On Filtering and Estimation of a Threshold Stochastic Volatility Model

Robert J. Elliott, Chuin Ching Liew and Tak Kuen Siu

Abstract

We derive a nonlinear filter and the corresponding filter-based estimates for a threshold autoregressive stochastic volatility, (TARSV), model. Using the technique of a reference probability measure, we derive a nonlinear filter for the hidden volatility and related quantities. The filter-based estimates for the unknown parameters are then obtained from the EM algorithm.

Keywords: Stochastic Volatility; Threshold Principle; Filtering; Change of Measures; Reference Probability; EM algorithm.
1 Introduction

Stochastic volatility, (SV), models constitute one of the major classes of financial time series models. SV models have been applied to study several important problems in finance, such as option valuation, portfolio selection and risk management. These are sometimes called the three pillars of finance. The origin of the SV models may be traced to the earlier work of Clark [1], where an information counting model was introduced to model the impact of intra-day market events and news on daily returns. Tauchen and Pitts [2] extended the information counting model of Clark by describing explicitly the relationships between trading volume and daily returns. The first practical form for SV models was proposed by Taylor [3], [4], [5], where the logarithmic volatility was modeled by a first-order autoregressive time series process and there are two independent random shocks, namely, the return shock and the volatility shock, per unit of time. The SV model proposed by Taylor is also called a product process, an autoregressive random variance (ARV) process or a lognormal autoregressive process. Continuous-time versions of the SV model of Taylor and its variants were considered by Hull and White [6], Heston [7] and Wiggins [8] in the context of option valuation. A fundamental difference between SV models and another important class of financial time series models, namely, autoregressive conditional heteroscedastic (ARCH) models pioneered by Engle [9], is that the changes in volatility in SV models are caused by market, or economic forces, while the changes in volatility in ARCH models are driven by information about past prices. From this perspective, SV models are more general than ARCH models.

Despite their generality, SV models are not as popular as ARCH models. One of the main reasons is that the volatility process in SV models is not observed directly. Consequently, the standard maximum likelihood estimation method cannot be apply directly to estimate SV models, but it can be applied directly to estimate ARCH models. Hence, it is more complicated to estimate SV models than ARCH models. Some methods for estimating SV models have been proposed. Melino and Turbull [10] and Andersen and Sørensen [11] used the generalised method of moments, (GMM), for estimating SV models. Harvey et al. [12] used the Kalman-filtering method in coupled with the quasi-maximum likelihood method to estimate SV models. Jacquier et al. [13] proposed a Bayesian computational approach with the Markov Chain Monte Carlo, (MCMC), to estimate SV models. Elliott and Miao [14] and Elliott et al. [15] considered a nonlinear filtering approach together with the EM algorithm to estimate SV models.
It is documented empirically that volatility responds differently to good and bad news. In particular, future volatility increases more when bad news emerges than when good news arrives. This is known as the asymmetry effect. Different explanations have been given in the literature for this stylized fact. Early explanations were given by the works of Black [16] and Christie [17], where they postulated that declines in share price increase the debt-equity ratio, (or financial leverage), which, in turn, increase the volatility. Nelson [18] and Glosten et al. [19] considered extensions to GARCH models, namely, the EGARCH model and the GJR-GARCH model, respectively, to incorporate the asymmetry in volatility. So et al. [20] and Asai and McAleer [21], [22] investigated extensions to SV models by incorporating asymmetry in volatility. So et al. [20] introduced a threshold autoregressive stochastic volatility (TARSV) model, which is regarded as the SV counterpart of the threshold ARCH model of Li and Li [23] ¹. More specifically, the threshold SV model of So et al. is an extension of the SV model of Taylor, and postulates that the autoregressive parameters in the SV dynamics change according to the sign of the past returns. Indeed, the basic idea of the TARSV comes from the threshold time series models pioneered by Tong [25], [26], [27]. Asai and McAleer [21] incorporated asymmetry in SV models by a threshold-effect indicator function, as proposed in the GJR-GARCH model of Glosten et al. [19]. It has been shown empirically that the threshold autoregressive SV model explains the asymmetry in volatility well. The Bayesian approach coupled with Markov Chain Monte Carlo, (MCMC), methods are used to estimate threshold autoregressive SV models in the literature. The advantage of the Bayesian approach together with the MCMC method is that they are flexible enough to deal with high dimensional estimation problems. However, the error bound of the approximation based on the MCMC method is random and the MCMC method can be computationally intensive.

In this paper, we propose a novel estimation theory and method to estimate threshold autoregressive SV models based on nonlinear filtering theory. In particular, we derive a nonlinear filter and the corresponding filter-based estimates for a threshold autoregressive stochastic volatility (TARSV) model. This model can incorporate asymmetry in stochastic volatility. It can also incorporate regime switching in the stochastic volatility process without introducing an additional stochastic factor. Using the technique of a reference probability measure, we derive a nonlinear filter for the hidden volatility and related quantities which are used to derive the filter-based estimates for the

¹A precursor of the threshold ARCH model was the second generation model SETAR-ARCH model introduced in Tong [24].
unknown parameters in the TARSV model. The filter-based estimates for the unknown parameters are then obtained from the EM algorithm. We obtain (semi)-analytical filtering and estimation formulae based on recursive formulae for unnormalized filters for quantities related to the hidden volatility process. A simulation study to examine the performance of the proposed estimation method is provided. In the simulation results, the relative errors for estimating the hidden volatility are low.

The proposed filter-based estimation method for the TARSV model is self-tunning and self-calibrating in the sense that the filter-based estimates of the unknown parameters in the TARSV model improve at each run of the EM algorithm. Furthermore, the filter-based estimates are (semi)-analytical in the sense that exact recursive formulae for the filtering quantities related to the filter-based estimates are obtained. This is less computationally intensive than the Bayesian MCMC method. The proposed filter-based estimation method is developed along the line of the modern theory of stochastic processes and has a solid theoretical basis.

This paper is organized as follows. Firstly, in Section 2, we present the model assumptions, price dynamics and information structures. In Section 3, we derive a nonlinear filter for the hidden volatility. We give filters for the quantities related to the hidden volatility and use these filters to derive the estimates of the model parameters using the EM algorithm in Section 4. We then provide an extended version of the TARSV model in Section 5. In Section 6, we give a simulation study to examine the performance of the proposed filter-based estimation method. The final section gives some concluding remarks.

2 The Model and Price Dynamics

In this section, we state the model assumptions and the price dynamics. We consider a discrete time economy with uncertainty described by a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and time described by an index set $\mathcal{T} = \{0, 1, 2, \ldots\}$. The price process $\{S_t\}$ is assumed to have dynamics

$$S_t = S_{t-1} \exp \left( \mu - \frac{\sigma_t^2}{2} + \sigma_t b_t \right), \quad (t \in \mathcal{T} \setminus \{0\}).$$
Here \( \mu \) is a real constant, \( \sigma_t \) is the volatility of price change between time \( t - 1 \) and \( t \) and \( b_t \) is a standard normal random variable.

The observation process \( \{y_t\} \) is defined as:

\[
y_t := \ln \left( \frac{S_t}{S_{t-1}} \right) = \mu - \frac{\sigma_t^2}{2} + \sigma_t b_t , \quad (t \in \mathcal{T} \setminus \{0\}) ,
\]

and the hidden log-volatility process \( \{x_t\} \) is:

\[
x_t := \ln \sigma_t , \quad (t \in \mathcal{T} \setminus \{0\}) .
\]

The observation process \( \{y_t\} \) represents the log-price change process.

Let \( \{w_t^{(1)}\} \) and \( \{w_t^{(2)}\} \) be two independent and identically distributed, (i.i.d.), sequences of standard normal random variables. We suppose that \( \{b_t\}, \{w_t^{(1)}\} \) and \( \{w_t^{(2)}\} \) are stochastically independent under \( \mathbb{P} \). Under the threshold autoregressive stochastic volatility (TARSV) model, the process \( \{x_t\} \) has the following dynamics:

\[
x_t = \begin{cases} 
  c_1 + d_1 x_{t-1} + \theta_1 w_t^{(1)} & \text{if } y_{t-1} \leq 0; \\
  c_2 + d_2 x_{t-1} + \theta_2 w_t^{(2)} & \text{if } y_{t-1} > 0
\end{cases} , \quad (t \in \mathcal{T} \setminus \{0\}) .
\]

for \( t = 2, 3, \ldots \).

Write \( \phi, \psi_1 \) and \( \psi_2 \) be the probability density functions for \( b_t, w_t^{(1)} \) and \( w_t^{(2)} \), respectively. Indeed,

\[
\phi(z) = \psi_1(z) = \psi_2(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} , \quad \forall z \in \mathbb{R} .
\]

However, to avoid confusing the notation, we shall distinguish them as previously stated.

The information structures of our model are specified now. For each \( t \in \mathcal{T} \setminus \{0\} \), write

\[
\mathcal{G}_t^0 := \sigma \{x_1, x_2, \ldots, x_t, y_1, y_2, \ldots, y_t\} ,
\]
\[
\mathcal{Y}_t^0 := \sigma \{y_1, y_2, \ldots, y_t\} .
\]

Write \( \mathcal{G} = \{\mathcal{G}_t \mid t \in \mathcal{T} \setminus \{0\}\} \) and \( \mathcal{Y} = \{\mathcal{Y}_t \mid t \in \mathcal{T} \setminus \{0\}\} \) for the completion of the filtrations \( \mathcal{G}_0^0 = \{\mathcal{G}_t^0 \mid t \in \mathcal{T} \setminus \{0\}\} \) and \( \mathcal{Y}_0^0 = \{\mathcal{Y}_t^0 \mid t \in \mathcal{T} \setminus \{0\}\} \), respectively. Lastly, we define

\[
\mathcal{G}_\infty := \bigvee_{t=1}^\infty \mathcal{G}_t ,
\]

5
and suppose that \( \mathcal{G}_\infty = \mathcal{F} \).

## 3 A Nonlinear Recursive Filter for Volatility

Suppose there exists a reference probability space \((\Omega, \mathcal{G}_\infty, \mathbb{P})\), such that under \(\mathbb{P}\), \(\{x_1, x_2, \ldots\}\) is a sequence of i.i.d. standard normal random variables; similarly, under \(\mathbb{P}\), \(\{y_1, y_2, \ldots\}\) is also a sequence of i.i.d. standard normal random variables.

For each \(t = 2, 3, \ldots\), let

\[
\lambda_t^{(1)} := \frac{\psi_1(x_t - c_1 - d_1 x_{t-1})}{\theta_1 \psi_1(x_t)} \frac{\phi\left(\frac{y_t - \mu + \frac{1}{2}e^{x_t}}{e^{x_t}}\right)}{e^{x_t} \phi(y_t)},
\]

and

\[
\lambda_t^{(2)} := \frac{\psi_2(x_t - c_2 - d_2 x_{t-1})}{\theta_2 \psi_2(x_t)} \frac{\phi\left(\frac{y_t - \mu + \frac{1}{2}e^{x_t}}{e^{x_t}}\right)}{e^{x_t} \phi(y_t)}.
\]

Define, for each \(t = 2, 3, \ldots\),

\[
\lambda_t := \lambda_t^{(1)} \chi_t + \lambda_t^{(2)} (1 - \chi_t).
\]

(4)

Here \(\chi_t\) is the characteristic random variable defined by:

\[
\chi_t := \begin{cases} 
1 & \text{if } y_{t-1} \leq 0; \\
0 & \text{if } y_{t-1} > 0 
\end{cases}
\]

for \(t = 2, 3, \ldots\). Note that \(\chi_t\) is \(\mathcal{Y}_{t-1}\)-measurable.

Let

\[
\lambda_1^{(1)} = \lambda_1^{(2)} = \lambda_1 = \frac{\phi\left(\frac{y_1 - \mu + \frac{1}{2}e^{x_1}}{e^{x_1}}\right)}{e^{x_1} \phi(y_1)},
\]

and

\[
\Lambda_t = \prod_{k=1}^{t} \lambda_k.
\]
Define $\mathbb{P}$ on $\mathcal{G}_t$ by putting:
\[ \frac{d\mathbb{P}}{d\mathbb{P}^t} \bigg|_{\mathcal{G}_t} := \Lambda_t. \]

Here the existence of $\mathbb{P}$ is guaranteed by Kolmogorov’s Extension Theorem.

Then we have the following lemma, which is the key result to prove that $\{\Lambda_t\}$ is a $(\mathcal{G}_t, \mathbb{P})$-martingale.

**Lemma 3.1.** For $t = 2, 3, \ldots$, $\mathbb{E}[\Lambda_t \mid \mathcal{G}_{t-1}] = 1$, where $\mathbb{E}$ is expectation under $\mathbb{P}$.

**Proof.** The proof is adapted from that in Elliott et al. [15]. Recall that $\lambda_t = \lambda^{(1)}_t \chi_t + \lambda^{(2)}_t (1 - \chi_t)$, so
\[ \mathbb{E}[\Lambda_t \mid \mathcal{G}_{t-1}] = \mathbb{E}[\lambda^{(1)}_t \mid \mathcal{G}_{t-1}] \chi_t + \mathbb{E}[\lambda^{(2)}_t \mid \mathcal{G}_{t-1}] (1 - \chi_t). \]

Firstly, we evaluate $\mathbb{E}[\lambda^{(1)}_t \mid \mathcal{G}_{t-1}]$:
\[
\begin{align*}
\mathbb{E}[\lambda^{(1)}_t \mid \mathcal{G}_{t-1}] &= \mathbb{E} \left[ \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) \right]_{\mathcal{G}_{t-1}} \\
&= \mathbb{E} \left[ \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) \mathbb{E} \left[ \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) \mid \mathcal{G}_{t-1} \right] \sigma \{ x_t \} \right]_{\mathcal{G}_{t-1}} \\
&= \mathbb{E} \left[ \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \int_{-\infty}^{\infty} \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) e^{-x_{t}} dy_t \mathbb{E} \left[ \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) \phi \left( y_t \right) \mid \mathcal{G}_{t-1} \right] \right]_{\mathcal{G}_{t-1}} \\
&= \mathbb{E} \left[ \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \int_{-\infty}^{\infty} \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) e^{-x_{t}} \phi \left( y_t \right) dy_t \mathbb{E} \left[ \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) \phi \left( y_t \right) \mid \mathcal{G}_{t-1} \right] \right]_{\mathcal{G}_{t-1}} \\
&= \mathbb{E} \left[ \int_{-\infty}^{\infty} \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) e^{-x_{t}} \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) e^{-x_{t}} dy_t \mathbb{E} \left[ \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) \phi \left( y_t \right) \mid \mathcal{G}_{t-1} \right] \right]_{\mathcal{G}_{t-1}} \\
&= \int_{-\infty}^{\infty} \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) e^{-x_{t}} \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) e^{-x_{t}} dy_t \mathbb{E} \left[ \phi \left( \frac{y_t - \mu + \frac{1}{2} x_{2t}}{e^{x_{t}}} \right) \phi \left( y_t \right) \mid \mathcal{G}_{t-1} \right]_{\mathcal{G}_{t-1}} \\
&= 1.
\end{align*}
\]
Similarly, replacing “1” by “2” appropriately in the above calculations gives:

$$\mathbb{E}[\lambda_t^{(2)} | \mathcal{G}_{t-1}] = 1 .$$

Consequently,

$$\mathbb{E}[\lambda_t | \mathcal{G}_{t-1}] = \chi_t + (1 - \chi_t) = 1 .$$

Hence the result follows. \(\Box\)

The following lemma gives the probability laws for \(\{b_t\}, \{w_t^{(1)}\}\), and \(\{w_t^{(2)}\}\) under the real-world probability \(\mathbb{P}\).

**Lemma 3.2.** Under \(\mathbb{P}\), \(\{w_t^{(1)}\}, \{w_t^{(2)}\}\), and \(\{b_t\}\) are independent sequences of i.i.d. standard normal random variables.

**Proof.** Let \(f\) and \(g\) be any measurable test functions. Then using a version of the Bayes’ rule,

$$\mathbb{E}[f(w_t^{(1)})g(b_t) | \mathcal{G}_{t-1}, y_{t-1} \leq 0] = \frac{\mathbb{E}[\Lambda_t f(w_t^{(1)})g(b_t) | \mathcal{G}_{t-1}, y_{t-1} \leq 0]}{\mathbb{E}[\Lambda_t | \mathcal{G}_{t-1}, y_{t-1} \leq 0]}$$

$$= \frac{\Lambda_{t-1} \mathbb{E}[\lambda_t f(w_t^{(1)})g(b_t) | \mathcal{G}_{t-1}, y_{t-1} \leq 0]}{\Lambda_{t-1} \mathbb{E}[\lambda_t | \mathcal{G}_{t-1}, y_{t-1} \leq 0]}$$

$$= \frac{\mathbb{E}[\lambda_t f(w_t^{(1)})g(b_t) | \mathcal{G}_{t-1}, y_{t-1} \leq 0]}{\mathbb{E}[\lambda_t | \mathcal{G}_{t-1}, y_{t-1} \leq 0]}$$

By definition, \(\chi_t = 1\) when \(y_{t-1} \leq 0\), so it follows that

$$\mathbb{E}[\lambda_t | \mathcal{G}_{t-1}, y_{t-1} \leq 0]$$

$$= \mathbb{E}[\chi_t \lambda_t^{(1)} + (1 - \chi_t) \lambda_t^{(2)} | \mathcal{G}_{t-1}, y_{t-1} \leq 0]$$

$$= \mathbb{E}[\lambda_t^{(1)} | \mathcal{G}_{t-1}, y_{t-1} \leq 0]$$

$$= \mathbb{E}[\lambda_t^{(1)} | I_{\{y_{t-1} \leq 0\}} | \mathcal{G}_{t-1}]$$

$$= \chi_t \mathbb{E}[\lambda_t^{(1)} | \mathcal{G}_{t-1}]$$

$$= \chi_t .$$

Now

$$\mathbb{E}[\lambda_t f(w_t^{(1)})g(b_t) | \mathcal{G}_{t-1}, y_{t-1} \leq 0]$$

$$= \mathbb{E}[(\chi_t \lambda_t^{(1)} + (1 - \chi_t) \lambda_t^{(2)}) f(w_t^{(1)})g(b_t) | \mathcal{G}_{t-1}, y_{t-1} \leq 0]$$

$$= \mathbb{E}[\lambda_t^{(1)} f(w_t^{(1)})g(b_t) | \mathcal{G}_{t-1}] .$$
Again by a version of the Bayes' rule,

\[
\chi_t = \text{E}
\]

since \( \chi_t = 1 \) when \( y_{t-1} \leq 0 \), this is:

\[
= \chi_t \text{E}\left[ f\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \psi_1\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \right] \times g(b_t) \frac{\phi\left(\frac{y_t - \mu + \frac{1}{2} e^{2x_t}}{e^{2x_t}}\right)}{e^{2x_t} \phi(y_t)} | \mathcal{G}_{t-1} \]

\[
= \chi_t \text{E}\left[ f\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \psi_1\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \right] \times \int_{-\infty}^{\infty} g\left(\frac{y_t - \mu + \frac{1}{2} e^{2x_t}}{e^{2x_t}}\right) \frac{\phi\left(\frac{y_t - \mu + \frac{1}{2} e^{2x_t}}{e^{2x_t}}\right)}{e^{2x_t} \phi(y_t)} \phi(y_t)dy_t | \mathcal{G}_{t-1} \]

\[
= \chi_t \text{E}\left[ f\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \psi_1\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \right] \int_{-\infty}^{\infty} g(z) \phi(z) dz | \mathcal{G}_{t-1} \]

\[
= \chi_t \int_{-\infty}^{\infty} f\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \psi_1\left(\frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \psi_1(x_t) dx_t \int_{-\infty}^{\infty} g(z) \phi(z) dz
\]

Since \( f, g \) are arbitrary, the claim is true for \( w^{(1)}_t \). We can prove the case for \( w^{(2)}_t \) by interchanging “1” and “2” appropriately in this proof.

We now derive a nonlinear recursive filter for the hidden logarithmic volatility. For any measurable function \( h \), we are interested in the quantity

\[
\text{E}[h(x_t) | \mathcal{Y}_t].
\]

Again by a version of the Bayes' rule,

\[
\text{E}[h(x_t) | \mathcal{Y}_t] = \frac{\text{E}[\Lambda_t h(x_t) | \mathcal{Y}_t]}{\text{E}[\Lambda_t | \mathcal{Y}_t]}.
\]

(9)

Suppose, for each \( t \in \mathcal{T} \setminus \{0\} \), there exists a density \( q_t(z) \) such that

\[
\text{E}[\Lambda_t h(x_t) | \mathcal{Y}_t] = \int_{-\infty}^{\infty} h(z) q_t(z) dz.
\]

(10)
The following theorem then gives a recursion for the density \( q_t(z) \).

**Theorem 3.1.** For each \( t = 2, 3, \ldots \),

\[
q_t(z) = D(y_t, z) \int_{-\infty}^{\infty} \left( \chi_t \psi_1 \left( \frac{z - c_1 - d_1 x}{\theta_1} \right) + (1 - \chi_t) \psi_2 \left( \frac{z - c_2 - d_2 x}{\theta_2} \right) \right) q_{t-1}(x) \, dx ,
\]

where

\[
D(y, z) = \frac{\phi \left( y - \mu + \frac{1}{2} e^z \right)}{e^z \phi(y)}. 
\]

**Proof.** Note that

\[
\int_{-\infty}^{\infty} h(z) q_t(z) \, dz = \mathbb{E}[\Lambda_t h(x_t) \mid \mathcal{Y}_t]
\]

\[
= \mathbb{E}[\Lambda_{t-1} \lambda_t h(x_t) \mid \mathcal{Y}_t]
\]

\[
= \mathbb{E}[\Lambda_{t-1} \lambda_1^{(1)} h(x_t) \mid \mathcal{Y}_t] + \mathbb{E}[\Lambda_{t-1} \lambda_1^{(2)} h(x_t) \mid \mathcal{Y}_t] (1 - \chi_t). 
\]

Now

\[
\mathbb{E}[\Lambda_{t-1} \lambda_1^{(1)} h(x_t) \mid \mathcal{Y}_t]
\]

\[
= \mathbb{E} \left[ \Lambda_{t-1} \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} e^{z_1}}{e^{z_1}} \right) h(x_t) \mid \mathcal{Y}_t \right]
\]

\[
= \frac{1}{\theta_1 \phi(y_t)} \mathbb{E} \left[ \Lambda_{t-1} \int_{-\infty}^{\infty} \psi_1 \left( \frac{x_t - c_1 - d_1 x_{t-1}}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} e^{z_1}}{e^{z_1}} \right) h(x_t) \psi_1(x_t) \, dx_t \mid \mathcal{Y}_t \right]
\]

\[
= \frac{1}{\theta_1 \phi(y_t)} \mathbb{E} \left[ \Lambda_{t-1} \int_{-\infty}^{\infty} e^{-z} \psi_1 \left( \frac{z - c_1 - d_1 x_{t-1}}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} e^{2z}}{e^z} \right) h(z) \, dz \mid \mathcal{Y}_t \right] 
\]

since the integral inside the expectation is a function of \( x_{t-1} \), this is:

\[
= \frac{1}{\theta_1 \phi(y_t)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-z} \psi_1 \left( \frac{z - c_1 - d_1 x_{t-1}}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} e^{2z}}{e^z} \right) h(z) \, dz \, q_{t-1}(x_{t-1}) \, dx_{t-1}
\]

\[
= \frac{1}{\theta_1 \phi(y_t)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-z} \psi_1 \left( \frac{z - c_1 - d_1 x}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} e^{2z}}{e^z} \right) h(z) \, dz \, q_{t-1}(x) \, dx
\]

\[
= \frac{1}{\theta_1 \phi(y_t)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-z} \psi_1 \left( \frac{z - c_1 - d_1 x}{\theta_1} \right) \phi \left( \frac{y_t - \mu + \frac{1}{2} e^{2z}}{e^z} \right) h(z) q_{t-1}(x) \, dx \, dz
\]

\[
= \int_{-\infty}^{\infty} h(z) D(y_t, z) \left( \int_{-\infty}^{\infty} \frac{\psi_1 \left( \frac{z - c_1 - d_1 x}{\theta_1} \right)}{\theta_1} q_{t-1}(x) \, dx \right) \, dz .
\]
Using the same arguments gives:

\[ E[\Lambda_{t-1} \lambda_t^{(2)} h(x_t) \mid \mathcal{Y}_t] = \int_{-\infty}^{\infty} D(y_t, z) h(z) \left( \int_{-\infty}^{\infty} \frac{\psi_2(z - c_2 - d_2 x_2)}{\theta_2} q_{t-1}(x) \, dx \right) \, dz . \]

Consequently,

\[ \int_{-\infty}^{\infty} h(z) q_t(z) \, dz = \int_{-\infty}^{\infty} h(z) D(y_t, z) \left[ \int_{-\infty}^{\infty} \left( \chi_t \frac{\psi_1(z - c_1 - d_1 x_1)}{\theta_1} + (1 - \chi_t) \frac{\psi_2(z - c_2 - d_2 x_2)}{\theta_2} \right) q_{t-1}(x) \, dx \right] \, dz . \]

Since this holds for any arbitrary measurable test function \( h \), the result follows.

Setting \( h(z) = 1 \) in (10) gives:

\[ E[\Lambda_t \mid \mathcal{Y}_t] = \int_{-\infty}^{\infty} q_t(z) \, dz . \]

Consequently, (9) gives:

\[ E[h(x_t) \mid \mathcal{Y}_t] = \frac{\int_{-\infty}^{\infty} h(z) q_t(z) \, dz}{\int_{-\infty}^{\infty} q_t(z) \, dz} , \]

and as a special case that \( h(z) = z \), we obtain

\[ E[x_t \mid \mathcal{Y}_t] = \frac{\int_{-\infty}^{\infty} z q_t(z) \, dz}{\int_{-\infty}^{\infty} q_t(z) \, dz} . \]  

(11)

### 4 The EM Algorithm

In this section, we first derive nonlinear recursive filters for some quantities and then use these filters to derive (semi)-analytical formulae for the estimates of the unknown parameters based on the EM algorithm.
4.1 Filters for some quantities

Suppose that $H, F, G$ be arbitrary test measurable functions. Define, for each $t = 2, 3, \ldots$,

$$S_t := \sum_{k=2}^{t} \chi_k H(y_k) F(x_k) G(x_{k-1}) .$$

For deriving the (semi)-analytical formulae for the estimates of the unknown parameters, we wish to evaluate the following quantity:

$$E[S_t | Y_t] = E\left[\sum_{k=2}^{t} \chi_k H(y_k) F(x_k) G(x_{k-1}) | Y_t\right] .$$

Again by a version of the Bayes’ rule,

$$E[S_t h(x_t) | Y_t] = \frac{E[\Lambda_t S_t h(x_t) | Y_t]}{E[\Lambda_t | Y_t]} .$$

Consider a measure associated with $S$ and suppose that there is then a density $L_t(z)$ such that

$$E[\Lambda_t S_t h(x_t) | Y_t] = \int_{-\infty}^{\infty} h(z) L_t(z) \, dz , \quad t \in T \setminus \{0\} ,$$

for any integrable function $h$.

The following theorem gives, under this assumption, a recursion for the density $L_t$.

**Theorem 4.1.** For each $t = 2, 3, \ldots$,

$$L_t(z) = D(y_t, z) \left[ \int_{-\infty}^{\infty} \left( \frac{\chi_t \psi_1(z-c_1-d_1 x)}{\theta_1} + \frac{(1-\chi_t) \psi_2(z-c_2-d_2 x)}{\theta_2} \right) L_{t-1}(x) \, dx 
+ H(y_t) F(z) \int_{-\infty}^{\infty} \left( \frac{\chi_t \psi_1(z-c_1-d_1 x)}{\theta_1} \right) G(x) q_{t-1}(x) \, dx \right] .$$

**Proof.** Let $h$ be an arbitrary measurable test function. Then

$$\int_{-\infty}^{\infty} h(z) L_t(z) \, dz = E[\Lambda_t S_t h(x_t) | Y_t] .$$
By the definition of $S_t$, we have:

\[
\int_{-\infty}^{\infty} h(z)L_t(z) \, dz
\]

\[
= \mathbb{E}[\Lambda_{t-1}\lambda_t S_{t-1} h(x_t) \mid \mathcal{Y}_t] + \mathbb{E}[\Lambda_{t-1}\lambda_t \chi_t H(y_t) F(x_t) G(x_{t-1}) h(x_t) \mid \mathcal{Y}_t]
\]

\[
= \chi_t \left\{ \mathbb{E}[\Lambda_{t-1}\lambda_t^{(1)} S_{t-1} h(x_t) \mid \mathcal{Y}_t] + \mathbb{E}[\Lambda_{t-1}\lambda_t^{(1)} H(y_t) F(x_t) G(x_{t-1}) h(x_t) \mid \mathcal{Y}_t] \right\} 
+ (1 - \chi_t) \mathbb{E}[\Lambda_{t-1}\lambda_t^{(2)} S_{t-1} h(x_t) \mid \mathcal{Y}_t],
\]

by using (4) and $\chi_t(1 - \chi_t) = 0$. Consider the first term:

\[
\mathbb{E}[\Lambda_{t-1}\lambda_t^{(1)} S_{t-1} h(x_t) \mid \mathcal{Y}_t] + \mathbb{E}[\Lambda_{t-1}\lambda_t^{(1)} H(y_t) F(x_t) G(x_{t-1}) h(x_t) \mid \mathcal{Y}_t]
\]

\[
= \mathbb{E} \left[ \Lambda_{t-1} S_{t-1} \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(x_t)} \phi \left( \frac{y_t - \mu + \frac{1}{2} \sigma^2}{\sigma} \right) h(x_t) \mid \mathcal{Y}_t \right] 
+ \mathbb{E} \left[ \Lambda_{t-1} \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(x_t)} \phi \left( \frac{y_t - \mu + \frac{1}{2} \sigma^2}{\sigma} \right) H(y_t) F(x_t) G(x_{t-1}) h(x_t) \mid \mathcal{Y}_t \right],
\]

using the towers property, this is:

\[
= \mathbb{E} \left[ \Lambda_{t-1} S_{t-1} \mathbb{E} \left[ \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(x_t)} D(y_t, x_t) h(x_t) \mid \mathcal{Y}_t \right] \right] 
+ \mathbb{E} \left[ \Lambda_{t-1} \mathbb{E} \left[ \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(x_t)} D(y_t, x_t) H(y_t) F(x_t) G(x_{t-1}) h(x_t) \mid \mathcal{Y}_t \right] \right]
\]

\[
= \mathbb{E} \left[ \Lambda_{t-1} S_{t-1} \int_{-\infty}^{\infty} D(y_t, z) \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(z_t)} h(z) \psi_1(z) \, dz \mid \mathcal{Y}_t \right] 
+ \mathbb{E} \left[ \Lambda_{t-1} \int_{-\infty}^{\infty} D(y_t, z) \psi_1(z_1 - c_1 d_1 x_t - 1) \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(z_t)} H(y_t) F(z) G(x_{t-1}) h(z) \psi_1(z) \, dz \mid \mathcal{Y}_t \right]
\]

\[
= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} D(y_t, z) \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(z_t)} h(z) \, dz \right) L_{t-1}(x) \, dx 
+ \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} D(y_t, z) \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(z_t)} H(y_t) F(z) G(x_t) h(z) \, dz \right) q_{t-1}(x) \, dx,
\]

interchanging the order of integration, this is:

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(y_t, z) \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(z_t)} h(z) L_{t-1}(x) \, dx \, dz 
+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(y_t, z) \frac{\psi_1(z_1 - c_1 d_1 x_t - 1)}{\theta_1 \psi_1(z_t)} H(y_t) F(z) G(x_t) h(z) q_{t-1}(x) \, dx \, dz.
\]
Consequently,
\[ \mathbb{E}[\Lambda_{t-1}\lambda_t^{(1)}S_{t-1}h(x_t) \mid \mathcal{Y}_t] + \mathbb{E}[\Lambda_{t-1}\lambda_t^{(1)}H(y_t)F(x_t)G(x_{t-1}) \mid \mathcal{Y}_t] \]
\[ = \int_{-\infty}^{\infty} h(z) D(y_t, z) \left[ \int_{-\infty}^{\infty} \psi_1\left(\frac{z-c_1-d_1 x}{\theta_1}\right) L_{t-1}(x) \, dx \right. \]
\[ + H(y_t)F(z) \int_{-\infty}^{\infty} \psi_1\left(\frac{z-c_1-d_1 x}{\theta_1}\right) G(x)q_{t-1}(x) \, dx \left. \right] \, dz . \]

Similarly,
\[ \mathbb{E}[\Lambda_{t-1}\lambda_t^{(2)}S_{t-1}h(x_t) \mid \mathcal{Y}_t] = \int_{-\infty}^{\infty} h(z) D(y_t, z) \int_{-\infty}^{\infty} \frac{\psi_2\left(\frac{z-c_2-d_2 x}{\theta_2}\right)}{\theta_2} L_{t-1}(x) \, dx \, dz . \]

Finally,
\[ \int_{-\infty}^{\infty} h(z) L_t(z) \, dx \]
\[ = \int_{-\infty}^{\infty} h(z) D(y_t, z) \left[ \int_{-\infty}^{\infty} \left( \chi_t \psi_1\left(\frac{z-c_1-d_1 x}{\theta_1}\right) \frac{1}{\theta_1} \right) + \left(1 - \chi_t \right) \psi_2\left(\frac{z-c_2-d_2 x}{\theta_2}\right) \frac{1}{\theta_2} \right] L_{t-1}(x) \, dx \]
\[ + H(y_t)F(z) \int_{-\infty}^{\infty} \left( \frac{\chi_t \psi_1\left(\frac{z-c_1-d_1 x}{\theta_1}\right)}{\theta_1} \right) G(x)q_{t-1}(x) \, dx \right] \, dz . \]

Since this identity holds for any arbitrary function \( h \), the result follows.

For each \( t = 2, 3, \ldots \), let
\[ R_t := \sum_{k=2}^{t} (1 - \chi_k) H(y_k)F(x_k)G(x_{k-1}) . \]

We then wish to evaluate the following expectation:
\[ \mathbb{E}[R_t \mid \mathcal{Y}_t] . \]

Again, consider a measure associated with \( R \) and suppose there exists a density \( M_t(z) \) such that
\[ \mathbb{E}[\Lambda_t R_t h(x_t) \mid \mathcal{Y}_t] = \int_{-\infty}^{\infty} h(z) M_t(z) \, dz , \quad t \in \mathcal{T} , \quad (14) \]
for any integrable function \( h \).

The following theorem gives a recursion for the density \( M_t(z) \).
Theorem 4.2. For each $t = 2, 3, \ldots$, 
\[ M_t(z) = D(y_t, z) \left[ \int_{-\infty}^{\infty} \left( \frac{\chi_t \psi_1(z - c_1 - d_1 x)}{\theta_1} + \frac{(1 - \chi_t) \psi_2(z - c_2 - d_2 x)}{\theta_2} \right) M_{t-1}(x) \, dx \right. \\
+ H(y_t) F(z) \int_{-\infty}^{\infty} \left( \frac{(1 - \chi_t) \psi_2(z - c_2 - d_2 x)}{\theta_2} \right) G(x) q_{t-1}(x) \, dx \right]. \]  

(15)

Proof. Let $h$ be an arbitrary measurable test function. Then 
\[
\int_{-\infty}^{\infty} h(z) M_t(z) \, dz \\
= \mathbb{E}[\Lambda_t R_th(x_t) \mid \mathcal{Y}_t] \\
= \mathbb{E}[\Lambda_{t-1} \chi_t R_{t-1}h(x_t) \mid \mathcal{Y}_t] + (1 - \chi_t) \mathbb{E}[\Lambda_{t-1} \chi_t H(y_t) F(x_t) G(x_{t-1}) \mid \mathcal{Y}_t] \\
= \chi_t \mathbb{E}[\Lambda_{t-1} \chi_t R_{t-1}h(x_t) \mid \mathcal{Y}_t] + (1 - \chi_t) \{ \mathbb{E}[\Lambda_{t-1} \chi_t H(y_t) F(x_t) G(x_{t-1}) h(x_t) \mid \mathcal{Y}_t] \\
+ \mathbb{E}[\Lambda_{t-1} \chi_t R_{t-1} h(x_t) \mid \mathcal{Y}_t] \}. 
\]

Applying similar arguments as in the previous proof yields the result. \hfill \Box

Putting $h(z) = 1$ in (12) gives:
\[
\mathbb{E}[\Lambda_t S_t \mid \mathcal{Y}_t] = \int_{-\infty}^{\infty} L_t(z) \, dz.
\]

By a version of the Bayes’ rule,
\[
\mathbb{E}[S_t \mid \mathcal{Y}_t] = \frac{\mathbb{E}[\Lambda_t S_t \mid \mathcal{Y}_t]}{\mathbb{E}[\Lambda_t \mid \mathcal{Y}_t]} = \frac{\int_{-\infty}^{\infty} L_t(z) \, dz}{\int_{-\infty}^{\infty} q_t(z) \, dz}.
\]  

(16)

Applying the same arguments to (14) then gives:
\[
\mathbb{E}[R_t \mid \mathcal{Y}_t] = \frac{\int_{-\infty}^{\infty} M_t(z) \, dz}{\int_{-\infty}^{\infty} q_t(z) \, dz}.
\]  

(17)

Each of the remarks below gives a special case of $S_t$, which is used to derive the (semi)-analytical formulae for the estimates of the model parameters using the EM algorithm.

Remarks:
1. Put \( H(z) = G(z) = 1 \) and \( F(z) = z \), write
\[
S_t^{(1)} = \sum_{k=2}^{t} \chi_k x_k, \quad R_t^{(1)} = \sum_{k=2}^{t} (1 - \chi_k) x_k;
\]
2. Put \( H(z) = F(z) = 1 \) and \( G(z) = z \), write
\[
S_t^{(2)} = \sum_{k=2}^{t} \chi_k x_{k-1}, \quad R_t^{(2)} = \sum_{k=2}^{t} (1 - \chi_k) x_{k-1};
\]
3. Put \( H(z) = 1 \) and \( F(z) = G(z) = z \), write
\[
S_t^{(3)} = \sum_{k=2}^{t} \chi_k x_k x_{k-1}, \quad R_t^{(3)} = \sum_{k=2}^{t} (1 - \chi_k) x_k x_{k-1};
\]
4. Put \( H(z) = F(z) = 1 \) and \( G(z) = z^2 \), write
\[
S_t^{(4)} = \sum_{k=2}^{t} \chi_k x_{k-1}^2, \quad R_t^{(4)} = \sum_{k=2}^{t} (1 - \chi_k) x_{k-1}^2;
\]
5. Put \( H(z) = G(z) = 1 \) and \( F(z) = z^2 \), write
\[
S_t^{(5)} = \sum_{k=2}^{t} \chi_k x_k^2, \quad R_t^{(5)} = \sum_{k=2}^{t} (1 - \chi_k) x_k^2;
\]
6. Put \( G(z) = 1 \), \( H(z) = z \) and \( F(z) = e^{-2z} \), write
\[
S_t^{(6)} = \sum_{k=2}^{t} \chi_k y_k e^{-2x_k}, \quad R_t^{(6)} = \sum_{k=2}^{t} (1 - \chi_k) y_k e^{-2x_k};
\]
7. Put \( H(z) = G(z) = 1 \) and \( F(z) = e^{-2z} \), write
\[
S_t^{(7)} = \sum_{k=2}^{t} \chi_k e^{-2x_k}, \quad R_t^{(7)} = \sum_{k=2}^{t} (1 - \chi_k) e^{-2x_k};
\]
8. For each \( k = 1, 2, \ldots, 7 \), let \( L_t^{(k)} \) and \( M_t^{(k)} \) be the corresponding density of \( S_t^{(k)} \) and \( R_t^{(k)} \) as in Theorem 4.1 and 4.2. In view of Theorem 4.1 and 4.2, we can compute each of the \( L_t^{(k)} \) and \( M_t^{(k)} \), subsequently estimate each \( \mathbb{E}[S_t^{(k)} \mid \mathcal{Y}_t] \) and \( \mathbb{E}[R_t^{(k)} \mid \mathcal{Y}_t] \) using (16) and (17).
4.2 An EM algorithm for parameter estimation

**Proposition 4.1.** The estimate for parameter $c_1$ at time $t$ is

$$\hat{c}_1(t) = \frac{\int_{-\infty}^{\infty} (L_t^{(1)}(z) - d_1 L_t^{(2)}(z)) dz}{\sum_{k=1}^{t} \chi_k \int_{-\infty}^{\infty} g_t(z) dz}.$$ 

**Proof.** The dynamics corresponding to the parameters $(c_1, d_1, \theta_1, c_2, d_2, \theta_2, \mu)$ are given by densities $\Lambda_{t}^{c_1}$ where

$$\Lambda_{t}^{c_1} = \prod_{k=2}^{t} D(y_k, x_k) \left[ \chi_k \frac{\psi_1(x_k-c_1-d_1 x_k-1)}{\psi_1(x_k)} + (1 - \chi_k) \frac{\psi_2(x_k-c_1-d_1 x_k-1)}{\psi_2(x_k)} \right].$$

Similarly, consider a density for the parameter set $(c'_1, d_1, \theta_1, c_2, d_2, \theta_2, \mu)$,

$$\Lambda_{t}^{c'_1} = \prod_{k=2}^{t} D(y_k, x_k) \left[ \chi_k \frac{\psi_1(x_k-c'_1-d_1 x_k-1)}{\psi_1(x_k)} + (1 - \chi_k) \frac{\psi_2(x_k-c'_1-d_1 x_k-1)}{\psi_2(x_k)} \right].$$

Therefore, a density which changes the parameter $c'_1$ to $c_1$ is given by considering the Radon-Nikodym derivative of the ratio of $\Lambda_{t}^{c_1}$ and $\Lambda_{t}^{c'_1}$,

$$\frac{d\pi^{c_1}}{d\pi^{c'_1}} \bigg|_{\mathcal{G}_t} = \frac{\Lambda_{t}^{c_1}}{\Lambda_{t}^{c'_1}} = \prod_{k=2}^{t} \left[ \chi_k \frac{\psi_1(x_k-c_1-d_1 x_k-1)}{\psi_1(x_k-c'_1-d_1 x_k-1)} + (1 - \chi_k) \right].$$

by noticing $f(\chi_t x + (1 - \chi_t)y) = \chi_t f(x) + (1 - \chi_t)f(y)$ as a special feature of the characteristic function we defined. Define set

$$\mathcal{Y}_t := \{ k : k \in \mathbb{Z}, 2 \leq k \leq t, \chi_k = 1 \},$$

and so,

$$\frac{d\pi^{c_1}}{d\pi^{c'_1}} \bigg|_{\mathcal{G}_t} = \prod_{k \in \mathcal{Y}_t} \frac{\psi_1(x_k-c_1-d_1 x_k-1)}{\psi_1(x_k-c'_1-d_1 x_k-1)}.$$ 

Our goal is to find $c_1$ so as to maximize the conditional expectation $E \left[ \ln \left( \frac{d\pi^{c_1}}{d\pi^{c'_1}} \right) \bigg| \mathcal{Y}_t \right]$. Note that

$$E \left[ \ln \left( \frac{d\pi^{c_1}}{d\pi^{c'_1}} \right) \bigg| \mathcal{Y}_t \right] = E \left[ \sum_{k \in \mathcal{Y}_t} -\frac{(x_k - c_1 - d_1 x_k-1)^2}{2\theta_1^2} + r(c'_1) \bigg| \mathcal{Y}_t \right],$$

17
where \( r(c'_1) \) is independent of \( c_1 \). Set

\[
\frac{d}{dc_1} E \left[ \sum_{k \in \mathcal{N}_t} \frac{(x_k - c_1 - d_1 x_{k-1})^2}{2\theta_1^2} + r(c'_1) \mid \mathcal{Y}_t \right] = 0 ,
\]

which gives:

\[
E \left[ \sum_{k \in \mathcal{N}_t} \frac{x_k - c_1 - d_1 x_{k-1}}{\theta_1^2} \mid \mathcal{Y}_t \right] + 0 \mid \mathcal{Y}_t = 0 ,
\]

and so

\[
E \left[ \sum_{k \in \mathcal{N}_t} x_k \mid \mathcal{Y}_t \right] - d_1 E \left[ \sum_{k \in \mathcal{N}_t} x_{k-1} \mid \mathcal{Y}_t \right] = c_1 E \left[ \sum_{k \in \mathcal{N}_t} 1 \mid \mathcal{Y}_t \right].
\]

Consequently, the estimate for parameter \( c_1 \) at time \( t \) is given by:

\[
\hat{c}_1(t) = \frac{E \left[ \sum_{k=2}^{t} \chi_k x_k \mid \mathcal{Y}_t \right] - d_1 E \left[ \sum_{k=2}^{t} \chi_k x_{k-1} \mid \mathcal{Y}_t \right]}{\sum_{k=2}^{t} \chi_k}
\]

\[= \frac{E[S(t) \mid \mathcal{Y}_t] - d_1 E[S(t) \mid \mathcal{Y}_t]}{\sum_{k=2}^{t} \chi_k}
\]

\[= \frac{\int_{-\infty}^{\infty} (L_{\mathcal{Y}_t}(z) - d_1 L_{\mathcal{Y}_t}(z))dz}{\sum_{k=2}^{t} \chi_k \int_{-\infty}^{\infty} q_t(z)dz}.\]

We can derive similar results for the estimates for \( c_2, d_1, d_2, \theta_1, \theta_2 \) and \( \mu \). We shall only outline the proof for \( c_2 \) and \( \mu \). Other cases follow similarly.

A density that changes \( c_2 \) to \( c'_2 \) is given by:

\[
\frac{d \pi^{c_2,c'_2}}{d \pi^{c_2}} \bigg| \mathcal{G}_t = \prod_{k=2}^{t} \left[ \chi_k + (1 - \chi_k) \frac{\psi_2(x_k - c_2 - d_2 x_{k-1})}{\psi_2(x_k - c'_2 - d_2 x_{k-1})} \right]
\]

\[= \prod_{k \notin \mathcal{N}_t} \frac{\psi_2(x_k - c_2 - d_2 x_{k-1})}{\psi_2(x_k - c'_2 - d_2 x_{k-1})}.\]

Using the same arguments as above gives:

\[
\hat{c}_2(t) = \frac{\int_{-\infty}^{\infty} (M_{\mathcal{Y}_t}(z) - d_1 M_{\mathcal{Y}_t}(z))dz}{\sum_{k=2}^{t} (1 - \chi_k) \int_{-\infty}^{\infty} q_t(z)dz}.\]
For the estimation of $\mu$, a density that changes $\mu$ to $\mu'$ is given by:

$$
\frac{dP_{\mu}}{dP_{\mu'}}|_{G_t} = \prod_{k=2}^{t} \frac{\phi(e^{-x_k}(y - \mu - e^{2x_k}/2))}{\phi(e^{-x_k}(y - \mu' - e^{2x_k}/2))}.
$$

The computation of $\hat{\mu}$ is similar to the two presented above, except that the summations and products do not involve $\chi_k$. Consequently, the quantities $S(k) + R(k)$ are used. The formula for the estimate $\hat{\mu}$ is given by:

$$
\hat{\mu}(t) = \frac{\int_{-\infty}^{\infty} (L_t^{(6)}(z) + M_t^{(6)}(z))dz + \frac{1}{2} \int_{-\infty}^{\infty} q_t(z)dz}{\int_{-\infty}^{\infty} q_t(z)dz}.
$$

Formulae for the estimates $\hat{d}_1, \hat{d}_2, \hat{\theta}_1, \hat{\theta}_2$ can be derived in the same way. Their formulae will be given in the next section, together with a method for estimating these parameters.

### 4.3 Summary of the estimation results and procedures

In this section, we summarize the results we obtained previously and give the procedures to estimate the model parameters and the hidden volatility. Suppose the current time is $T$, and we have the data set that contains the log-price changes $\{y_1, y_2, \ldots, y_T\}$. Choose an initial value $q_0 = \rho_0$, which is the a priori distribution of $x_0$. We first compute the function $q_t(z)$ for all $t = 1, 2, \ldots, T$ using the following recursion:

$$
q_t(z) = D(y_t, z) \int_{-\infty}^{\infty} \left( \chi_t \frac{\psi_1(z - \chi_t \frac{d_1}{\theta_1})}{\theta_1} + (1 - \chi_t) \frac{\psi_2(z - \chi_t \frac{d_2}{\theta_2})}{\theta_2} \right) q_{t-1}(x)dx
$$

where

$$
D(y, z) = \frac{\phi(e^{-z}(y + \frac{1}{2}e^{2z}))}{e^z\phi(y)}.
$$

Write $s_T = \sum_{k=2}^{T} \chi_k$. Using the method in the last section, we obtain:

$$
\hat{c}_1 = \frac{\int_{-\infty}^{\infty} (L_T^{(1)}(z) - d_1 L_T^{(2)}(z))dz}{s_T \int_{-\infty}^{\infty} q_T(z)dz}
$$
The functions $L_t^{(i)}$ and $M_t^{(i)}$ are defined as

\[
L_t^{(i)}(z) = D(y_t, z) \left[ \int_{-\infty}^{\infty} \left( \frac{\chi_i \psi_1(z-x)}{\theta_1} + \frac{(1-\chi_i) \psi_2(z-x)}{\theta_2} \right) M_{t-1}^{(i)}(x) \, dx + H_i(y_t) F_i(z) \int_{-\infty}^{\infty} \left( \frac{\chi_i \psi_1(z-x)}{\theta_1} \right) G_t(x) q_{t-1}(x) \, dx \right]
\]

and

\[
M_t^{(i)}(z) = D(y_t, z) \left[ \int_{-\infty}^{\infty} \left( \frac{\chi_i \psi_1(z-x)}{\theta_1} + \frac{(1-\chi_i) \psi_2(z-x)}{\theta_2} \right) M_{t-1}^{(i)}(x) \, dx + H_i(y_t) F_i(z) \int_{-\infty}^{\infty} \left( \frac{(1-\chi_i) \psi_2(z-x)}{\theta_2} \right) G_t(x) q_{t-1}(x) \, dx \right].
\]

where $H_i$, $G_i$ and $F_i$ are given by:
The estimates $\hat{c}_i$, $\hat{d}_i$, $\hat{\theta}_i$ and $\hat{\mu}$, $i = 1, 2$, can be computed recursively using the three steps of the filter-based EM algorithm described below:

**Step 1:** Select the initial values $\hat{c}_i(0)$, $\hat{d}_i(0)$, $\hat{\theta}_i(0)$ and $\hat{\mu}(0)$.

**Step 2:** Compute the MLEs $\hat{c}_i(k + 1)$, $\hat{d}_i(k + 1)$, $\hat{\theta}_i(k + 1)$ and $\hat{\mu}(k + 1)$ by running the filter bank above, where $k$ represents the $k^{th}$ iteration of the algorithm.

**Step 3:** Stop when $|\hat{c}_i(k + 1) - \hat{c}_i(k)| < \epsilon_1$, $|\hat{d}_i(k + 1) - \hat{d}_i(k)| < \epsilon_2$, $|\hat{\theta}_i(k + 1) - \hat{\theta}_i(k)| < \epsilon_3$, $|\hat{\mu}(k + 1) - \hat{\mu}(k)| < \epsilon_4$; otherwise, continue from Step II, where $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ are the desirable levels of accuracy and $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 > 0$.

The sequence of log-likelihoods in the above EM algorithm is increasing and converges. Readers may refer to Dembo and Zeitouni [28] and Zeitouni and Dembo [29] for details about the convergence of the sequence of estimates.

Finally, we estimate the log-volatility

$$\hat{\xi} = \frac{\int_{-\infty}^{\infty} z q_T(z) \, dz}{\int_{-\infty}^{\infty} q_T(z) \, dz}$$

and the hidden volatility

$$\hat{\sigma} = e^{\frac{\hat{\xi}}{2}}.$$
4.4 Special Cases

We now look at some special cases of the threshold autoregression stochastic volatility (TARSV) model.

If \( c = c_1 = c_2, d = d_1 = d_2, \theta_1 = \theta_2 \), the model reduces to the stochastic volatility model in Elliott et al. [15], that is

\[
    x_t = c + dx_{t-1} + \theta w_t .
\]

The filters and densities reduce to exactly the same forms as those derived in Elliott et al. [15].

If \( \mu = 0, c_1 = c_2 = 0, \theta_1 = \theta_2 = 1 \), our model reduces to

\[
    y_t = -e^{2x_t} + e^{x_t}b_t \\
    x_t = \begin{cases} 
        d_1 x_{t-1} + w_{t}^{(1)} & \text{if } y_{t-1} \leq 0, \\
        d_2 x_{t-1} + w_{t}^{(2)} & \text{if } y_{t-1} > 0.
    \end{cases}
\]

This is similar to the TARSV model discussed in Diop and Guegan [30].

5 Extended TARSV Model

In this section, we discuss an extension to the TARSV model we considered before. Let \( E_1, E_2, \ldots, E_N \) be measurable subsets of the real line such that \( E_i \cap E_j = \emptyset \) whenever \( i \neq j \) and \( \bigcup_{k=1}^{N} E_k = \mathbb{R} \). That is, \( \{E_1, E_2, \ldots, E_N\} \) is a “measurable” partition of the real line.

For example, one could take that \( E_1 = (-\infty, e_0], E_2 = (e_0, e_1], \ldots, E_k = (e_{k-1}, e_k], \ldots, E_N = (e_N, \infty) \), where \( e_0 < e_1 < \ldots < e_N \) are real numbers.

We set

\[
    x_t = \begin{cases} 
        c_1 + d_1 x_{t-1} + \theta_1 w_{t}^{(1)} & \text{if } y_{t-1} \in E_1; \\
        c_2 + d_2 x_{t-1} + \theta_2 w_{t}^{(2)} & \text{if } y_{t-1} \in E_2; \\
        \vdots & \text{if } y_{t-1} \in E_{N-1}; \\
        c_N + d_N x_{t-1} + \theta_N w_{t}^{(N)} & \text{if } y_{t-1} \in E_N
    \end{cases}
\]
where \( w_t^{(k)} \) are i.i.d. standard normal random variables for \( k = 1, 2, \ldots, N \).

We now derive the filters and densities.

Let
\[
\chi_{k,t} = \begin{cases} 
1 & y_{t-1} \in E_k \\
0 & y_{t-1} \notin E_k 
\end{cases},
\]
and \( \psi_k \) be the probability density function of \( w_t^{(k)} \). The recursive filter is
\[
q_t(z) = D(y_t, z) \int_{-\infty}^{\infty} q_{t-1}(x) \sum_{k=1}^{N} \chi_{k,t} \frac{\psi_k(z-c_k-d_kx)}{\theta_k} \, dx,
\]
\[
\hat{x}_T = \frac{\int_{-\infty}^{\infty} z q_T(z) \, dz}{\int_{-\infty}^{\infty} q_T(z) \, dz}.
\]

For any measurable functions \( H, F, G \), define
\[
S_t(m, H, F, G) = \sum_{k=1}^{t} \chi_{m,k} H(y_k) F(x_k) G(x_{k-1}).
\]

Suppose there is a density \( L_t^{m,H,F,G}(z) \) such that
\[
L_t^{m,H,F,G}(z) = D(y_t, z) \left[ \int_{-\infty}^{\infty} L_{t-1}^{m,H,F,G}(x) \sum_{k=1}^{N} \chi_{k,t} \frac{\psi_k(z-c_k-d_kx)}{\theta_k} \, dx \\
+ H(y_t) F(z) \int_{-\infty}^{\infty} \frac{\chi_{m,t} \psi_m(z-c_m-d_mx)}{\theta_m} G(x) q_{t-1}(x) \, dx \right].
\]

Write \( L_t^{m,(k)} \) for the density corresponding to the functions \( H, F \) and \( G \) in the \( k \)-th point in the remark of Section 4.1. Define \( s_m^{T} = \sum_{k=1}^{T} \chi_{m,k} \). Then the parameters are estimated as:
\[
\hat{c}_m = \frac{\int_{-\infty}^{\infty} (L_{T}^{m,(1)}(z) - d_m L_{T}^{m,(2)}(z)) \, dz}{s_m^{T} \int_{-\infty}^{\infty} q_T(z) \, dz},
\]
\[
\hat{d}_m = \frac{\int_{-\infty}^{\infty} (L_{T}^{m,(3)}(z) - c_m L_{T}^{m,(2)}(z)) \, dz}{s_m^{T} \int_{-\infty}^{\infty} L_{T}^{m,(4)}(z) \, dz},
\]
\[
\hat{\theta}_m = \frac{\int_{-\infty}^{\infty} (L_{T}^{m,(5)}(z) - 2c_m L_{T}^{m,(1)}(z) + c_m^2 s_m^{T} - 2d_m L_{T}^{m,(3)}(z) + 2c_m d_m L_{T}^{m,(2)}(z) + d_m^2 L_{T}^{m,(4)}(z)) \, dz}{s_m^{T} \int_{-\infty}^{\infty} q_T(z) \, dz},
\]
\[
\hat{\mu} = \frac{\int_{-\infty}^{\infty} \sum_{m=1}^{N} L_{T}^{m,(6)}(z) \, dz + \frac{T}{2} \int_{-\infty}^{\infty} q_T(z) \, dz}{\int_{-\infty}^{\infty} q_T(z) \, dz}.
\]
The generalized TARSV model provides more flexibility to incorporate various asymmetric behaviors in stochastic volatility than the TARSV model. For different financial assets or even different shares in the same market, the number of cut-off and the division points in volatility may be different. This generalised TARSV model allows one to pertain the pros of the original TARSV model while providing more flexibility.

6 A Simulation Study

In this section we provide a simulation study to examine the performance of the proposed filter-based estimation method. Firstly, we consider the case of a fixed volatility level. Then we study the case of a varying volatility level. In both cases, we suppose that the model parameters are given as if they were the “true” model parameters. Simulated data based on models with these “true” parameters are used as if they were the real data to examine the performance of the proposed filter-based estimation method.

6.1 Fixed volatility level

In this case, we fix the level of the volatility $\sigma_t = 0.25$ and simulate three sets of returns data $\{y_t\}$ with 300 data points for each set. We consider the following configurations of the parameters values:

$$\mu = 0.0001; \quad c_1 = 0.2; \quad c_2 = -0.2; \quad d_1 = 0.1; \quad d_2 = 0.15; \quad \theta_1 = 0.1; \quad \theta_2 = -0.05.$$ 

The estimation results of the proposed estimation method and their errors are shown in the following table.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$\hat{\sigma}_{300}$</th>
<th>Relative Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.236265</td>
<td>5.49</td>
</tr>
<tr>
<td>2</td>
<td>0.239480</td>
<td>4.21</td>
</tr>
<tr>
<td>3</td>
<td>0.266071</td>
<td>6.42</td>
</tr>
</tbody>
</table>

We see that the average relative error of the estimation of the volatility in the three cases is around 5%, which is quite low.
6.2 Varying volatility level

In this case we examine the performance of the proposed estimation method when the volatility level is varying. We consider the following configurations of model parameters:

\[ c_1 = -0.5 ; \quad c_2 = -0.3 ; \quad d_1 = 0.1 ; \quad d_2 = 0.15 ; \]
\[ \theta_1 = 0.1 ; \quad \theta_2 = -0.05 ; \quad \mu = 0.0001 ; \quad \sigma_1 = 0.6 \, . \]

We simulate three sets of returns data \( \{y_t\} \) and volatility data \( \{\sigma_t\} \) with 300 data points each.

The filter-based estimates of the volatility are given in the following table.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>( \sigma_{300} )</th>
<th>( \hat{\sigma}_{300} )</th>
<th>Relative Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.565693</td>
<td>0.606538</td>
<td>7.22</td>
</tr>
<tr>
<td>2</td>
<td>0.621370</td>
<td>0.606507</td>
<td>2.39</td>
</tr>
<tr>
<td>3</td>
<td>0.667094</td>
<td>0.670320</td>
<td>0.48</td>
</tr>
</tbody>
</table>

We can see that the relative errors of the estimation in the three cases are low. This illustrates that the performance of the proposed method is, at least, at an acceptable level.

7 Conclusion

We have derived a nonlinear recursive filter for the hidden volatility and filter-based estimates for model parameters in a TARSV model based on the EM algorithm. (Semi)-analytical formulae filters and estimates are derived. Further, we illustrated that our model can be thought of as an extension to two SV models, and consequently, similar filtering and filter-based estimation methods can be applied. Lastly, an extension to the model was given, and we outlined the derivation of the filters in the extended model. We examined the proposed estimation method based on simulations. The results revealed that the performance of the proposed method is, at least, at an acceptable level.

Acknowledgments: The authors would like to thank the referee for helpful
Robert J. Elliott and Tak Kuen Siu wish to acknowledge the Discovery Grant from the Australian Research Council (ARC), (Project No.: DP1096243). Robert J. Elliott would also like to thank the ARC for support through grant DP0877639.

References


