conditions which ensure convergence rates. In Section 5 we formulate conclusions and discuss forthcoming results and ideas.

2 The forward operator

2.1 Option pricing formula

In this section we consider option prices for European call options in the case that the price of the underlying asset follows Assumption 1.1. It may be a little bit surprising that the (with respect to the time-dependent volatility) generalized Black-Scholes option pricing formula still holds in the given situation.

Definition 2.1

For parameters \( \tilde{P} > 0, K > 0, r \geq 0, t \geq 0 \) and \( \tilde{S} \geq 0 \) we define the Black-Scholes function as

\[
U_{BS}(\tilde{P}, K, r, t, \tilde{S}) := \begin{cases} 
\tilde{P}\Phi(d_1) - Ke^{-rt}\Phi(d_2) & \tilde{S} > 0 \\
\max(\tilde{P} - Ke^{-rt}, 0) & \tilde{S} = 0 
\end{cases}
\]

(2.1)

with

\[
d_1 := \frac{\ln\left(\frac{\tilde{P}}{K}\right) + rt + \frac{\tilde{S}}{2}}{\sqrt{\tilde{S}}}, \quad d_2 := d_1 - \sqrt{\tilde{S}}. \tag{2.2}
\]

In (2.1) \( \Phi \) denotes the distribution function of the standard normal distribution.

In the sequel we follow a notation introduced in [3] and [4]. First we set \( a(u) := \sigma^2(u) \) and express by that not directly observable function the volatility term structure. Furthermore we introduce the auxiliary function

\[
S(t) := \int_{0}^{t} a(u) \, du \quad (t \in I). \tag{2.3}
\]

Theorem 2.2

Let the price of the underlying asset follow Assumption 1.1. With respect to the following considerations we restrict ourselves to the cases \( \lambda = 0 \) and \( \gamma = 0 \), respectively. Using the initial asset price \( P_0 = P(0) = e^{\varphi(0)} \) the fair option price \( u(t) \) is given by

\[
u(t) = U_{BS}(P_0, K, r, t, S(t)) \quad (t \in I), \tag{2.4}
\]

where \( S(t) \) is defined by formula (2.3) with the underlying volatility term structure \( a(\cdot) \).
\textbf{Proof:} Consider the general model for the price of a financial asset \( P(s) \) as solution of the stochastic differential equation

\[ dP(s) = \sigma(s)P(s)\,dW_q(s) + P(s)\,dZ(s), \quad s \in I, \]

where \( Z(s) \) is a continuous random process of zero square variation (possibly dependent on \( P \)) and \( \sigma(s) > 0 \) is a continuous deterministic function. We mention that the considered models with \( \lambda = 0 \) and \( \gamma = 0 \) are special cases of this model with

\[ dZ(s) = \left( -\gamma (\ln P(s) - \mu s) + \frac{\sigma^2(s)}{2} + \mu \right) \, ds \]

in the case \( \lambda = 0 \) (which means that \( Z(s) \) is adapted to the filtration generated by \( W_q \)) and

\[ dZ(s) = \left( \lambda X(s) + \frac{\sigma^2(s)}{2} + \mu \right) \, ds \]

in the case \( \gamma = 0 \) (in this case, \( Z(s) \) is independent on \( W_q \)).

On the one hand we mention that it is possible to show that in both situations there exists an equivalent measure such that the discounted price process \( e^{-rs}P(s) \) is a martingale with respect to this measure. This, however, will not be part of this paper.

Following the argumentation of [8], instead the existence of a self-financing trading strategy with an initial capital calculated by the Black-Scholes equation is shown. We fix a value \( t \in I \) recalling that \( t \) plays the role of the maturity. Suppose that \( C(s, x) \) satisfies the Black and Scholes parabolic differential equation

\[ C_s + \frac{1}{2}x^2\sigma^2(s)C_{xx} + rxC_x - rC = 0 \quad \text{(2.5)} \]

with final value \( C(t, x) = \max(x - K, 0) \). Note that from the theory of option pricing in the classical model of a geometric Brownian motion with time-dependent volatility it is well-known that

\[ C(s, x) = U_{BS}(x, K, r, t - s, S(t) - S(s)) \]

is a function with this property.

We show that there exists a replicating self-financing portfolio of the claim with value \( C(s, P(s)) \) at time \( s \). For this, we consider the portfolio containing

\[ C_x(s, P(s)) \]

shares of the stock and an amount of money equal to

\[ C(s, P(s)) - P(s)C_x(s, P(s)). \]

Then the total value of the portfolio \( V(s, P(s)) \) is equal to \( C(s, P(s)) \) for all \( s, 0 \leq s \leq t \). Furthermore it holds \( V(t, P(t)) = \max(P(t) - K, 0) \).
By Itô’s Lemma it follows that we have in the considered situations
\[ dV(s, P(s)) = C_s(s, P(s)) \, ds + C_x(s, P(s)) \, dP(s) + \frac{1}{2} \sigma^2(s) P_x(x(s, P(s))) \, ds , \]
where (2.5) has been used. We see that the constructed portfolio is self-financing and replicates the pay-off-value of the call option with probability one. Hence it follows that
\[ C(0, P_0) = U_{BS}(P_0, K, r, t, S(t)) \]
is indeed a fair option price at time zero.

Next we will derive explicit formulas for the theoretical moments of the logarithmic returns \( r_\tau(t) \).

### 2.2 The special case \( \lambda = 0 \)

In this subsection we will consider the case (i), where \( \lambda = 0 \), i.e. the stochastic differential equation (1.2) for the detrended log price process \( q(t) \) reduces to
\[ dq(t) = -\gamma q(t) \, dt + \sigma(t) \, dW_q(t). \quad (2.6) \]

Thus, the detrended log price process \( q(t) \) is mean-reverting, i.e. if \( q(t) \) deviates from zero the term \( -\gamma q(t) \, dt \) ensures that it is pulled back with a rate proportional to its deviation, where the parameter \( \gamma \) characterizes the speed of adjustment. The last term in (2.6) adds the randomness to the increment \( dq(t) \). Here \( \sigma(t) \) is the time-dependent volatility.

The explicit solution of (2.6),
\[ q(t) = e^{-\gamma t} q_0 + \int_0^t \sigma(s) e^{-\gamma (t-s)} \, dW_q(s), \quad (2.7) \]
shows that the influence of the initial value decreases with increasing \( t \).

Straightforward calculation gives
\[ E q(t) = E \left\{ e^{-\gamma t} q_0 + \int_0^t \sigma(s) e^{-\gamma (t-s)} \, dW_q(s) \right\} = 0 \]
and using \( q(t) = p(t) - \mu t \) we get
\[ E r_\tau(t) = E \{ p(t) - p(t - \tau) \} = E \{ \mu \tau + q(t) - q(t - \tau) \} = \mu \tau . \]

The results for the variance, covariance and correlation structures of \( r_\tau(t) \) are shown in Lemma 2.3. For the proof we refer to the appendix. Note, that we will use only covariances and correlations of the form \( \text{Cov} (r_\tau(t), r_\tau(t + k\tau)) \) or \( \text{Corr} (r_\tau(t), r_\tau(t + k\tau)) \).
Lemma 2.3
Using the notation
\[ H(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \sigma^2(s) e^{-2\gamma(t_2-s)} \, ds = \int_{\tau_1}^{\tau_2} a(s) e^{-2\gamma(t_2-s)} \, ds \]
it holds
\[ D^2 r_\tau(t) = (1 - e^{\gamma\tau})^2 \left( e^{-2\gamma t} E q_0^2 + e^{-2\gamma \tau} H(0, t-\tau) \right) + H(t-\tau, t) \] (2.8)
and for \( k \in \mathbb{N} := \{1, 2, 3, \ldots \} \)
\[ \text{Cov} (r_\tau(t), r_\tau(t+k\tau)) = e^{-\gamma(k\tau)} \left( (1 - e^{-\gamma\tau})^2 H(0, t-\tau) + (1 - e^{\gamma\tau}) H(t-\tau, t) \right) + e^{-\gamma(2t+k\tau)} E q_0^2 (1 - e^{\gamma\tau})^2 \] (2.9)
as well as
\[ \text{Corr} (r_\tau(t), r_\tau(t+k\tau)) = \frac{\text{Cov} (r_\tau(t), r_\tau(t+k\tau))}{\sqrt{D^2 r_\tau(t)} \sqrt{D^2 r_\tau(t+k\tau)}}. \] (2.10)

For \( \gamma = 0 \) we get the (with respect to a time-dependent volatility generalized) Black-Scholes model, where
\[ dq(t) = \sigma(t) \, dW(t). \]
In this case (2.8) and (2.9) simplify to
\[ D^2 r_\tau(t) = \int_{\tau-\tau}^{\tau} \sigma^2(s) \, ds = \int_{\tau-\tau}^{\tau} a(s) \, ds \] (2.11)
and
\[ \text{Cov} (r_\tau(t), r_\tau(t+k\tau)) = 0, \quad k \in \mathbb{N}. \] (2.12)
The next lemma characterizes the case \( \sigma(t) \equiv \sigma \), where the integral \( H(\tau_1, \tau_2) \) can be computed explicitly.

Lemma 2.4
In the case of a constant volatility function \( \sigma(t) \equiv \sigma \) and \( \gamma > 0 \) variance and covariance of the logarithmic returns are given by
\[ D^2 r_\tau(t) = e^{-2\gamma t} (1 - e^{\gamma\tau})^2 \left( E q_0^2 - \frac{\sigma^2}{2\gamma} \right) + \frac{\sigma^2}{\gamma} (1 - e^{-\gamma\tau}) \] (2.13)
and for \( k \in \mathbb{N} \)
\[ \text{Cov} (r_\tau(t), r_\tau(t+k\tau)) = e^{-\gamma(2t+k\tau)} (1 - e^{\gamma\tau})^2 \left( E q_0^2 - \frac{\sigma^2}{2\gamma} \right) - \frac{\sigma^2}{2\gamma} (1 - e^{-\gamma\tau})^2 e^{-\gamma(k-1)\tau}. \] (2.14)
Furthermore, in this special case the processes \( q(t) \) and \( r_\tau(t) \) are asymptotically stationary, where roughly speaking a random process \( X(t) \) is called asymptotically stationary if the distribution of \( (X(t_1 + s), \ldots, X(t_n + s)) \) does not depend on \( s \) when \( s \) is large.

Hence for \( t \to \infty \) the empirical moments converge to finite values, which are shown in the following table.

<table>
<thead>
<tr>
<th>( \gamma &gt; 0 )</th>
<th>( \gamma = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lim_{t \to \infty} D^2 r_\tau(t) )</td>
<td>( \frac{\sigma^2}{\gamma} (1 - e^{\gamma\tau}) )</td>
</tr>
<tr>
<td>( \lim_{t \to \infty} \text{Cov}(r_\tau(t), r_\tau(t + k\tau)) )</td>
<td>( -\frac{\sigma^2}{2\gamma} (1 - e^{\gamma\tau})^2 e^{\gamma(k-1)\tau} )</td>
</tr>
<tr>
<td>( \lim_{t \to \infty} \text{Corr}(r_\tau(t), r_\tau(t + k\tau)) )</td>
<td>( \frac{1 - e^{\gamma\tau}}{2e^{\gamma k\tau}} )</td>
</tr>
</tbody>
</table>

The calculations which are necessary for the proof of Lemma 2.4 are given in the appendix.

However, if we admit the volatility to be time-dependent, clearly the processes \( r_\tau(t) \) and \( q(t) \) are not asymptotically stationary.

Next, we will assume that there exists at least one interval on which \( \sigma(t) \) and thus also \( a(t) = \sigma^2(t) \) is constant. That is, we have a decomposition of the interval \( I \) in \( n \) intervals

\[
I_i = \begin{cases} 
[b_{i-1}, b_i] & 1 \leq i < n \\
[b_{i-1}, b_n] & i = n,
\end{cases}
\]  

(2.15)

where \( b_0 := 0 \) and \( b_n := T \), and functions \( a_i : I_i \to \mathbb{R} \) such that \( a(t) \) satisfies

\[
a(t) = a_i(t) \leq c_i \quad \forall t \in I_i
\]  

(2.16)

and at least one of the functions \( a_i(t) \) is constant. We will assume that the number of intervals \( n \) is small, especially \( n << N \).

For the moment we will assume that all \( a_i(t) \) are constant, i.e. \( a_i(t) \equiv c_i \) for some \( c_i \in \mathbb{R} \). It should be mentioned that although this assumption is in contrast to the assumptions in Section 2.1 this has clearly no consequences for the validity of the option pricing formula. Using the indicator function

\[
\chi_i(t) = \begin{cases} 
1 & t \in I_i \\
0 & t \notin I_i
\end{cases}
\]  

(2.17)

in this case we can write \( a(t) \) as

\[
a(t) = \sum_{i=1}^{n} c_i \chi_i(t).
\]
Thus, the integral $H(0, t)$ for $t \in I$ can be computed explicitly as

$$H(0, t) = \int_0^t a(s)e^{-2\gamma(t-s)} \, ds$$

$$= \sum_{i=1}^n c_i \int_{\min(b_i,t)}^{\min(b_i,t+1)} e^{-2\gamma(t-s)} \, ds = \frac{1}{2\gamma} \sum_{i=1}^n c_i \left( e^{-2\gamma(t-\min(t,b_i))} - e^{-2\gamma(t-\min(t,b_i-1))} \right).$$

For $\gamma > 0$ and times $t$ with $b_{j-1} < t - \tau < t \leq b_j$ for some $j$ we can write $D^2r_\tau(t)$ as

$$D^2r_\tau(t) = (1 - e^{\gamma \tau})^2 \left( e^{-2\gamma t} E_{b_0}^2 + \frac{e^{2\gamma \tau}}{2\gamma} \sum_{i=1}^{j-1} c_i \left( e^{-2\gamma(t-\tau-b_i)} - e^{-2\gamma(t-\tau-b_{i-1})} \right) + c_j \left( 1 - e^{-2\gamma(t-\tau-b_{j-1})} \right) \right) + \frac{1}{2\gamma} (c_j (1 - e^{-2\gamma \tau}))$$

This formula shows that the influence of $E_{b_0}^2$ decreases for increasing $t$. Moreover, the influence of terms $c_i \left( e^{-2\gamma(t-\tau-b_i)} - e^{-2\gamma(t-\tau-b_{i-1})} \right)$ decreases if $b_i << t$. Hence, if we have a piecewise constant volatility function $\sigma(t)$, where the interval $I_j$ is relatively large and $b_{j-1} << t - \tau < t \leq b_j$ the value $D^2r_\tau(t)$ can be approximated by

$$(1 - e^{\gamma \tau})^2 \frac{1}{2\gamma} c_j e^{-2\gamma \tau} \left( 1 - e^{-2\gamma(t-\tau-b_{j-1})} \right) + \frac{1}{2\gamma} c_j \left( 1 - e^{-2\gamma \tau} \right)$$

$$= \frac{c_j}{\gamma} (1 - e^{-\gamma \tau}) - \frac{c_j}{2\gamma} e^{-2\gamma(t-\tau-b_{j-1})} (1 - e^{\gamma \tau})^2 \approx \frac{c_j}{\gamma} (1 - e^{-\gamma \tau}).$$

We will now admit functions $a_i(t)$ ($i < j$) in (2.16) which are not constant, but still bounded from above. Then we have for $b_{j-1} < t - \tau < t \leq b_j$

$$(1 - e^{\gamma \tau})^2 \left( e^{-2\gamma t} E_{b_0}^2 + \frac{e^{2\gamma \tau}}{2\gamma} \sum_{i=1}^{j-1} c_i \left( e^{-2\gamma(t-\tau-b_i)} - e^{-2\gamma(t-\tau-b_{i-1})} \right) \right) + \frac{1}{2\gamma} (c_j (1 - e^{-2\gamma \tau}))$$

$$\leq D^2r_\tau(t)$$

$$\leq (1 - e^{\gamma \tau})^2 \left( e^{-2\gamma t} E_{b_0}^2 + \frac{e^{2\gamma \tau}}{2\gamma} \sum_{i=1}^{j-1} c_i \left( e^{-2\gamma(t-\tau-b_i)} - e^{-2\gamma(t-\tau-b_{i-1})} \right) \right) + c_j \left( 1 - e^{-2\gamma(t-\tau-b_{j-1})} \right) + \frac{1}{2\gamma} (c_j (1 - e^{-2\gamma \tau}))$$

As above the lower bound and the upper bound can be approximated by $\frac{c_j}{\gamma} (1 - e^{-\gamma \tau})$ provided $b_{j-1} << t - \tau < t \leq b_j$.

Furthermore we see from (2.11) that for $\gamma = 0$ the variance $D^2r_\tau(t)$ is equal to $c_j \tau$ provided $b_{j-1} < t - \tau < t \leq b_j$. Hence, defining the auxiliary function

$$\hat{h}_\tau(c, \gamma) := \begin{cases} \frac{c}{\gamma} (1 - e^{-\gamma \tau}) & \gamma > 0 \\ \frac{c}{\gamma^2} & \gamma = 0 \end{cases}$$

(2.18)
and assuming that the interval $I_j$ is sufficiently long we have $D^2 r_\tau(t) \approx \hat{h}_\tau(c_j, \gamma)$ for all $t$, for which $0 \leq b_i - t$ is sufficiently small. Thus, $\hat{h}_\tau(c_j, \gamma)$ can be seen as approximation of the variance at the point $t$ if $t < b_j$ and $t \approx b_j$. To illustrate that this approximation is in practical situations sufficiently good, we consider model (2.6) with $T = 1000$, $q_0 \equiv 0$, $\gamma = 0.104$, $\tau = 0.1$ and the volatility function

$$\sigma(t) = \begin{cases} 
0.0175 & t < 200 \\
0.0325 & 200 \leq t < 400 \\
0.0300 & 400 \leq t < 600 \\
0.0375 & 600 \leq t < 800 \\
0.0225 & 800 \leq t < 1000 
\end{cases},$$

which is displayed in Figure 2.1. The corresponding variance $D^2 r_\tau(t)$ is shown in Figure 2.2. For this parameter constellation we do not find any deviation of $D^2 r_\tau(t)$ from its approximation $\frac{\sigma^2(t)}{\gamma} (1 - e^{-\gamma \tau}) = \frac{a(t)}{\gamma} (1 - e^{-\gamma \tau})$, which is piecewise constant on the interval $I_i$ ($1 \leq i \leq 5$).

As the approximation $D^2 r_\tau(t) \approx \hat{h}_\tau(a, \gamma)$ seems to be sufficiently good we will use the term

$$\frac{1}{2n} \sum_{k=j-n}^{j+n} \left( r_\tau(t_k) - \frac{1}{2n+1} \sum_{i=j-n}^{j+n} r_\tau(t_i) \right)^2,$$  \hspace{1cm} (2.19)

which can be viewed as estimation of the variance of $r_\tau(t_j)$, to approximate $\hat{h}_\tau(a(t_j), \gamma)$. The approximation is expected to be quite satisfactory for long intervals $I_i$ on which $a_i(t)$ is constant.

With respect to the inverse problem we will assume that we know at least one interval $I_i$ on which the volatility term structure $a(t) = a_i(t)$ is constant. For notational convenience
we introduce a set $J_0 \subset \{1, 2, \ldots, n\}$ such that we know in advance that for all $i \in J_0$ the function $a(t)$ is constant on the interval $I_i$.

### 2.3 The special case $\gamma = 0$

In this section, we consider the case (ii), where $\gamma = 0$, i.e. $q(t)$ is a stochastic process with mean-reverting drift $X(t)$:

$$
\begin{align*}
\, dq(t) &= \lambda X(t) \, dt + \sigma(q) \, dW_q(t) \\
\, dX(t) &= -\beta X(t) \, dt + \sigma_X \, dW_X(t).
\end{align*}
$$

We can set $\lambda = 1$ without loss of generality. Indeed, for $\lambda \neq 1$ we could consider the drift process $\tilde{X}(t)$ determined by the stochastic differential equation

$$
\, d\tilde{X}(t) = -\beta \tilde{X}(t) \, dt + \tilde{\sigma} X \, dW_X(t),
$$

where $\tilde{X}(0) = \tilde{X}_0 := \lambda X_0$ and $\tilde{\sigma} := \lambda \sigma_X$. Then we would get

$$
\, dq(t) = \tilde{X}(t) \, dt + \sigma(t) \, dW_q(t).
$$

Using the explicit solutions

$$
\begin{align*}
\, X(t) &= e^{-\beta t} X_0 + \sigma_X \int_0^t e^{-\beta(t-s)} \, dW_X(s) \quad \text{(2.21)} \\
\, q(t) &= q_0 + \int_0^t X(s) \, ds + \int_0^t \sigma(s) \, dW_q(s) \quad \text{(2.22)}
\end{align*}
$$

we get

$$
\begin{align*}
\, q(t) &= q_0 - \frac{1}{\beta} \left[ e^{-\beta t} - 1 \right] X_0 - \frac{\sigma_X}{\beta} \int_0^t \left( e^{-\beta(t-s)} - 1 \right) \, dW_X(s) + \int_0^t \sigma(s) \, dW_q(s). \quad \text{(2.23)}
\end{align*}
$$

We can therefore easily calculate the first-order moments

$$
EX(t) = 0, \quad Eq(t) = 0
$$

and thus

$$
E r_\tau(t) = E \{ p(t) - p(t - \tau) \} = E \{ \mu t + q(t) - \mu (t - \tau) - q(t - \tau) \} = \mu \tau.
$$

The results for the variance and covariance functions of $r_\tau$ are given in Lemma 2.5, for the proof we refer again to the appendix.
Lemma 2.5
For the model considered in this section it holds
\[
D^2 r_\tau(t) = \frac{1}{\beta^2} \left[ e^{-\beta(t-\tau)} - e^{-\beta t} \right]^2 \left( \frac{\sigma^2 X_0}{2\beta} \right) + \int_{t-\tau}^t \sigma^2(s) \, ds \\
+ \frac{\sigma^2 X}{\beta^3} \left( -1 + e^{-\beta \tau} + \beta \tau \right),
\]
and for \( k \in \mathbb{N} \) we have
\[
\text{Cov} \left( r_\tau(t), r_\tau(t + k\tau) \right) = \frac{1}{\beta^2} e^{-\beta k\tau} \left( e^{-\beta(t-\tau)} - e^{-\beta t} \right)^2 \left( \frac{\sigma^2 X_0}{2\beta} \right) \\
+ \frac{\sigma^2 X}{2\beta^3} e^{-\beta(k+1)\tau} \left( e^{\beta \tau} - 1 \right)^2.
\]

The last equation in Lemma 2.5 shows, that for \( \gamma = 0 \) the volatility does not influence the covariance between \( r_\tau(t) \) and \( r_\tau(t + k\tau) \). Furthermore, the influence of the time point \( t \) in \( \text{Cov} \left( r_\tau(t), r_\tau(t + k\tau) \right) \) tends to zero if \( t \) tends to infinity. Thus, it seems natural to use the asymptotic covariance
\[
\lim_{\tau \to \infty} \text{Cov} \left( r_\tau(t), r_\tau(t + i\tau) \right) = \frac{\sigma^2 X}{2\beta^3} e^{-\beta(i+1)\tau} \left( e^{\beta \tau} - 1 \right)^2 =: \tilde{h}_i((\beta, \sigma_X)) \quad (2.24)
\]
for several lags \( i \geq 1 \) as data for the calibration of the parameters \( \sigma_X \) and \( \beta \). For the seek of a uniform notation with respect to the following sections we introduce again a set \( J_0 \subset \{1, 2, \ldots, n\} \), which contains in this situation the used lags.

Furthermore, we obtain asymptotically
\[
D^2 r_\tau(t) \approx \int_{t-\tau}^t \sigma^2(s) \, ds + \frac{\sigma^2 X}{\beta^3} \left( -1 + e^{-\beta \tau} + \beta \tau \right).
\]
We can therefore expect that for smooth \( \sigma(t) \) the estimation of \( D^2 r_\tau(t) \) given in (2.19) contains some information about the shape of \( \sigma(t) \).

2.4 Formulation of the inverse problem

We are now able to define the inverse problem under consideration. In order to do this in a general way, which includes the case \( \lambda = 0 \) as well as the case \( \gamma = 0 \) we introduce some notation. Let \( \beta^*, \sigma^*_X, \gamma^* \) and \( \sigma^*(t) \) denote the exact parameter values of the model (1.2). Similarly we write \( a^*(t) = (\sigma^*(t))^2 \).

Let \( d_p \) denote the number of parameters in (1.2) which have to be determined in addition to \( \sigma(t) \), i.e.
\[
d_p = \begin{cases} 
1 & \lambda = 0 \\
2 & \gamma = 0
\end{cases}
\]
and $d_m$ the number of second order moments, which we use as data.

In the parameter vector $p^* \in \mathbb{R}^{d_p}$ we collect for each special case the exact real valued parameters in the considered model. Furthermore, we collect in the vector $m^* \in \mathbb{R}^{d_m}$ the asymptotic theoretical moments of $r_0(t)$ which we use as data for the estimation.

That is, in the special case (i) ($\lambda = 0$), where we assume $a^* = a_i^* = c_i$ on the interval $I_i$, we introduce the indices $i_1 < i_2 < \ldots < i_{d_m}$ such that $J_0 = \{i_1, i_2, \ldots, i_{d_m}\}$ and set

$$p^* := (\gamma^*) \quad \text{and} \quad m^* := \left( \frac{1}{b_{i_j} - b_{i_{j-1}}} \int_{b_{i_{j-1}}}^{b_{i_j}} a^*(t) \, dt, \gamma^* \right)_{j=1}^{d_m}.$$ 

On the other hand, in the special case (ii) ($\gamma = 0$), where we want to use the covariances for several lags $i \in J_0$, we introduce also indices $i_1 < i_2 < \ldots < i_{d_m}$ such that $J_0 = \{i_1, i_2, \ldots, i_{d_m}\}$ and set

$$p^* := (\beta^*, \sigma_X^*)^T \quad \text{and} \quad m^* := \left( \tilde{h}_{i_j}(\beta^*, \sigma_X^*) \right)_{j=1}^{d_m}.$$ 

Our aim is to identify the volatility term structure $a^*$ and the parameter vector $p^*$ from noisy data $(u^\delta, m^\delta)$. Using the notation $(v)_j$ for the $j$-th component of the vector $v$ we can formulate this task as the following specific inverse problem.

**Definition 2.6**

**Specific inverse problem – SIP**

Given a square-integrable noisy data function $u^\delta(t)$ ($t \in I$) and empirical moments $m^\delta$ with noise level $\delta > 0$

$$\| (u^\delta, m^\delta) - (u^*, m^*) \|_{L^2(I) \times \mathbb{R}^{d_p}} = \left( \int_I \left( u^\delta(t) - u^*(t) \right)^2 \, dt + \sum_{j=1}^{d_m} \xi_j \left( (m^\delta)_j - (m^*)_j \right)^2 \right)^{1/2} \leq \delta,$$ (2.25)

find appropriate approximations $a^\delta, p^\delta$ of the function $a^*$ and the parameter vector $p^*$, where both $a^\delta$ and $a^*$ are integrable and almost everywhere nonnegative functions defined on $I$ and

$$p^\delta \in \begin{cases} \{(p_1) : p_1 \geq 0\} & \lambda = 0 \\ \left\{ \left( \begin{array}{c} p_1 \\ p_2 \end{array} \right) : p_1 > 0, p_2 \geq 0 \right\} & \gamma = 0 \end{cases}.$$ 

The accuracy of the pair $(a^\delta, p^\delta)$ is measured by

$$\| (a^\delta, p^\delta) - (a^*, p^*) \|_{B_1} = \left( \int_0^T (a^\delta(t) - a^*(t))^2 \, dt + \sum_{j=1}^{d_p} \left( (p^\delta)_j - (p^*)_j \right)^2 \right)^{1/2}. \quad (2.26)$$
It should be mentioned that the space $B_1 = L^2(I) \times \mathbb{R}^{d_p}$ with the norm (2.26) is a Hilbert space. The corresponding scalar product is given by

$$\langle (a_1, p_1), (a_2, p_2) \rangle_{B_1} = \int_0^T a_1(t) a_2(t) \, dt + \sum_{j=1}^{d_p} (p_1)_j (p_2)_j .$$

Analogously we define a scalar product in $B_2 = L^2(I) \times \mathbb{R}^{d_m}$ compatible with the norm used in (2.25). Furthermore, we will use the Hilbert spaces $X_p = \{\mathbb{R}^{d_p}, \| \cdot \|_{X_p} \}$ and $X_m = \{\mathbb{R}^{d_m}, \| \cdot \|_{X_m} \}$ attributed with the scalar products

$$\langle p, q \rangle_{X_p} = \sum_{j=1}^{d_p} (p)_j (q)_j$$

and

$$\langle p, q \rangle_{X_m} = \sum_{j=1}^{d_m} \xi_j (p)_j (q)_j .$$

The factors $\xi_j > 0$ are weights, which enable us to weight the information contained in the $j$-th component of the vector $m$. Especially in the case (i) ($\lambda = 0$) we could use large $\xi_j$ for long intervals $I_{ij}$ and small $\xi_j$ for short intervals $I_{ij}$, which reflects that the empirical variance is expected to be closer to the theoretical variance for long intervals $I_{ij}$.

Note that we defined the Hilbert spaces $B_1$, $B_2$, $X_m$ and $X_p$ such that it holds

$$\langle (a, p), (b, q) \rangle_{B_1} = \langle a, b \rangle_{L^2(I)} + \langle p, q \rangle_{X_p} \quad (2.27a)$$

and

$$\langle (a, p), (b, q) \rangle_{B_2} = \langle a, b \rangle_{L^2(I)} + \langle p, q \rangle_{X_m} \quad (2.27b)$$

respectively, where

$$\langle a, b \rangle_{L^2(I)} := \int_I a(t)b(t) \, dt$$

denotes the common scalar product in $L^2(I)$.

From now on let $L^2_+(I)$ denote the space of all nonnegative square integrable functions of $I$. Using these Hilbert spaces $B_1$, $B_2$ we can write the inverse problem (SIP) as a nonlinear operator equation

$$F \left( (a, p) \right) = (u, m) \quad ((a, p) \in D(F) \subset B_1, \ (u, m) \in L^2(I) \times \mathbb{R}^{d_m}_+ \subset B_2) , \quad (2.28)$$

where the nonlinear forward operator

$$F = F_1 \times F_2 : D(F) \subset B_1 \rightarrow B_2 \quad (2.29)$$
with the convex domain

\[ D(F) = \begin{cases} 
L^2(I) \times \{(p_1) \in \mathbb{R}^1 : p_1 \geq 0\} & \lambda = 0 \\
L^2(I) \times \left\{ \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \in \mathbb{R}^2 : p_1 > 0, p_2 \geq 0 \right\} & \gamma = 0 
\end{cases} \]

is decomposed into an operator

\[ F_1 : D(F_1) = L^2(I) \subset L^2(I) \rightarrow L^2(I) \quad F_1 : a \mapsto u , \]

with \( u \) given by formulas (2.1)-(2.4), and an operator

\[ F_2 : D(F_2) = D(F) \subset B_1 \rightarrow L^2(I) \times \mathbb{R}^m_+, \]

\[ (a, p) \mapsto m = \left\{ \begin{pmatrix} \hat{h}_r \left( \frac{1}{b_{ij} - b_{i-1}} \int_{b_{i-1}}^{b_{ij}} a(t) \ dt, p_1 \right) \\ \hat{h}_i \left( p \right) \right\} \Bigg|_{j=1}^{d_m} \lambda = 0 \]

\[ \gamma = 0 . \quad (2.30) \]

In the case (i) (\( \lambda = 0 \)) we will write \( p \) instead of \( p_1 \). Furthermore we will use the notation \( \tilde{F}_1, \tilde{F}_2 \) and \( \tilde{F}_1, \tilde{F}_2 \) for the operators \( F, F_2 \) with \( \lambda = 0 \) and \( \gamma = 0 \) respectively. The pair \((u^*, m^*)\) will be called exact right hand side of the operator equation (2.28)

Note that the operator \( \tilde{F}_2 \) does not depend on \( a^* \). That is, we can introduce an operator \( \tilde{F}_2 : X_p \rightarrow X_m \) as follows:

\[ \tilde{F}_2 (p) := \tilde{F}_2 ((a^*, p)) . \]

Therefore, in that case the operator equation

\[ F ((a, p)) = (u, m) \]

can be decomposed and occurs in form of the two independent equations

\[ F_1 (a) = u \quad \text{(SIP1)} \]

and

\[ \tilde{F}_2 (p) = m . \quad \text{(SIP2)} \]

## 3 Properties of the forward operator and ill-posedness of the inverse problem

In order to characterize the forward operator \( F \) in the pair of Hilbert spaces \( B_1 \) and \( B_2 \), we can use the following results of [3]:

**Proposition 3.1**

The nonlinear operator \( F_1 : D(F_1) \subset L^2(I) \rightarrow L^2(I) \) is compact, continuous and injective. Thus, the inverse operator \( F_1^{-1} \) defined on the range \( F_1(D(F_1)) \) of \( F_1 \) exists.
From the injectivity of the operator $F_1$ it follows that the volatility term structure $a^*(\tau)$ is uniquely determined by the exact option prices $u^*(t)$ $(t \in I)$. Furthermore, Theorem 3.4 in [3] gives a sufficient condition for the solvability of the operator equation

$$F_1(a) = u \quad (a \in D(F_1), u \in L^2_+(I)).$$

Next we study properties of the operator $F_2$. First we consider the case $(i)$ $(\lambda = 0)$. In order to study the function $\hat{h}_\tau$ and its derivatives we introduce the auxiliary functions

$$f_n(y) := \begin{cases} 
(-1)^n \frac{n!}{y^{n+1}} \sum_{k=0}^{\frac{1}{y^n}} 1^{y^k} & y > 0 \\
(-1)^n \frac{1}{n+1} & y = 0
\end{cases}.$$  \hspace{1cm} (3.2)

Some properties of these auxiliary functions, which we will use later on, are presented in the following lemma. For the proof we refer again to the appendix.

**Lemma 3.2**

The functions $f_n(y)$ have the following properties

1. $(f_n(y))' = f_{n+1}(y)$
2. For $y \geq 0$ we have $\begin{cases} f_n(y) > 0 & \text{if } n \text{ is even} \\
f_n(y) < 0 & \text{if } n \text{ is odd}
\end{cases}$
3. The function $f_n(y)$ is strictly monotonically $\begin{cases} \text{decreasing} & \text{if } n \text{ is even} \\
\text{increasing} & \text{if } n \text{ is odd}
\end{cases}$
4. $\lim_{y \to 0} f_n(y) = (-1)^n \frac{1}{n+1}$ and $\lim_{y \to \infty} f_n(y) = 0$

Substituting $y = \gamma \tau$ we can now write

$$\hat{h}_\tau(c, \gamma) = c \tau f_0(\gamma \tau).$$

Therefore we get immediately the following corollary.

**Corollary 3.3**

The function $\hat{h}_\tau(\cdot, \cdot) : \mathbb{R}_+ \times \mathbb{R}_+ \to [0, \infty)$ is continuous. Furthermore, for every $c > 0$ the function

$$\hat{h}_{c,\tau}(\cdot) := \hat{h}_\tau(c, \cdot), \quad \hat{h}_{c,\tau} : \mathbb{R}_+ \to [0, c\tau] \subset \mathbb{R}_+$$

is continuous, strictly decreasing and surjective.

From the injectivity and surjectivity of $\hat{h}_{c,\tau}(\gamma)$ follows that if $d_m = 1$ and we assume to know the exact volatility term structure $a^*(t)$ and therefore the exact value

$$e_{i_1}^* = \frac{1}{b_{i_1} - b_{i_1-1}} \int_{b_{i_1-1}}^{b_{i_1}} a^*(t) \, dt,$$
the operator equation $\hat{F}_2(p) = m$ is uniquely solvable for $0 \leq m_1 \leq c^{*}_i \tau$. Thus, if we have exact data, we are able to find the exact solution.

In order to study convergence and weak convergence of the operator $F_2$, we need a characterization of weak convergence in the Hilbert spaces $B_1$ and $B_2$. Using (2.27) we see that weak convergence $(a_n, p_n) \xrightarrow{B_1} (a, p)$ of a sequence $(a_n, p_n)$ to some element $(a, p)$ in $B_1$ is equivalent to the weak convergences

$$\begin{align*}
a_n &\xrightarrow{L^2(I)} a \quad \text{and} \quad p_n \xrightarrow{X_p} p.
\end{align*}$$

Furthermore, in the finite dimensional spaces $X_p$ and $X_m$ a weakly convergent sequence is also strongly convergent. Thus, we have

$$\begin{align*}
(a_n, p_n) &\xrightarrow{B_2} (a, p) \iff a_n \xrightarrow{L^2(I)} a \quad \text{and} \quad p_n \xrightarrow{X_m} p.
\end{align*}$$

Now we are able to prove continuity and weak continuity of the operator $\hat{F}_2$.

**Lemma 3.4**

*The operator $\hat{F}_2$ is continuous and weakly continuous.*

**Proof:** Let

$$\begin{align*}
(a_n, p_n) &\xrightarrow{B_1} (a_0, p_0).
\end{align*}$$

We will show $\hat{F}_2 ((a_n, p_n)) \to \hat{F}_2 ((a_0, p_0))$.

As every strong convergent sequence is also weak convergent this implication shows the continuity of $\hat{F}$ as well as the weak continuity of $\hat{F}$.

The convergence (3.3) implies $a_n \to a_0$ in $L^2(I)$ and therefore

$$\begin{align*}
\left\langle a_n, \frac{1}{b_{i_j} - b_{i_j-1}} \chi_{i_j} \right\rangle_{L^2(I)} &\to \left\langle a_0, \frac{1}{b_{i_j} - b_{i_j-1}} \chi_{i_j} \right\rangle_{L^2(I)},
\end{align*}$$

where $\chi_{i_j}$ denotes the indicator function defined in (2.17). Using the continuity of $\hat{h}_\tau (c, \gamma)$ we get

$$\hat{F}_2 ((a_n, p_n)) \to \hat{F}_2 ((a_0, p_0)).$$

Now we consider the case (ii) ($\gamma = 0$). There we use the auxiliary functions $\tilde{h}_{i_j} ((\beta, \sigma X))$ for several values $i_j \in J_0$. The next lemma gives some properties of these functions. For the proof we refer again to the appendix.
Lemma 3.5
The functions \( \tilde{h}_k : (0, \infty) \times \mathbb{R}_+ \to (0, \infty) \) are for every \( k \in \mathbb{N} \) continuous. Furthermore the limit relations
\[
\lim_{\beta \to 0} \tilde{h}_k ((\beta, \sigma_X)) = \infty \quad \text{and} \quad \lim_{\beta \to \infty} \tilde{h}_k ((\beta, \sigma_X)) = 0
\]
hold.

Now we can prove the following lemma concerning continuity of the forward operator and solvability of the operator equation (2.30).

Lemma 3.6
The operator \( \tilde{F}_2 \) has the following properties.

1. \( \tilde{F}_2 \) is continuous.

2. If \( d_m = 2 \) the range of \( \tilde{F}_2 \) is given by
\[
\tilde{F}_2 \left( D \left( \tilde{F}_2 \right) \right) = \left\{ m \in \mathbb{R}_+^2 : (m)_1 > (m)_2 > 0 \right\} \cup \{ 0 \}.
\]

Furthermore, the operator equation \( \tilde{F}_2 (a, p) = m \) has a unique solution for \( p \) if and only if \( (m)_1 > (m)_2 > 0 \).

Proof:

1. The continuity of the operator \( F_2 \) follows from the continuity of the functions
\[
\tilde{h}_i(\beta, \sigma_X) = \frac{\sigma_X^2}{2\beta^3} e^{-\beta(i+1)\tau} (e^{\beta\tau} - 1)^2
\]
in the set
\[
\left\{ \left[ \begin{array}{c} \beta \\ \sigma_X \end{array} \right] : \beta > 0, \sigma_X \geq 0 \right\}
\]
for all \( i \in \mathbb{N} \).

2. A solution \( p \) of the operator equation (2.30) with \( \gamma = 0 \) and \( d_m = 2 \) must satisfy
\[
\frac{\sigma_X^2}{2\beta^3} e^{-(i_1+1)\beta\tau} (e^{\beta\tau} - 1)^2 = m_1 \tag{3.4}
\]
\[
\frac{\sigma_X^2}{2\beta^3} e^{-(i_2+1)\beta\tau} (e^{\beta\tau} - 1)^2 = m_2. \tag{3.5}
\]
For \( m_1 = 0 \) or \( m_2 = 0 \) we get \( p_2 = \sigma_X = 0 \), which implies \( m \equiv 0 \); \( p_1 = \beta \) is not determined. Therefore \( F_2^{-1} (\{0\}) = \{ (\beta, 0) : \beta > 0 \} \). For \( m_1 \neq 0 \) we divide Equation (3.5) by Equation (3.4) and get
\[
e^{-(i_2-i_1)\beta\tau} = \frac{m_2}{m_1}, \tag{3.6}
\]
which has for $0 < m_2 < m_1$ the unique solution $\beta = -\frac{1}{(i_2 - i_1)^\tau} \ln \left( \frac{m_2}{m_1} \right) > 0$. Substituting this into Equation (3.4) and solving for $\sigma_X$ we get (under the restriction $\sigma_X \geq 0$) the unique solution $\sigma_X = \frac{\sqrt{2^\beta m_1}}{\sqrt{e^{i_1 + 1}}} \sqrt{e \sigma^2}$. For $m_2 \geq m_1$ Equation (3.6) has no positive solution for $\beta$. Thus, in this case the system of equations (3.4) and (3.5) has no solution $p = \left( \frac{\beta}{\sigma_X} \right) \in D(F_2)$.  

The above results can be summarized in the following way:

**Corollary 3.7**

The operator $F$ in (2.29) is always continuous and injective if

$$d_m \geq \begin{cases} 1 & \lambda = 0 \\ 2 & \gamma = 0 \end{cases}. \quad (3.7)$$

Thus, in this case the inverse operator $F^{-1}$ defined on the range $F(D(F))$ exists.

Now we address the question whether the operator $F : D(F) \subset B_1 \rightarrow B_2$ is compact or not, i.e. whether for each bounded sequence $\{(a_n, p_n)\}_{n \in \mathbb{N}}$ there exists a subsequence $\{(a_{n_k}, p_{n_k})\}$ and an element $(u_0, m_0) \in B_2$ such that

$$F\left((a_{n_k}, p_{n_k})\right) \rightarrow (u_0, m_0) \text{ in } B_2.$$  

The answer is given in the following theorem.

**Theorem 3.8**

In the case (i) ($\lambda = 0$) the operator $F = \hat{F} : D(\hat{F}) \subset B_1 \rightarrow B_2$ is compact. In the case (ii) ($\gamma = 0$) the operator $\tilde{F} : D(\tilde{F}) \subset B_1 \rightarrow B_2$ is not compact.

**Proof:** First we consider the case (i) ($\lambda = 0$). Let $\{(a_n, p_n)\}_{n \in \mathbb{N}}$ be a bounded sequence in $B_1$, i.e. $\exists C$ such that $\| (a_n, p_n) \|_{B_1} \leq C$. This implies $\| a_n \|_{L^2(I)} \leq C$ and $\| p_n \|_{\mathbb{R}^1} \leq C$. Since $F_1 : D(F_1) \subset L^2(I) \rightarrow \mathbb{L}^2(I)$ is compact, there exists a subsequence of $\{a_n\}_{n \in \mathbb{N}}$ and an element $u_0$ such that

$$F_1(a_{n_k}) \rightarrow u_0 \text{ in } L^2(I).$$

From Corollary 3.3 we get

$$\left\| \hat{F}_2\left(a_{n_k}, p_{n_k}\right) \right\|_{X_m} \leq \sqrt{d_m} \max_{1 \leq j \leq d_m} \left( \frac{1}{b_{i_j} - b_{i_{j-1}}} \int_{b_{i_{j-1}}}^{b_{i_j}} a_{n_k}(t) \, dt \right)^\tau \leq \max_{1 \leq j \leq d_m} \frac{\sqrt{d_m}}{\sqrt{b_{i_j} - b_{i_{j-1}}}} \| a_{n_k} \|_{L^2(I)}^\tau.$$
Since in the finite dimensional space $X_m$ each bounded sequence contains a convergent subsequence, we get a subsequence $\{n_{k'}\} \subset \{n_k\}$ and an element $m_0 \in X_m$ such that

$$F_1 \left( a_{n_{k'}} \right) \xrightarrow{L^2(I)} u_0, \quad \tilde{F}_2 \left( \left( a_{n_{k'}}, p_{n_{k'}} \right) \right) \xrightarrow{X_m} m_0 \quad k' \to \infty.$$ 

By the definition of the Hilbert space $B_2$ this implies

$$\tilde{F} \left( \left( a_{n_{k'}}, p_{n_{k'}} \right) \right) \to (u_0, m_0) \quad k' \to \infty.$$ 

Now we turn to the case (ii) ($\gamma = 0$). We have to show that there exists a bounded sequence $\{ (a_n, p_n) \} \subset B_1$ for which there exists no subsequence $\left( a_{n_k}, p_{n_k} \right)$ such that

$$\tilde{F} \left( \left( a_{n_k}, p_{n_k} \right) \right)$$

is convergent in $B_2$. We can take an arbitrary sequence $\{ a_n \} \subset L^2(I)$ and the sequence $\{ p_n = \left( \begin{array}{c} 1 \\ \frac{1}{n} \end{array} \right) \}_{n=1}^\infty$. From the limit $\lim_{\beta \to 0} \tilde{h}_j ((\beta, \sigma_X)) = \infty$ (c.f. Lemma 3.5) we get

$$\left\| \tilde{F}_2 \left( (a_n, p_n) \right) \right\|_{X_m} \to \infty.$$ 

Therefore, we also have

$$\left\| \tilde{F} \left( (a_n, p_n) \right) \right\|_{B_1} \to \infty. \quad (3.8)$$

Since every convergent sequence has to be bounded, the sequence $\left\{ \tilde{F} \left( (a_n, p_n) \right) \right\}_{n=1}^\infty$ cannot contain any convergent subsequence.

Now we consider the question whether the operator $F$ is weakly continuous or at least weakly closed.

**Theorem 3.9**

The operator $F$ in (2.29) is weakly closed, i.e. for each in $B_1$ weakly convergent sequence $(a_n, p_n) \overset{B_1}{\to} (a, p)$ with $(a_n, p_n) \in D(F)$ and $F \left( (a_n, p_n) \right) \overset{B_2}{\to} (u, m)$, we have $(a, p) \in D(F)$ and $F \left( (a, p) \right) = (u, m)$. In the case (i) ($\lambda = 0$) the operator $F = \tilde{F}$ is even weakly continuous.

**Proof:** In the case (i) ($\lambda = 0$) the weak continuity of the operator $F = \tilde{F}$ follows from the weak continuity of the operators $F_1$ and $\tilde{F}_2$. Therefore we only need to prove the weak closedness of $F = \tilde{F}$ in the case (ii) ($\gamma = 0$). Let $(a_n, p_n) \in D(F)$, $(a_n, p_n) \overset{B_1}{\to} (a, p)$ be a weakly convergent sequence and $\tilde{F} \left( (a_n, p_n) \right) \overset{B_2}{\to} (u, m)$. This implies $a_n \in D(F_1)$, $(a_n, p_n) \in D(\tilde{F}_2)$ and

$$a_n \xrightarrow{L^2(I)} a, \quad F_1(a_n) \xrightarrow{L^2(I)} u \quad \text{as well as} \quad p_n \xrightarrow{X_p} p, \quad \tilde{F}_2 \left( a_n, p_n \right) \xrightarrow{X_m} m.$$ 

Since $F_1 : L^2_+(I) \to L^2(I)$ is weakly closed, we get $a \in D(F_1) = L^2_+(I)$ and $F_1(a) = u$. 

Writing \( p := \left( \frac{\beta}{\sigma_X} \right) \) it remains to show \( \beta > 0, \sigma_X \geq 0 \) and \( \left( \tilde{h}_j ((\beta, \sigma_X)) \right)^{d_m}_{j=1} = m \), where \( \tilde{h}_j \) was defined in (2.24). Because of \( (a_n, p_n) \in D(\tilde{F}_2) \) we have \( (p_n)_{1} > 0 \) and \( (p_n)_{2} \geq 0 \). Thus, \( p_n \to p = \left( \frac{\beta}{\sigma_X} \right) \) implies \( \beta \geq 0 \) and \( \sigma_X \geq 0 \).

Assume \( \beta = 0, \) i.e. \( p_n \xrightarrow{X_n} \left( \frac{0}{\sigma_X} \right) \). Because of \( \lim_{\beta \to 0} \tilde{h}_j ((\beta, \sigma_X)) = \infty \) (c.f. Lemma 3.5) we have then \( \left\| \tilde{F}_2 (a_n, p_n) \right\|_{X_n} \to \infty \), which contradicts \( \tilde{F}_2 (a_n, p_n) \to m \). Hence we have \( \beta > 0 \) and thus \( (a, p) \in D(\tilde{F}_2) \).

Finally, from the continuity of \( \tilde{h}_j \) on the set \( \left\{ \left( \frac{p_1}{p_2} \right) : p_1 > 0, p_2 \geq 0 \right\} \) we get

\[
\tilde{h}_j (p_n) \to \tilde{h}_j (p) \quad \text{and therefore also} \quad \tilde{F}_2 (a_n, p_n) \to \tilde{F}_2 (a_0, m_0),
\]

which completes the proof.

Now we are in search of solution points \( (a^*, p^*) \in D(F) \), for which the inverse problem (SIP) written as an operator equation (2.29) is locally ill-posed or locally well-posed in the sense of Definition 2 in [5]), which we repeat in the following.

**Definition 3.10** We call the operator equation (2.28) between the Hilbert spaces \( B_1 \) and \( B_2 \) locally ill-posed at the point \( (a^*, p^*) \in D(F) \), if for all balls \( B_r \) \( ((a^*, p^*)) \) with radius \( r > 0 \) and center \( (a^*, p^*) \) there exist sequences \( \{ (a_n, p_n) \}_{n=1}^{\infty} \subset B_r \) \( ((a^*, p^*)) \cap D(F) \) satisfying the condition

\[
F \left( (a_n, p_n) \right) \to F \left( (a^*, p^*) \right) \quad \text{in} \ B_2, \ \text{but} \quad (a_n, p_n) \not\to (a^*, p^*) \quad \text{in} \ B_1 \quad \text{as} \ n \to \infty.
\]

Otherwise, we call the equation (2.28) locally well-posed at \( (a^*, p^*) \in D(F) \). That means that there exists a radius \( r > 0 \) such that

\[
F \left( (a_n, p_n) \right) \to F \left( (a^*, p^*) \right) \quad \text{in} \ B_2 \implies (a_n, p_n) \to (a^*, p^*) \quad \text{in} \ B_1 \quad \text{as} \ n \to \infty,
\]

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

Corollary 3.7 stated that the operator \( F \) is injective if and only if condition (3.7) is fulfilled. In this case local ill-posedness at \( (a^*, p^*) \) implies discontinuity of \( F^{-1} \) at the point \( (u^*, m^*) = F \left( (a^*, p^*) \right) \) and vice versa continuity of \( F^{-1} \) at \( (u^*, m^*) \) implies local well-posedness at \( (a^*, p^*) \).

The next proposition is a simple consequence of Corollary 5.2 in [3], which states that the operator equation (3.1) is locally ill-posed at every point \( a^* \in D(F_1) \), whenever we have

\[
\text{ess inf}_{t \in I} a^*(t) > 0.
\]

(3.9)
Proposition 3.11
The operator equation (2.28) associated with the inverse problem (SIP) is locally ill-posed at the solution point \((a^*, p^*) \in D(F)\) whenever \(a^*\) satisfies (3.9).

According to Proposition 3.11 the inverse problem (SIP) is in general ill-posed, this means that even if we had a sequence \(\{(u_n, m_n)\}_{n=1}^\infty\) converging to the data \((u^*, m^*)\) the pair \((a_n, p_n) = F^{-1}(u_n, m_n)\) need not converge to the solution of (2.28). This ill-posedness results from the ill-posedness of the operator equation (3.1), i.e. the identification of the volatility term structure \(a(t) \ (t \in I)\) is the ill-posed part of (SIP).

4 Applicability of Tikhonov regularization

As the operator equation (2.28) is in general ill-posed a regularization approach is required. Here we will use the standard Tikhonov regularization for nonlinear equations. For data \((u^\delta, m^\delta)\) with
\[
\| (u^\delta, m^\delta) - (u^*, m^*) \|_{B_2} \leq \delta
\]
and a fixed initial guess \((\bar{a}, \bar{p}) \in B_1\) we will use regularized solutions \((a^\alpha^\delta, p^\alpha^\delta)\) as minimizers of the Tikhonov functional, i.e. we solve the extremal problem
\[
\| F((a, p)) - (u^\delta, m^\delta) \|_{B_2}^2 + \alpha \| (a, p) - (\bar{a}, \bar{p}) \|_{B_1}^2 \rightarrow \min, \quad \text{subject to} \quad (a, p) \in D(F),
\]
for regularization parameters \(\alpha > 0\).

Since the operator \(F\) is continuous and weakly closed we can apply the theory of [1] and obtain the following two results concerning continuous dependence of the regularized solutions \((a^\alpha^\delta, p^\alpha^\delta)\) on the data \((u^\delta, m^\delta)\) and convergence of the regularized solutions \((a^\delta, p^\delta)\) to the unique solution \((a^*, p^*)\) of operator equation (2.28). Note that regularized solutions exist for all \(\alpha > 0\).

**Proposition 4.1**

Let \(\alpha > 0\) and let \(\{ (u_k, m_k) \}\) and \((a^\alpha^k, p^\alpha^k)\) be sequences with
\[
(u_k, m_k) \overset{B_2}{\rightarrow} (u^\delta, m^\delta).
\]
and \((a^\alpha^k, p^\alpha^k)\) is a minimizer of (4.1) with \((u^\delta, m^\delta)\) replaced by \((u_k, m_k)\). Then there exists a convergent subsequence of \(\{(a^\alpha^k, p^\alpha^k)\}\) and the limit of every convergent subsequence is a minimizer of (4.1).

**Proposition 4.2**

Let \((u^\delta, m^\delta) \in B_2\) with \(\| (u^\delta, m^\delta) - (u^*, m^*) \|_{B_2} \leq \delta\) be given and let \(\alpha(\delta)\) be such that
\( \alpha(\delta) \to 0 \) and \( \delta^2 \alpha(\delta) \to 0 \) as \( \delta \to 0 \). Then the regularized solutions converge towards the unique solution \((a^*, p^*)\) of operator equation (2.28), i.e.

\[
\lim_{\delta \to 0} \left( a_{\alpha(\delta)}^\delta, p_{\alpha(\delta)}^\delta \right) = (a^*, p^*) .
\]

The convergence formulated in Proposition 4.2 may be arbitrarily slow. For obtaining convergence rates, we restrict the domain of \( F \) in the form

\[
\tilde{D}(F) := \{ (a, p) \in B_1 : a(t) \geq \xi > 0 \text{ a.e. in } I \} .
\]

Using a well known modification of Theorem 2.4 in [2] we have the following proposition.

**Proposition 4.3** Under the conditions stated above we obtain for the parameter choice \( \alpha \sim \delta \) the convergence rate

\[
\left\| \left( a_{\alpha(\delta)}^\delta, p_{\alpha(\delta)}^\delta \right) - (a^*, p^*) \right\| = \mathcal{O} \left( \sqrt{\delta} \right)
\]

if there exists a continuous linear operator \( G : L^2(I) \times X_m \subset B_1 \to B_2 \) with adjoint \( G^* \) and a positive constant \( L \) such that

i) \( \left\| F \left( (a, p) \right) - F \left( (a^*, p^*) \right) - G \left( (a, p) - (a^*, p^*) \right) \right\|_{B_2} \leq \frac{1}{2} \left\| (a, p) - (a^*, p^*) \right\|_{B_1}^2 \)

holds for all \( (a, p) \in \tilde{D}(F) \cap U_\rho \left( (a^*, p^*) \right) \) for some \( \rho > 2 \left\| (a^*, p^*) - (\bar{a}, \bar{p}) \right\|_{B_1} \)

ii) there exists an element \( (w, r) \in B_2 \) satisfying \( (a^*, p^*) - (\bar{a}, \bar{p}) = G^*(w, r) \) and

iii) \( L \left\| (w, r) \right\|_{B_1} < 1 \).

In the rest of this section we are going to analyze the conditions i)-ii) for the specific inverse problem (SIP). Here we will concentrate mainly on the special case (i) \( (\lambda = 0) \).

Writing

\[
G \left( (v, q) \right) = \left( G_1 \left( (v, q) \right), G_2 \left( (v, q) \right) \right)
\]

(4.2)

we will reformulate condition i) in form of conditions for \( G_1 \) and \( G_2 \). Excluding the singular situation of at-the-money options, where the initial asset price \( P_0 \) is equal to the strike price \( K \), we will give an operator \( G \) and a constant \( L \) such that i) is fulfilled. After that we are going to interpret the conditions ii) and iii) from Proposition 4.3. As we need therefore the adjoint operator \( G^* \), part (2) is devoted to the derivation of \( G^* \). Using this result we can analyze the source condition ii) in part (3). In part (4) we will consider some specific situations and analyze for these specific situations necessary conditions imposed by ii). We conclude this section by a remark concerning the special case (ii) \( (\gamma = 0) \).
Using the decomposition (4.2) of the operator $G$ and applying the definition of the norms in $B_1$ and $B_2$ we can write condition 4.3 as

\[
\|F_1(a) - F_1(a^*) - G_1((a - a^*, p - p^*))\|_{L^2(I)}^2
\]

\[
+ \|F_2((a, p)) - F_2((a^*, p^*)) - G_2((a - a^*, p - p^*))\|_{X_m}^2
\]

\[
\leq \left( \frac{L}{2} \| (a, p) - (a^*, p^*) \|_{B_1}^2 \right)^2 \quad \forall (a, p) \in \hat{D}(F). \quad (4.3)
\]

Setting $p = p^*$ and $L_1 = L$ this implies

\[
\|F_1(a) - F_1(a^*) - G_1((a - a^*, 0))\|_{L^2(I)} \leq \frac{L_1}{2} \| a - a^* \|^2_{L^2(I)}. \quad (4.4)
\]

Defining an integral operator $J : L^2(I) \rightarrow L^2(I)$ as

\[
[J(v)](t) := \int_0^t v(\tau) \, d\tau \quad (t \in I)
\]

and a multiplier function

\[
m(0) = 0, \quad m(t) = \frac{\partial U_{BS}(X, K, r, t, J(a^*))(t)}{\partial s} > 0 \quad (0 < t \leq T)
\]

we introduce the linear operator $G_1 : B_1 \rightarrow L^2(I)$ by

\[
[G_1(v, q)](t) = m(t)[J(v)](t) \quad (t \in I, \ v \in L^2(I), \ q \in X_p).
\]

The next proposition is a direct consequence of Theorem 5.4 in [3].

**Proposition 4.4**

If $P_0 \neq K$ the operator $G_1 : B_1 \rightarrow L^2(I)$ is a continuous linear operator and it satisfies (4.4) with a constant $L_1 = TC$, where

\[
C := \sup_{(t, \bar{S}) \in \mathcal{M}_c} \left| \frac{\partial^2 U_{BS}(X, K, r, t, \bar{S})}{\partial \bar{S}^2} \right| < \infty
\]

is defined from the set

\[
\mathcal{M}_c := \left\{ (t, \bar{S}) \in \mathbb{R}^2 : \bar{S} \geq c_t, 0 < t \leq T \right\}.
\]

We are now in search of an operator $G_2$ satisfying (4.3) and therefore

\[
\|F_2((a, p)) - F_2((a^*, p^*)) - G_2((a - a^*, p - p^*))\|_{X_m}
\]

\[
\leq \frac{L_2}{2} \| (a - a^*, p - p^*) \|_{B_1}^2. \quad (4.5)
\]
Here, we will concentrate our considerations on the special case (i) ($\lambda = 0$). Analyzing the $j$-th component of $F(a, p)$ we get by Taylor Expansion

$$
(F_2 ((a, p)))_j - F_2 ((a^*, p^*))_j = \frac{\partial \hat{h}(c, \gamma)}{\partial c} \bigg|_{c^j_\gamma} \left( \frac{1}{b_{ij} - b_{ij-1}} \right) \xi_{ij}, a - a^* - \frac{\partial \hat{h}(c, \gamma)}{\partial \gamma} \bigg|_{c^j_\gamma} (p_1 - p_1^*) + \frac{1}{2} \left( \frac{1}{b_{ij} - b_{ij-1}} \right) \xi_{ij}, a - a^* \right) (p_1 - p_1^*)^2,
$$

where $c_j^* = \left( \frac{1}{b_{ij} - b_{ij-1}} \xi_{ij}, a^* \right)$, $c_j^\theta = \left( \frac{1}{b_{ij} - b_{ij-1}} \xi_{ij}, a^* + \theta (a - a^*) \right)$ and $p_1^\theta = p_1^* + \theta (p_1 - p_1^*)$ for some $\theta \in (0, 1)$.

Using the auxiliary functions $f_n(y)$ defined in (3.2) and Lemma 3.2 we can now write the function $\hat{h}$ and its partial derivatives as follows

$$
\hat{h}(c, \gamma) = c \tau f_0(\gamma \tau) \quad (4.7a)
$$

$$
\frac{\partial \hat{h}(c, \gamma)}{\partial c} = \tau f_0(\gamma \tau) \quad \in (0, \tau) \quad (4.7b)
$$

$$
\frac{\partial \hat{h}(c, \gamma)}{\partial \gamma} = c \tau^2 f_1(\gamma \tau) \quad \in \left(- \frac{\tau^2 c}{2}, 0\right) \quad (4.7c)
$$

$$
\frac{\partial^2 \hat{h}(c, \gamma)}{\partial c^2} = 0 \quad (4.7d)
$$

$$
\frac{\partial^2 \hat{h}(c, \gamma)}{\partial c \partial \gamma} = \tau^2 f_1(\gamma \tau) \quad \in \left(- \frac{\tau^2}{2}, 0\right) \quad (4.7e)
$$

$$
\frac{\partial^2 \hat{h}(c, \gamma)}{\partial \gamma^2} = c \tau^3 f_2(\gamma \tau) \quad \in \left(0, \frac{\tau^3}{3} c\right) \quad (4.7f)
$$

Thus, we set

$$
\dot{G}_2 ((v, q)) := \frac{\partial \hat{h}(c, \gamma)}{\partial c} \bigg|_{c^j_\gamma} \left( \frac{1}{b_{ij} - b_{ij-1}} \xi_{ij}, v \right) + \frac{\partial \hat{h}(c, \gamma)}{\partial \gamma} \bigg|_{c^j_\gamma} (q) = \left( \tau f_0(p_1^* \tau) \left( \frac{1}{b_{ij} - b_{ij-1}} \xi_{ij}, v \right) + \left( \frac{1}{b_{ij} - b_{ij-1}} \xi_{ij}, a^* \right) \tau^2 f_1(p_1^* \tau) \right) q \right|_{\gamma = 0}^{\gamma = \tau}.
$$

The relations (4.7b) and (4.7c) show that $G_2$ is a bounded and hence continuous linear operator mapping from $B_1$ into $X_m$. Furthermore from (4.6) and (4.7) we
get the estimation
\[ \left\| F_2((a, p)) - F_2((a^*, p^*)) - \hat{G}_2((a - a^*, p - p^*)) \right\|_{X_m} \leq \sqrt{\sum_{j=1}^{d_m} \xi_j \left( \tau^2 \|a - a^*\| \|p_1 - p_1^*\| + \frac{1}{2} \frac{\tau^3 (\|a^*\| + \rho) \|p_1 - p_1^*\|^2}{\sqrt{b_i - b_{i-1}}} \right)^2} \]
\[ \leq \max_{1 \leq j \leq d_m} \frac{\sqrt{\xi_j}}{2\sqrt{b_i - b_{i-1}}} \left( \tau^2 + \frac{\tau^3}{3} \left( \|a^*, p^*\|_{B_1} + \rho \right) \|a, p\| - \|a^*, p^*\|_{B_1} \right) \]
\[ \leq \frac{L_2}{2} \| (a, p) - (a^*, p^*) \|^2_{B_1} \]

with \( L_2 \) defined as
\[ L_2 := \sqrt{d_m} \frac{1}{2} \left( \tau^2 + \frac{\tau^3}{3} \left( \|a^*, p^*\|_{B_1} + \rho \right) \right) \max_{1 \leq j \leq d_m} \frac{\sqrt{\xi_j}}{\sqrt{b_i - b_{i-1}}} . \]

Setting
\[ G(v, q) = \left( G_1(v, q), \hat{G}_2(v, q) \right) \quad (v \in L^2(I), q \in X_p) \]

(4.4) and (4.8) imply that (4.3) is fulfilled with \( L := L_1 + L_2 \).

(2) Now we are going to study the adjoint operator \( G^* : B_2 \to B_1 \) of \( G : B_1 \to B_2 \). Defining the operator \( \hat{G}_1 : L^2(I) \to L^2(I) \) as
\[ \hat{G}_1(v) = G_1(v, 0) \]

we can write the general condition for an adjoint operator as
\[ \langle G^*(w, \ell), (v, q) \rangle_{B_1} = \langle (w, \ell), G(v, q) \rangle_{B_2} = \langle w, \hat{G}_1v \rangle_{L^2(I)} + \langle \ell, \hat{G}_2(v, q) \rangle_{X_m} \quad \forall (w, \ell) \in B_2, \forall (v, q) \in B_1. \quad (4.9) \]

Setting \( q = 0 \) (4.9) implies
\[ \langle G^*(w, \ell), (v, 0) \rangle_{B_1} = \langle w, \hat{G}_1v \rangle_{L^2(I)} + \langle \ell, \hat{G}_2(v, 0) \rangle_{X_m} = \langle w, \hat{G}_1v \rangle_{L^2(I)} + \langle \ell, \hat{G}_2v \rangle_{X_m} \]
\[ = \langle G_1^* w, v \rangle_{L^2(I)} + \langle \hat{G}_2^* \ell, v \rangle_{L^2(I)} \quad \forall (w, \ell) \in B_2, \forall v \in L^2(I), \]

where \( \hat{G}_2^* : L^2(I) \to X_m \) is defined by
\[ \hat{G}_2^* v := \hat{G}_2(v, 0) = \left( \tau f_0(p_1^*) \frac{1}{b_i - b_{i-1}} \langle \chi_{i_j}, v \rangle \right)^{d_m}_{j=1} \quad (v \in L^2(I)) \]
and the adjoint $\hat{G}_2^* : X_m \to L^2(I)$ has the form
\[
[\hat{G}_2^* \mathcal{E}](t) = \tau f_0 (p_1^* \tau) \sum_{j=1}^{d_m} \frac{\xi_j}{b_{ij} - b_{ij-1}} (\mathcal{E})_j x_i(t) \quad (\mathcal{E} \in X_m).
\]

Analogously we get from (4.9)
\[
\langle G^*(w, \mathcal{E}), (0, q) \rangle_{B_1} = \langle w, G_1(0) \rangle_{L^2(I)} + \langle \mathcal{E}, G_2(0, q) \rangle_{X_m}
\]
\[
= 0 + \langle \mathcal{E}, \hat{G}_2(0, q) \rangle_{X_m} = \langle \hat{G}_2^* \mathcal{E}, q \rangle_{X_p},
\]

where the operator $\hat{G}_2 : X_p \to X_m$ is defined by
\[
\hat{G}_2 q := \hat{G}_2(0, q) = (c_j^* \tau^2 f_1(p_1^* \tau)) \sum_{j=1}^{d_m} q_j (\mathcal{E})_j \xi_j \quad (q \in X_p)
\]
and the adjoint is given by
\[
\tilde{\hat{G}}_2^* \mathcal{E} = \sum_{j=1}^{d_m} c_j^* \xi_j \tau^2 f_1(p_1^* \tau) (\mathcal{E})_j = \tau^2 f_1(p_1^* \tau) \sum_{j=1}^{d_m} c_j^* (\mathcal{E})_j \xi_j \quad (\mathcal{E} \in X_m)
\]
with $c_j^* = \left( \frac{1}{b_{ij} - b_{ij-1}} \right)_{X_m} x_i, a^*$. Hence, for $(v, q) \in B_1$ we can write condition (4.9) as
\[
\langle G^*(w, \mathcal{E}), (v, q) \rangle_{B_1} = \langle G^*(w, \mathcal{E}), (v, 0) \rangle_{B_1} + \langle G^*(w, \mathcal{E}), (0, q) \rangle_{B_1}
\]
\[
= \langle \hat{G}_1^* w, v \rangle_{L^2(I)} + \langle \hat{G}_2^* \mathcal{E}, v \rangle_{L^2(I)} + \langle \tilde{\hat{G}}_2^* \mathcal{E}, q \rangle_{X_p}. \tag{4.10}
\]

Since Equation (4.10) holds for all $(v, q) \in B_1$ we have
\[
G^*(w, \mathcal{E}) = \begin{pmatrix}
\hat{G}_1^* w \\
\hat{G}_2^* \mathcal{E} \\
\tilde{\hat{G}}_2^* \mathcal{E}
\end{pmatrix} _{\mathcal{E} \in L^2(I)} _{\mathcal{E} \in X_p}.
\]

(3) Now we can write the source condition 4.3 as
\[
(a^*, p^*) - (\bar{a}, \bar{p}) = G^*(w, \mathcal{E}) = \begin{pmatrix}
\hat{G}_1^* w + \hat{G}_2^* \mathcal{E} \\
\tilde{\hat{G}}_2^* \mathcal{E}
\end{pmatrix},
\]
or equivalently as
\[
a^* - \bar{a} = \hat{G}_1^* w + \hat{G}_2^* \mathcal{E}, \tag{4.11a}
\]
\[
p^* - \bar{p} = \tilde{\hat{G}}_2^* \mathcal{E}. \tag{4.11b}
\]
Equation (4.11a) is equivalent to
\[
(a^* - \overline{a})(t) = \int_t^T m(y) w(y) \, dy + \tau f_0(p_1^* \tau) \sum_{j=1}^{d_m} \frac{\xi_j}{b_{ij} - b_{ij-1}} (r_j \chi_{ij})(t), \tag{4.12}
\]
which implies
\[
\frac{(a^* - \overline{a} - \tau f_0(p_1^* \tau) \sum_{j=1}^{d_m} \frac{\xi_j}{b_{ij} - b_{ij-1}} (r_j \chi_{ij})') \, dy}{m} \in L^2(I)
\]
and therefore also
\[
a^* - \overline{a} - \tau f_0(p_1^* \tau) \sum_{j=1}^{d_m} \frac{\xi_j}{b_{ij} - b_{ij-1}} (r_j \chi_{ij}) \in H^1(I) \subset C(I). \tag{4.13}
\]
This means that the difference \(a^* - \overline{a}\) has to be continuous in the intervals \((b_{ij-1}, b_{ij})\)
and may have jumps at the points \(b_{ij}\).

From (4.12) we get furthermore
\[
\begin{align*}
\left\{ (a^* - \overline{a})(T) = \tau f_0(p_1^* \tau) \frac{\xi_{d_m}}{b_{i_{d_m}} - b_{i_{d_m}-1}} (r_{d_m}) & \quad i_{d_m} = n \\
(a^* - \overline{a})(T) = 0 & \quad i_{d_m} < n, 
\end{align*}
\tag{4.14}
\]
i.e. we find the condition
\[
\begin{align*}
(a^*(T) = \overline{a}(T) & \quad i_{d_m} < n \\
(r)_{d_m} = \frac{a^*(T) - \overline{a}(T)}{\tau \xi_{d_m} f_0(p_1^* \tau)} (b_{i_{d_m}} - b_{i_{d_m}-1}) & \quad i_{d_m} = n.
\end{align*}
\tag{4.15}
\]

The case \(i_{d_m} < n\) occurs if we do not expect \(a^*(t)\) to be constant on the last interval and therefore cannot use the logarithmic returns on that interval. Then (4.15) implies that we have to know the exact solution at the point \(T\) in advance, which seems to be not realistic.

The case \(i_{d_m} = n\) occurs if we expect \(a^*(t)\) to be constant on the last interval and therefore we can use the logarithmic returns on that interval. Then (4.15) determines the last component of the vector \(r\).

We derive furthermore from (4.13) that the function
\[
a^* - \overline{a} - \tau f_0(p_1^* \tau) \sum_{j=1}^{d_m} \frac{\xi_j}{b_{ij} - b_{ij-1}} (r_j \chi_{ij}) \in C(\mathbb{R}) \tag{4.16}
\]
has to be continuous at the points \(b_k\) \((k = 1, \ldots, n - 1)\). This gives the \(n - 1\) equations
\[
\lim_{t \to b_k^+} \{a^*(t) - \overline{a}(t)\} - \lim_{t \to b_k^-} \{a^*(t) - \overline{a}(t)\} = \tau f_0(p_1^* \tau) \left( \lim_{t \to b_k^+} \sum_{j=1}^{d_m} \frac{\xi_j}{b_{ij} - b_{ij-1}} (r_j \chi_{ij})(t) \right)
\]
\[
- \lim_{t \to b_k^-} \sum_{j=1}^{d_m} \frac{\xi_j}{b_{ij} - b_{ij-1}} (r_j \chi_{ij})(t) \right) \tag{4.17}
\]
\((k = 1, \ldots, n - 1)\).
Clearly, these equations are conditions on the vector $\mathbf{r}$ and the jump heights of the function $a^* - \overline{a}$. The right hand sides of (4.17) simplify if some of the intervals $I_k$ and $I_{k+1}$ are not used, i.e. if $k \notin J_0$ or $k + 1 \notin J_0$. The following situations can occur.

<table>
<thead>
<tr>
<th>$k = i_j \in J_0$</th>
<th>$k + 1 = i_{j+1} \in J_0$</th>
<th>right hand side of Equation (4.17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = i_j \in J_0$</td>
<td>$k + 1 \notin J_0$</td>
<td>$\tau f_0 (p^*<em>j \tau_j) \left( \frac{\xi</em>{j+1}}{b_{k+1} - b_k} (r)<em>{j+1} - \frac{\xi_j}{b_k - b</em>{k-1}} (r)_j \right)$</td>
</tr>
<tr>
<td>$k \notin J_0$</td>
<td>$k + 1 = i_{j+1} \in J_0$</td>
<td>$-\tau f_0 (p^*<em>j \tau_j) \frac{\xi_j}{b_k - b</em>{k-1}} (r)_j$</td>
</tr>
<tr>
<td>$k \notin J_0$</td>
<td>$k + 1 \notin J_0$</td>
<td>$\tau f_0 (p^*<em>j \tau_j) \frac{\xi</em>{j+1}}{b_{k+1} - b_k} (r)_{j+1}$</td>
</tr>
</tbody>
</table>

(4) Finally we will consider some specific situations. Let us first assume that $J_0 = \{1, \ldots, n\}$, i.e. we assume $a^*$ to be piecewise constant and use the logarithmic returns from all intervals $I_k$. In this situation there exists for each $a^*$ and $\overline{a}$ a uniquely defined vector $\mathbf{r}$ which satisfies the equations (4.17) and (4.15).

Now we consider an interval $I_k$ such that $k = i_j \in J_0$ but $k - 1 \notin J_0$ and $k + 1 \notin J_0$. Then the equations (4.17) at the points $b_{k-1}$ and $b_k$ require

$$
\lim_{t \to b_k^-} \{a^*(t) - \overline{a}(t)\} - \lim_{t \to b_k^+} \{a^*(t) - \overline{a}(t)\} = \tau f_0 (p^*_j \tau_j) \frac{\xi_j}{b_k - b_{k-1}} (r)_j
$$

and

$$
\lim_{t \to b_k^-} \{a^*(t) - \overline{a}(t)\} - \lim_{t \to b_k^+} \{a^*(t) - \overline{a}(t)\} = -\tau f_0 (p^*_j \tau_j) \frac{\xi_j}{b_k - b_{k-1}} (r)_j
$$

and therefore

$$
\lim_{t \to b_{k-1}^+} \{(a^* - \overline{a}) (t)\} - \lim_{t \to b_{k-1}^-} \{(a^* - \overline{a}) (t)\} = -\lim_{t \to b_k^-} \{(a^* - \overline{a}) (t)\} + \lim_{t \to b_k^+} \{(a^* - \overline{a}) (t)\},
$$

which can be viewed as a condition on the jump heights of $a^* - \overline{a}$.

Furthermore we see the following. If $a^* - \overline{a}$ is continuous then the left hand side of Equation (4.17) is zero for all $k = 1, \ldots, n - 1$. If there is an interval $k \notin J_0$, then either $k + 1 \notin J_0$ or $k + 1 = i_{j+1} \in J_0$. In the last case equation (4.17) requires $(r)_{j+1} = 0$. If we proceed in this way we can show $(r)_l = 0$ for all $l \leq j + 1$ and analogously also $(\overline{r})_l = 0$ for all $l \leq j + 1$ and therefore $\mathbf{r} = 0$. Particularly, (4.14) gives $a^*(T) = \overline{a}(T)$.

We are now going to study the source condition for functions $a^*$ and $\overline{a}$ which are piecewise constant on the intervals $[b_{j-1}, b_j]$ ($j = 1, \ldots, n$), i.e.

$$
a^*(t) \equiv c_j \quad \overline{a}(t) \equiv \overline{c}_j \quad j = 1, \ldots, n.
$$

However, we do not assume that we knew in advance that the volatility is constant on all intervals $I_i$, i.e. we do not impose the restriction $J_0 = \{1, 2, \ldots, n\}$. 

As \( a^* \) and \( \overline{a} \) are constant on each interval \( I_i \) the function (4.16) is a step function. Hence, the continuity of this function implies that it is constant on the whole interval \( I \) and because of (4.14) equal to zero, i.e.

\[
(a^* - \overline{a}) (t) = \tau f_0(p_1^* \tau) \sum_{j=1}^{d_m} \frac{\xi_j}{b_{i_j} - b_{i_j-1}} (\overline{r})_j \chi_{i_j}(t). \tag{4.18}
\]

Therefore the vector \( \overline{r} \) is uniquely determined and it holds

\[
(\overline{r})_j = p_1^* \frac{c_j - \overline{c}_j}{1 - e^{-p_1^* \tau}} \frac{b_{i_j} - b_{i_j-1}}{\xi_j} \quad (j = 1, \ldots, d_m).
\]

Furthermore (4.18) implies \( c_k = \overline{c}_k \) whenever \( k \notin J_0 \). Hence, the equation

\[
a^*(t) = \overline{a}(t) \quad (t \in I_k, k \notin J_0) \tag{4.19}
\]

is required.

If (4.19) is satisfied, then (4.11a) is also fulfilled with \( w \equiv 0 \) and \( \overline{r} \). Condition (4.11b) gives

\[
p_1^* - \overline{p}_1 = -\frac{1}{p_1^* (1 - e^{-p_1^* \tau})} \sum_{j=1}^{d_m} c_j (c_j - \overline{c}_j) (b_{i_j} - b_{i_j-1}),
\]

which can be interpreted as a condition for \( \overline{p} \). The smallness condition 4.3 implies then \( L \| \overline{r} \| < 1 \) or equivalently

\[
\frac{p_1^*}{1 - e^{-p_1^* \tau}} \sqrt{\sum_{j=1}^{d_m} \frac{1}{\xi_j} (c_j - \overline{c}_j)^2 (b_{i_j} - b_{i_j-1})^2} < \frac{1}{L}
\]

with \( L \) defined above.

Now we turn to the special case (ii), where the inverse problem (SIP), defined in Definition 2.6, can be decomposed into two independent inverse problems (SIP1) and (SIP2), both defined at the end of Section 2. In this special case the two inverse problems (SIP1) and (SIP2) can be solved separately. Sufficient conditions for convergence rates \( O(\sqrt{\delta}) \) for the inverse problem (SIP1) are given in [3]. Especially they also impose a terminal condition of the form

\[
\overline{a}(T) = a^*(T),
\]

which seems to be not realistic to us. For the inverse problem (SIP2) it can be shown that the solution depends stably on the data and we have a convergence rate \( O(\delta) \), provided uniqueness and solvability are guaranteed. Therefore it is not necessary to analyze the source condition 4.3 in detail for the second special case.
5 Conclusions

In this paper, we have derived an operator equation that describes the inverse problem of parameter estimation in the considered two special cases of a generalized bivariate Ornstein-Uhlenbeck model. It has been proven that this operator equation is in general ill-posed and requires a regularization approach. We have shown that the Tikhonov-regularized solutions depend stably on the data and converge to the true solution, provided the regularization parameter is chosen appropriately. Furthermore, we have analyzed the form of sufficient conditions for convergence rates and have found that they are in most situations very restrictive.

Further studies are needed to treat not only the two special cases considered here, but also the general case where both parameters $\gamma$ and $\lambda$ are positive. Furthermore we intend to compare the method presented above with a 2-step approach that reconstructs the volatility function $\sigma(t)$ from the option price data in a first step and calibrates the remaining (real-valued) parameters in a second step. In this context we also plan analytical and numerical studies concerning the stability and condition of the remaining parameter estimation problem after the identification of the volatility function.

A Proofs

A.1 Proof of Lemma 2.3

With the explicit solution (2.7) of the corresponding stochastic differential equation (2.6) we obtain

$$D^2 r_\tau(t) = E \left\{ (r_\tau(t) - E r_\tau(t))^2 \right\} = E \left\{ (q(t) - q(t - \tau))^2 \right\} = E \left\{ \left( e^{-\gamma t} q_0 + \int_0^t \sigma(s) e^{-\gamma(t-s)} dW_q(s) - e^{-\gamma(t-\tau)} q_0 - \int_0^{t-\tau} \sigma(s) e^{-\gamma(t-\tau-s)} dW_q(s) \right)^2 \right\}$$

$$= (e^{-\gamma t} - e^{-\gamma(t-\tau)})^2 E q_0^2 + \int_0^t \sigma^2(s) e^{-2\gamma(t-s)} ds - 2 \int_0^{t-\tau} \sigma^2(s) e^{-\gamma(2t-\tau-2s)} ds$$

$$+ \int_0^{t-\tau} \sigma^2(s) e^{-2\gamma(t-\tau-s)} ds$$

$$= e^{-2\gamma t} (1 - e^{\gamma \tau})^2 E q_0^2 + e^{-2\gamma \tau} (1 - e^{\gamma \tau})^2 \int_0^{t-\tau} \sigma^2(s) e^{-2\gamma(t-\tau-s)} ds$$

$$+ \int_{t-\tau}^t \sigma^2(s) e^{-2\gamma(t-s)} ds.$$
Using the notation $H(t_1, t_2)$ introduced in Lemma 2.3 the result for the variance is proven. For $k \in \mathbb{N} \cup \{0\}$ the covariance of the detrended log price process is given by

\[
E \{q(t)q(t + k\tau)\} = E \left\{ \left( e^{-\gamma t} q_0 + \int_0^t \sigma(s)e^{-\gamma(t-s)} dW_q(s) \right) \left( e^{-\gamma(t+k\tau)} q_0 + \int_0^{t+k\tau} \sigma(s)e^{-\gamma(t+k\tau-s)} dW_q(s) \right) \right\} \\
= e^{-\gamma(2t+k\tau)} E q_0^2 + e^{-\gamma(k\tau)} H(0, t) \\
= e^{-\gamma k\tau} \left( e^{-2\gamma t} E q_0^2 + H(0, t) \right),
\]

and it follows for the covariance of the returns for $k \in \mathbb{N}$

\[
\text{Cov} \{r_\tau(t), r_\tau(t + k\tau)\} = E \{(q(t) - q(t - \tau))(q(t + k\tau) - q(t + (k-1)\tau))\} \\
= E \{q(t)q(t + k\tau)\} - E \{q(t)q(t + (k-1)\tau)\} \\
- E \{q(t-\tau)q(t + k\tau)\} + E \{q(t-\tau)q(t + (k-1)\tau)\} \\
= e^{-\gamma k\tau} \left( e^{-2\gamma t} E q_0^2 + H(0, t) \right) - e^{-\gamma(k-1)\tau} \left( e^{-2\gamma t} E q_0^2 + H(0, t) \right) \\
- e^{-\gamma(k+1)\tau} \left( e^{-2\gamma(t-\tau)} E q_0^2 + H(0, t-\tau) \right) \\
+ e^{-\gamma k\tau} \left( e^{-2\gamma(t-\tau)} E q_0^2 + H(0, t-\tau) \right). 
\]

This gives (2.9).

### A.2 Proof of Lemma 2.4

In the case of a constant volatility function $\sigma(t) \equiv \sigma$ the integral $H(t_1, t_2)$ can be computed explicitly. Therefore, for $\gamma > 0$ equations (2.8) and (2.9) reduce to

\[
D^2 r_\tau(t) = (1 - e^{\gamma \tau})^2 \left( e^{-2\gamma t} E q_0^2 + e^{-2\gamma \tau} \int_0^{t-\tau} e^{-2\gamma(t-\tau-s)} ds \right) + \sigma^2 \int_{t-\tau}^t e^{-2\gamma(t-s)} ds \\
= (1 - e^{\gamma \tau})^2 \left( e^{-2\gamma t} E q_0^2 + e^{-2\gamma \tau} \frac{\sigma^2}{2\gamma} (1 - e^{-2\gamma(t-\tau)}) \right) + \frac{\sigma^2}{2\gamma} (1 - e^{-2\gamma t}) \\
= (1 - e^{\gamma \tau})^2 e^{-2\gamma t} \left( E q_0^2 - \frac{\sigma^2}{2\gamma} \right) + \frac{\sigma^2}{2\gamma} (e^{-2\gamma t} (1 - e^{\gamma \tau})^2 + 1 - e^{-2\gamma t}) \\
= e^{-2\gamma t} (1 - e^{\gamma \tau})^2 \left( E q_0^2 - \frac{\sigma^2}{2\gamma} \right) + \frac{\sigma^2}{\gamma} (1 - e^{-\gamma \tau}).
\]
and
\[
\text{Cov} (r_\tau (t), r_\tau (t + k\tau)) = \frac{\sigma^2}{2\gamma} e^{-\gamma k\tau} (1 - e^{-\gamma \tau})^2 \left( 1 - e^{-2\gamma(t-\tau)} \right) \\
+ \frac{\sigma^2}{2\gamma} e^{-\gamma k\tau} (1 - e^{-\gamma \tau}) \left( 1 - e^{-2\gamma\tau} \right) \\
+ e^{-\gamma (2t + k\tau)} (1 - e^{-\gamma \tau})^2 E q_0^2 \\
= e^{-\gamma (2t + k\tau)} (1 - e^{-\gamma \tau})^2 \left( E q_0^2 - \frac{\sigma^2}{2\gamma} \right) \\
+ \frac{\sigma^2}{2\gamma} e^{-\gamma k\tau} (1 - e^{-\gamma \tau}) \left( e^{-2\gamma\tau} - e^{-\gamma \tau} + 1 - e^{-2\gamma\tau} \right) \\
= e^{-\gamma (2t + k\tau)} (1 - e^{-\gamma \tau})^2 \left( E q_0^2 - \frac{\sigma^2}{2\gamma} \right) - \frac{\sigma^2}{2\gamma} (1 - e^{-\gamma \tau})^2 e^{-\gamma (k-1)\tau}.
\]

A.3 Proof of Lemma 2.5

Outgoing from (2.23) a straightforward calculation gives
\[
D^2 r_\tau (t) = E \left\{ (q(t) - q(t - \tau))^2 \right\} \\
= E \left\{ \left( q_0 - \frac{1}{\beta} [e^{-\beta \tau} - 1] X_0 + \int_0^t \sigma(s) dW_q(s) - \frac{\sigma X}{\beta} \int_0^t (e^{-\beta(t-s)} - 1) dW_X(s) \right)^2 \right\} \\
= E \left\{ \left( \frac{1}{\beta} [e^{-\beta (t-\tau)} - e^{-\beta \tau}] X_0 + \int_0^t \sigma(s) dW_q(s) + \frac{\sigma X}{\beta} \int_0^{t-\tau} (e^{-\beta(t-\tau-s)} - 1) dW_X(s) \right)^2 \right\} \\
= \frac{1}{\beta^2} \left[ e^{-\beta(t-\tau)} - e^{-\beta \tau} \right]^2 E X_0^2 + \int_{t-\tau}^t \sigma^2(s) ds \\
+ \frac{\sigma^2 X}{\beta^2} \left[ e^{-\beta(t-\tau)} - e^{-\beta \tau} \right]^2 \int_0^{t-\tau} e^{2\beta s} ds + \frac{\sigma^2 X}{\beta^2} \int_{t-\tau}^t (e^{-\beta(t-s)} - 1)^2 ds \\
= \frac{1}{\beta^2} \left[ e^{-\beta(t-\tau)} - e^{-\beta \tau} \right]^2 E X_0^2 + \int_{t-\tau}^t \sigma^2(s) ds \\
+ \frac{\sigma^2 X}{\beta^2} \left[ e^{-\beta(t-\tau)} - e^{-\beta \tau} \right]^2 \frac{1}{2\beta} (e^{2\beta(t-\tau)} - 1) + \frac{\sigma^2 X}{\beta^2} \frac{1}{2\beta} (1 - e^{-2\beta \tau} - 4 e^{-\beta \tau} + 2\beta \tau) \\
= \frac{1}{\beta^2} \left[ e^{-\beta(t-\tau)} - e^{-\beta \tau} \right]^2 \left( E X_0^2 - \frac{\sigma^2 X}{2\beta} \right) + \int_{t-\tau}^t \sigma^2(s)
\[ + \frac{\sigma^2}{\beta^2} \frac{1}{2\beta} (1 - 2e^{-\beta \tau} + e^{-2\beta \tau} + 1 - e^{-2\beta \tau} - 4 + 4e^{-\beta \tau} + 2\beta \tau) \]

\[ = \frac{1}{\beta^2} [e^{-\beta(t-\tau)} - e^{-\beta t}]^2 \left( EX_0^2 - \frac{\sigma^2}{2\beta} \right) + \int_{t-\tau}^{t} \sigma^2(s) \, ds + \frac{\sigma^2}{\beta^2} (-1 + e^{-\beta \tau} + \beta \tau) \]

and the result for the variance is proven.

For \( k \geq 1 \) it holds

\[ \text{Cov} (r_\tau(t), r_\tau(t + k \tau)) \]

\[ = E \{ (q(t) - q(t - \tau)) (q(t + k \tau) - q(t + (k - 1) \tau)) \} \]

\[ = E \left( \left( g_0 - \frac{1}{\beta} [e^{-\beta t} - 1] X_0 + \int_0^t \sigma(s) \, dW_q(s) \right) - \frac{\sigma X}{\beta} \int_0^t (e^{-\beta(t-s)} - 1) \, dW_X(s) \right) \]

\[ - g_0 + \frac{1}{\beta} [e^{-\beta(t-\tau)} - 1] X_0 - \int_0^{t-\tau} \sigma(s) \, dW_q(s) + \frac{\sigma X}{\beta} \int_0^{t-\tau} (e^{-\beta(t-\tau-s)} - 1) \, dW_X(s) \]

\[ + \frac{\sigma X}{\beta} \int_0^{t-\tau} (e^{-\beta(t-\tau-s)} - 1) \, dW_X(s) \left( g_0 - \frac{1}{\beta} [e^{-\beta(t+k \tau)} - 1] X_0 + \int_0^{t+k \tau} \sigma(s) \, dW_q(s) \right) \]

\[ = E \left( \left( \frac{1}{\beta} [e^{-\beta(t-\tau)} - e^{-\beta t}] X_0 + \int_{t-\tau}^{t} \sigma(s) \, dW_q(s) + \frac{\sigma X}{\beta} \int_{t-\tau}^{t} (1 - e^{-\beta(t-s)}) \, dW_X(s) \right) \right. \]

\[ + \frac{\sigma X}{\beta} (e^{-\beta(t-\tau)} - e^{-\beta t}) \int_0^{t-\tau} e^{\beta s} \, dW_X(s) \left. \right) \]

\[ + \frac{1}{\beta} (e^{-\beta(t+k \tau)} - e^{-\beta(t+k \tau)}) X_0 + \int_{t+(k-1) \tau}^{t+k \tau} \sigma(s) \, dW_q(s) \]

\[ + \frac{\sigma X}{\beta} \int_{t+(k-1) \tau}^{t+k \tau} (1 - e^{-\beta(t+k \tau-s)}) \, dW_X(s) \]

\[ + \frac{\sigma X}{\beta} (e^{-\beta(t+k \tau)} - e^{-\beta(t+k \tau)}) \int_0^{t+(k-1) \tau} e^{\beta s} \, dW_X(s) \right) \}

\[ = \frac{1}{\beta^2} e^{-\beta k \tau} (e^{-\beta(t-\tau)} - e^{-\beta t})^2 EX_0^2 + \frac{\sigma X}{\beta^2} (e^{-\beta k \tau} (e^{-\beta(t-\tau)} - e^{-\beta t})^2 \int_0^{t-\tau} e^{2\beta s} \, ds \]

\[ + e^{-\beta k \tau} (e^{-\beta(t-\tau)} - e^{-\beta t}) \int_{t-\tau}^{t} (e^{\beta s} - e^{\beta(2s-\tau)}) \, ds \]
3. This result directly follows from properties 1. and 2.
4. Using l’Hospital’s Rule we get

\[
\lim_{y \to 0} f_n(y) = \lim_{y \to 0} (-1)^n n! \frac{1 - e^{-y} \left( \sum_{k=0}^{n} \frac{1}{k!} y^k \right)}{y^{n+1}}
\]

\[
= \lim_{y \to 0} (-1)^n n! \frac{e^{-y} \left( \sum_{k=0}^{n} \frac{1}{k!} y^k \right) - e^{-y} \sum_{k=1}^{n} \frac{1}{(k-1)!} y^{k-1}}{(n+1)y^n}
\]

\[
= \lim_{y \to 0} (-1)^n \frac{n!}{n+1} \frac{e^{-y} \left( \frac{1}{n!} y^n \right)}{y^n} = (-1)^n \frac{1}{n+1}.
\]

Furthermore, because of

\[
|f_n(y)| = n! \frac{1 - e^{-y} \left( \sum_{k=0}^{n} \frac{1}{k!} y^k \right)}{y^{n+1}} \leq \frac{n!}{y^{n+1}}
\]

we have \(\lim_{y \to \infty} |f_n(y)| \leq \lim_{y \to \infty} \frac{n!}{y^{n+1}} = 0\) and therefore the assertion.

### A.5 Proof of Lemma 3.5

The continuity of the function \(\tilde{h}_k((\beta, \sigma_X))\) follows from the continuity of the functions \(e^{-\beta(k+1)\tau}, (e^{\beta\tau} - 1)^2\) and \(\frac{1}{2\beta^3}\) on the interval \((0, \infty)\).

For the first limit relation we write \(\tilde{h}_k((\beta, \sigma_X)) = \sigma_X^2 e^{-\beta(k+1)\tau} \frac{(e^{\beta\tau} - 1)^2}{2\beta^3}\) and use two times l’Hospital’s Rule to obtain

\[
\lim_{\beta \to 0} \frac{1}{2\beta^3} e^{-\beta(k+1)\tau} (1 - e^{-\beta\tau})^2 = \infty.
\]

For the second limit relation we write \(\tilde{h}_k((\beta, \sigma_X)) = \sigma_X^2 e^{-\beta(k-1)\tau} \frac{(1 - e^{-\beta\tau})^2}{2\beta^3}\). From the relations \((1 - e^{-\beta\tau})^2 \leq 1, e^{-\beta(k-1)\tau} \leq 1\) for all \(\beta, \tau > 0\) and \(k \geq 1\) and the limit relation \(\lim_{\beta \to \infty} \beta^3 = \infty\) we get

\[
\lim_{\beta \to \infty} \frac{1}{2\beta^3} e^{-\beta(j+1)\tau} (1 - e^{-\beta\tau})^2 = 0.
\]

### References


